

DURABLE RESEARCH CONCEPT

20 Watts Was Enough

A biologically inspired R&D blueprint for sparse, grounded, continual, energy-efficient AI.

EDITION

PDF v0.3.0 · main snapshot

CONTENTS

20 concept documents + 30 mathematical notes + 1 generated appendix

LENGTH

145,594 words

SOURCE

Git main snapshot

[SUPPORT](#) [FOR MAIN](#) [How to help](#) [Report this edition](#)

Contents

00	Research readiness
CONCEPT	
01	20 Watts Was Enough
02	Thesis and design principles
03	Working architecture: three coupled loops
04	Biology is a launchpad, not a ceiling
05	Cross-domain convergence and principle deduplication
06	Structural growth, specialization, and conditional routing
07	Multimodal sensorimotor grounding
08	Representative adaptive performance
09	Active acoustic inference
10	Operator-qualified sensing and physical inference
11	Active chemical sensing
12	Reliability under mission profiles
13	Physical computation requires six boundaries
14	Sparse prediction and adaptive compute
15	Fast memory, replay, and consolidation
16	Maturity, grokking, and reversible structural consolidation
17	Hardening, reflex paths, and factual memory
18	System synthesis
19	Energy model and efficiency evaluation contract
20	Research roadmap
MATHEMATICAL NOTES	
M01	Boundary-qualified physical-computation contract
M02	Burden-qualified contestable decisions
M03	First testable efficiency model
M04	History-conditioned modular succession: mathematical contract
M05	Interface-qualified retroactivity and insulation
M06	Interface-qualified scale symmetry

M07	Intervention-qualified mechanism equivalence
M08	Horizon-qualified learning outcomes
M09	Material commitment and realized service
M10	Measurement contracts, uncertainty, and invalidation
M11	Budgeted memory lifecycle
M12	Mission-profile-qualified device reliability contract
M13	Multiscale reduction contract
M14	Notation and units
M15	Occupancy-qualified spatial transition
M16	Operator-qualified active acoustic inference
M17	Operator-qualified active chemical sensing
M18	Operator-qualified optical inference contract
M19	Opportunity- and history-qualified adaptive action
M20	Cohort-aware population and historical observation
M21	Predictive temporal co-adaptation without a privileged clock
M22	Proof-discovery and verification contract
M23	Query-registered semantic preservation
M24	Regime-qualified flow inference and control contract
M25	One-sided topology signatures under repeated stimulation
M26	Representative adaptive-performance contract
M27	State-qualified intervention and withdrawal
M28	Typed tolerance, response, and recovery lifecycle
M29	Versioned reconstructive design contract
M30	Visual models

APPENDIX

A1	Global field coverage

GENERATED RESEARCH STATUS

Scientific claims are not runnable tests

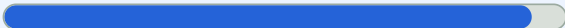
Evidence status describes support for a source-domain claim. Test readiness describes whether this project can execute a frozen experiment. The two axes are deliberately reported separately.

CLAIM DENOMINATOR: 1,571

Written protocols versus executable evidence

1,475 of 1,571

Protocol-covered claims



At least one linked candidate or fixture contains all eight required written protocol facets. This is design coverage, not an experimental result.

0 of 1,571

Workstation-executable claims



A linked artifact must have a checked workstation-ready manifest and a runnable scientific harness. Smoke plumbing cannot promote a claim.

ARTIFACT DENOMINATOR: 49

Experiment implementation lane

49

COMPLETE WRITTEN
PROTOCOLS

11

VALIDATED SMOKE HARNESS

0

WORKSTATION-READY
EXPERIMENTS

INDEPENDENT AXES

Evidence status by highest test tier

HIGHEST TEST TIER	ESTABLISHED	PLAUSIBLE	SPECULATIVE	DISPUTED	UNKNOWN	TOTAL
Ledger only	83	9	0	4	0	96
Linked description	0	0	0	0	0	0
Protocol complete	1,145	210	64	47	9	1,475
Workstation executable	0	0	0	0	0	0

"Established" means the cited observation is established within its stated scope. It does not mean the proposed AI translation has passed this project's experiment.

NOT MISSING AT RANDOM

96 ledger-only claims

Evidence inputs

85

Source reproductions

11

New experiment needed

0

No ledger-only record currently requires a new experiment family. The remaining records are evidence inputs or source-domain reproductions and are not falsely counted as project tests.

[Inspect the disposition record](#)

NEAREST EXECUTABLE TARGET

candidate-010: smoke-ready

Candidate 010: reset-coupled staged verification

6/9 machine-checkable structural promotion gates currently pass



Explicit execution-track claim scope

execution track is limited to C-170



Frozen full profile

data.full_profile schema 1, development identity, opportunities_per_seed workload cardinality, direct SHA-256, and any declared embedded seed/config hashes match repository files



Full experiment-path tests

implementation.full_tests is part of the workstation suite executed by npm test



Corruption-evident output ledger

hash-chain output is wired through the tested checkpoint ledger



Deterministic resume

runner imports the declared, tested checkpoint/resume implementation



Measured-energy or explicit non-energy boundary

external-meter provider capability, exclusions, test, and frozen config hash match



Frozen confirmation seeds

seeds.confirmation is explicitly pending the implementation/pilot freeze and has no sealed reveal



Frozen held-out seeds

seeds.held_out is explicitly pending the implementation/pilot freeze and has no sealed reveal



Strict recomputed promotion evidence

strict promotion evidence is pending and no claim-eligible bundle exists

[Open the experiment contract](#)

[Open the harness notes](#)

The public Git repository contains the complete artifact table and generated coverage report.

20 Watts Was Enough

CI

GitHub Pages

OpenSSF Scorecard

Latest release

Code: EUPL 1.2

Original content: CC BY-SA 4.0

A biologically inspired R&D blueprint for sparse, grounded, continual, energy-efficient AI.

The adult human brain operates within a whole-organ power budget of roughly 17–20 watts (C-001). That observation does not provide a fair brain-versus-silicon benchmark by itself; it establishes a serious engineering question: which computational constraints behind adaptive biological intelligence can be translated into measurable requirements for artificial systems? The [energy evaluation chapter](#) defines the boundary required for a comparison.

This repository turns that question into a traceable research programme: primary evidence becomes scoped claims, cross-domain mechanisms are deduplicated, and retained ideas face explicit null models and failure rules.

Current status: the evidence framework, architecture and bounded experiment machinery exist. An integrated AI system and claim-eligible workstation result do not. Development smoke checks remain `NO_RESULT`. The generated [coverage report](#) and [readiness summary](#) record the exact current boundary.

Read the project

GOAL	START HERE
Understand the question and proposed system	Working architecture
Browse the research online	Research portal
Read continuously	Full HTML book
Read offline or print	Download the A4 PDF
Inspect evidence	Claim ledger and bibliography
Inspect the plan and proposed tests	Research roadmap and experiment coverage
Run a released experiment image	Container execution contract
Help with a bounded task	Public contribution map
Find a specific area	Repository map

The portal and web book list all 24 official EU languages, but publish only translations tied to an exact canonical source digest and human review. English remains canonical. Where a reviewed translation does not yet exist, the language control opens a focused contribution issue instead of unlabelled machine output. See the [translation contract](#).

What the project tests

Useful adaptive intelligence may require much less active computation and data movement than current systems routinely spend. The project investigates whether that gap can be attacked through a system that:

- separates total capacity from active capacity;
- grounds representations in aligned perception, action, and outcome;
- allocates computation according to uncertainty and task demand;
- separates rapid episodic learning from slow structural learning;
- consolidates before pruning or hardening;
- keeps mutable facts in inspectable memory rather than forcing everything into weights; and
- measures quality, energy, data movement, lifecycle cost, and uncertainty together.

Nature is a source of mechanisms, not a shortcut around engineering controls. Observations from neuroscience, biology, ecology, physics, chemistry, mathematics, social systems, and other fields are normalized by their causal operation, deduplicated into shared principle bundles, and tested against strong ordinary engineering alternatives.

Plan and progress

The [research roadmap](#) is the canonical dependency order. Its stages run from evidence and experiment contracts through isolated mechanisms, adaptation, grounding, structural maturation and substrate co-design. They are gates, not promised dates.

[GitHub milestones](#) project those stages into issues and pull requests. They answer *what work is open and who is doing it*; they do not decide whether a scientific claim is established. The authorities remain separate:

QUESTION	AUTHORITY
What should happen next?	Research roadmap
Which bounded tasks are open or finished?	GitHub milestones projected from the milestone manifest and repository-bound issue map
What does the evidence support?	Claim ledger
Which experiment paths are actually ready?	Coverage report and readiness summary

Closing an issue updates operational progress. Only a reviewed evidence change updates a claim, and only a qualifying run updates result status.

The default normative context is the European Union and Germany. Legal, standards, and conformity statements remain qualified by jurisdiction, role, version, applicability, and date under the [normative baseline](#).

Scientific conduct follows the [research integrity baseline](#), which deduplicates the recurring ALLEA, EU, DFG, and large European university rules for disclosure, stewardship, ethics, correction, and review.

How the repository works

```
primary sources -> scoped claims -> deduplicated principles
                  -> engineering translations -> equations and architecture
                  -> equal-budget tests -> bounded evidence -> revised claims
```

- `research/claims.md` gives major assertions stable `C-` identities and evidence status.
- `research/principle-registry.md` groups equivalent cross-domain mechanisms under stable `P-` identities.
- `concept/` develops the maintained system synthesis.
- `math/` defines notation, units, derivations, and testable efficiency models.
- `experiments/` separates written protocols, development plumbing, and result authority.
- `decisions/` records durable choices without rewriting their history.

- `sources/` preserves provenance and publication boundaries; imported discussions and summaries are leads, not evidence.

Canonical source and editing

Git `main` is the canonical source. The Pages portal, HTML book, and PDF are generated views of the same committed material; Google Docs and imported chats are not synchronized document stores. Work incrementally on one chapter, claim, equation, diagram, test, or decision rather than regenerating the whole project.

Read `AGENTS.md` and the [engineering and research contract](#) before editing. The [research-writing skill](#) keeps project prose direct and evidence-aware, with a deliberately narrow automated check for high-confidence generated filler. The [repository-rule crosswalk](#) explains which conventions from the maintainer's other projects were adopted, adapted, staged, or rejected here.

For a local live preview:

```
npm ci
npm run dev:github-pages
```

The preview runs at `http://localhost:5173/` and reloads canonical Markdown, equations, tables, diagrams, and plots as their source files change.

Qualifying releases publish repository tooling and individual experiments as small container images. Use the complete `image@sha256:...` identity from the release's checksum-bound `oci-images.json`, not a mutable tag. The [release contract](#) explains admission, provenance and supported platforms; the [experiment instructions](#) provide the exact isolated commands. For example, the static validation image runs without network access:

```
image='ghcr.io/lusoris/20-watts-was-enough-20w@sha256:...'
docker pull "$image"
docker run --rm --network none \
  --pull never \
  --mount "type=bind,src=$PWD,dst=/repo,readonly" \
  "$image" \
  validate docs --root /repo
```

Releases may also attach a verified native Go binary as a convenience. The containers are the portable public default and remain scoped per experiment.

Before a local commit, follow the [validation sequence](#): check the changed contracts and their downstream consumers, combining scopes for mixed changes. Unknown, unsafe or shared-authority scope requires the full gate.

Run the complete `npm` and Go gates before marking a pull request ready, integrating into `main`, merging or releasing.

Changes to book sources also require:

```
npm run generate:book-pdf
npm run validate:book-pdf
```

See `CONTRIBUTING.md` for the complete workflow and `SUPPORT.md` for issue routing. The repository does not maintain a PowerShell, C# or host-specific execution lane.

Citation, support, and licence

Use `CITATION.cff` and identify the exact release or commit when citing the project. Scientific use should also cite the original sources that support the specific claim; citing this synthesis does not replace them.

If the project is useful, support is available through [Ko-fi](#) or [GitHub Sponsors](#).

Project-authored software, scripts, tests, configurations, schemas, and workflows are licensed under the EUPL v1.2 or later. Original project prose, mathematics, diagrams, plots, and presentation are licensed under CC BY-SA 4.0. Third-party and source material retains its own terms. Read [LICENSING.md](#) for the controlling boundary.

Thesis and design principles

Scope

This chapter states the project's argument, its engineering requirements, and the limits that keep a biological analogy from becoming mythology.

Thesis

The brain demonstrates that adaptive behavior, continual learning, perception, memory, and action can coexist under a tight energy and communication budget (C-001). Contemporary AI demonstrates a different strength: highly parallel optimization over vast data and parameter spaces.

The project hypothesis is that a capable artificial system should combine the search capacity of modern learning with constraints that biology cannot evade:

1. only a small, relevant fraction of capacity should be active for an event;
2. communication and memory movement must be priced, not hidden behind FLOPs;
3. learning must begin from aligned perception, action, and consequence rather than language alone;
4. rapid acquisition must not directly rewrite stable long-term structure;
5. consolidation must precede destructive compression; and
6. stable skills, mutable facts, and active reasoning should use different storage and execution paths.

The intended result is a system whose cost scales primarily with the information and uncertainty relevant to a task, rather than with its full stored capacity.

Biological observation

Neural signaling is metabolically costly, cortical activity is constrained, and biological learning spans multiple timescales. Sensory and motor experience precedes mature language. These observations motivate design constraints.

They do **not** establish a single brain algorithm, an equivalent number of digital operations, or a guarantee that every biological mechanism is efficient on silicon.

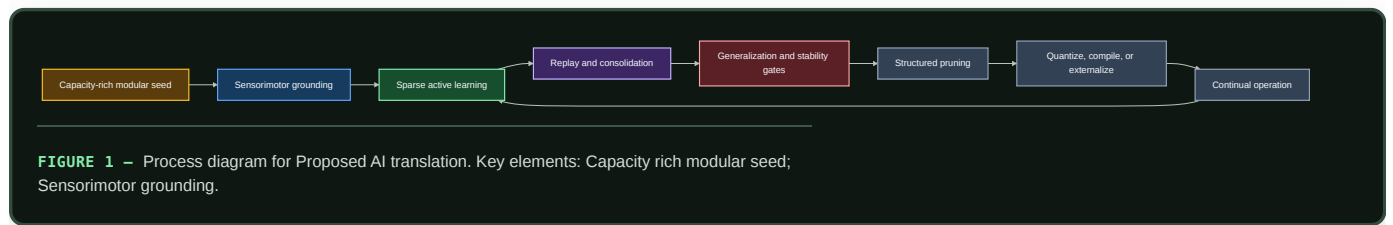
Proposed AI translation

The blueprint combines four separations:

- **capacity versus activity:** sparse routers select modules, depth, memory, and precision;

- **prediction versus correction:** predictable state follows a cheap path; residual uncertainty receives more computation;
- **experience versus structure:** fast memory captures episodes while slow learning integrates regularities under regression tests; and
- **reasoning versus knowledge:** stable skills may be compiled or quantized, while mutable facts remain retrievable and attributable.

Development is a controlled lifecycle:



Editable source: `../assets/diagrams/developmental-pipeline.mmd`.

Efficiency mechanism

The optimization target is constrained quality per unit of total system energy, with latency and uncertainty limits. A model is not efficient merely because it has sparse weights; the runtime must avoid loading and operating on inactive state, and routing overhead must remain smaller than the work skipped.

The top-level objective is developed in the energy model:

$$\max \frac{Q}{E_{\text{facility}}} \quad \text{subject to} \quad L_{p95} \leq L_{\max}, R \leq R_{\max}, Q \geq Q_{\min}.$$

Here Q is task quality, E_{facility} is measured facility energy, L_{p95} is tail latency, and R is a defined risk or calibration metric.

Evidence status

- Metabolic constraints on neural activity are established within the scope of C-001.
- Conditional computation is established in narrower engineered systems under C-003 and C-004.
- The complete developmental pipeline is a speculative integration. No cited paper validates it as one system.

Speculative extensions

- Learn a global energy price that modules must “bid” against when requesting compute.
- Let stable routing motifs become physically colocated or compiled for lower data-movement cost.
- Allow controlled structural growth when no existing module can absorb a new regime without interference.

Failure modes

- Anthropomorphic labels conceal incompatible mechanisms.
- Sparsity saves theoretical FLOPs but increases real communication or latency.
- Energy optimization suppresses rare but important computation.
- Early specialization prevents transfer and produces brittle modality silos.
- “Hardening” turns uncertain or mutable behavior into an uncorrectable path.

Measurable predictions

The integrated hypothesis survives only if, at matched quality:

- active parameters and memory traffic grow more slowly than total capacity;
- average energy falls without unacceptable tail-risk or calibration loss;
- sequential learning retains prior capability better than a matched single-timescale baseline;
and
- grounded intervention tasks improve beyond gains explained by additional data or parameters alone.

Working architecture: three coupled loops

A capable system should execute, adapt, and consolidate on different timescales while one controller prices energy, latency, and risk.

Scope

This is the shortest complete description of the proposed system. It joins the project's individual mechanisms into one operating architecture and shows where information moves, where learning happens, and which decisions remain reversible.

Architecture at a glance

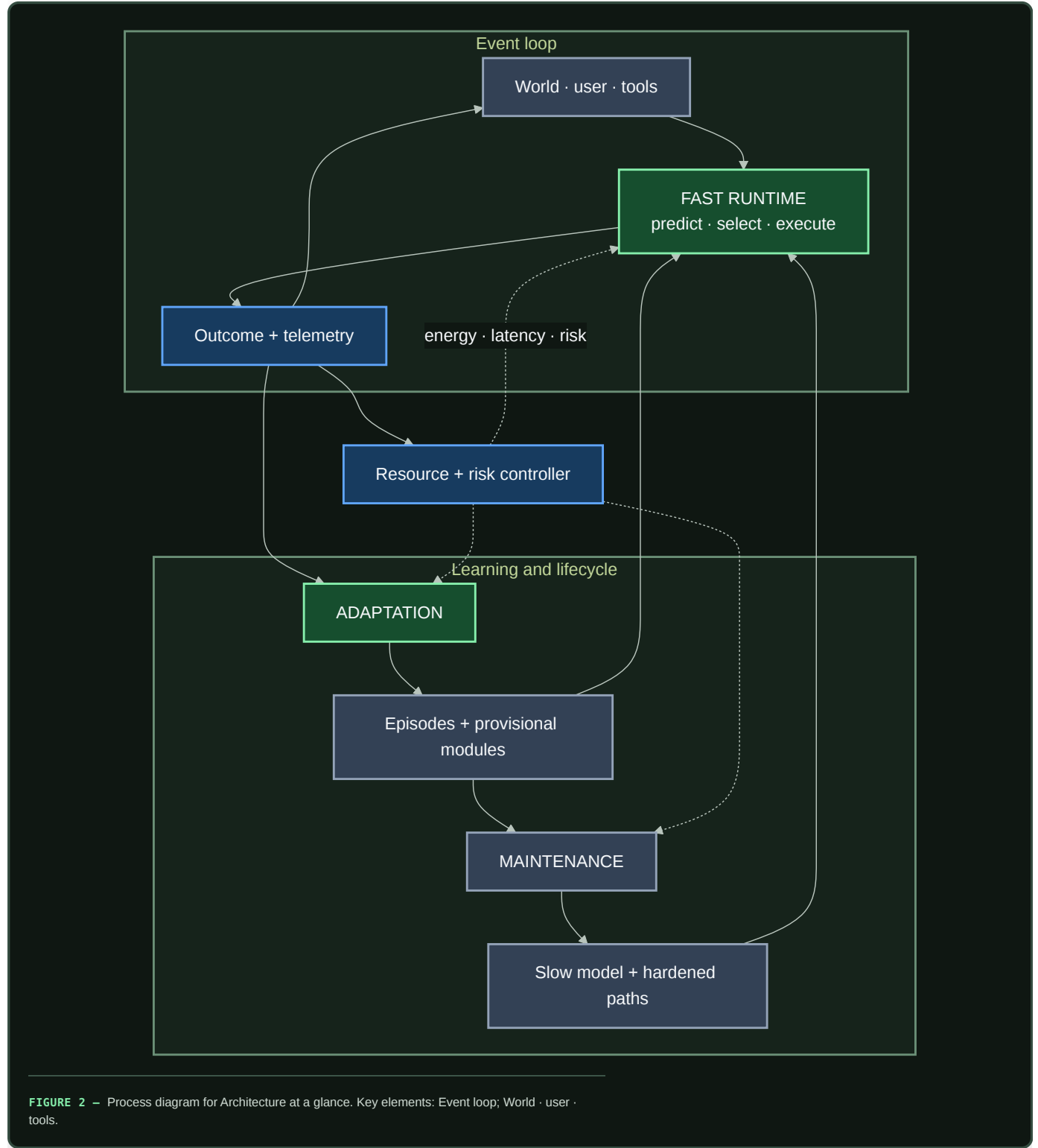


FIGURE 2 – Process diagram for Architecture at a glance. Key elements: Event loop; World · user · tools.

Editable source: [../assets/diagrams/three-coupled-loops.mmd](#).

The three loops solve different problems:

LOOP	TIMESCALE	PRIMARY DECISION	NORMAL OUTPUT
Fast runtime	event	What should run now?	action, answer, or escalation
Adaptation	episode to session	What did this outcome teach us?	attributable episode, hypothesis, or provisional module

LOOP	TIMESCALE	PRIMARY DECISION	NORMAL OUTPUT
Maintenance	replay to release cycle	What should become stable, move, weaken, or disappear?	consolidated memory and restructured execution

The resource controller crosses all three. It can restrict an action because of energy, latency, communication, or risk, but it cannot silently replace the task objective. Measured outcomes return to the controller so estimated costs can be corrected.

Biological observation

Living systems repeatedly separate processes by timescale and locality:

- many components are available, while only a small subset is strongly active together (C-001, C-003);
- local circuits resolve routine events while larger systems coordinate exceptions (C-017, C-024);
- fast traces change behavior before slow structure is rewritten (C-008, C-019);
- replay, forgetting, and structural maintenance are selective processes (C-036–C-042);
- stable structure can be protected and later reopened under bounded intervention (C-043–C-045); and
- local demand can recruit adjacent resource supply (C-049–C-051).

The shared principle is separation of concerns: execution, adaptation, memory, structure, and physical supply change at different speeds and respond to different signals. Digital systems add exact copying, external stores, checkpoints, shadow evaluation, and rollback. Those capabilities make the separations easier to test and safer to revise.

Proposed AI translation

1. Fast runtime

For each event, the runtime forms a predictive state containing expected observations, uncertainty, current goals, and relevant recent context. A hierarchical selector then chooses the cheapest admissible path:

- a hardened low-cost path for a familiar, low-risk event;
- one or more conditionally active experts;
- another layer or reasoning step;
- episodic or factual retrieval;
- a sensor, tool, or environment intervention; or
- escalation to a human or a more capable controller.

The output is more than an answer. It includes measured latency, energy, communication, calibration, and any observed consequence. That telemetry is the input to later learning and resource decisions.

2. Adaptation

An outcome is captured with provenance before it can alter slow parameters. It may update temporary state, create a provisional adapter, change a local connection, or simply become a replay candidate. A surprising event is not automatically a lesson; it may be noise, attack, contradiction, or evidence of a new regime.

There is no ex-nihilo generator in this design. Every proposal is built from inherited structure and experience: complete patterns, fragments, relations, or abstractions are copied and transformed. The generative cycle is therefore explicit:

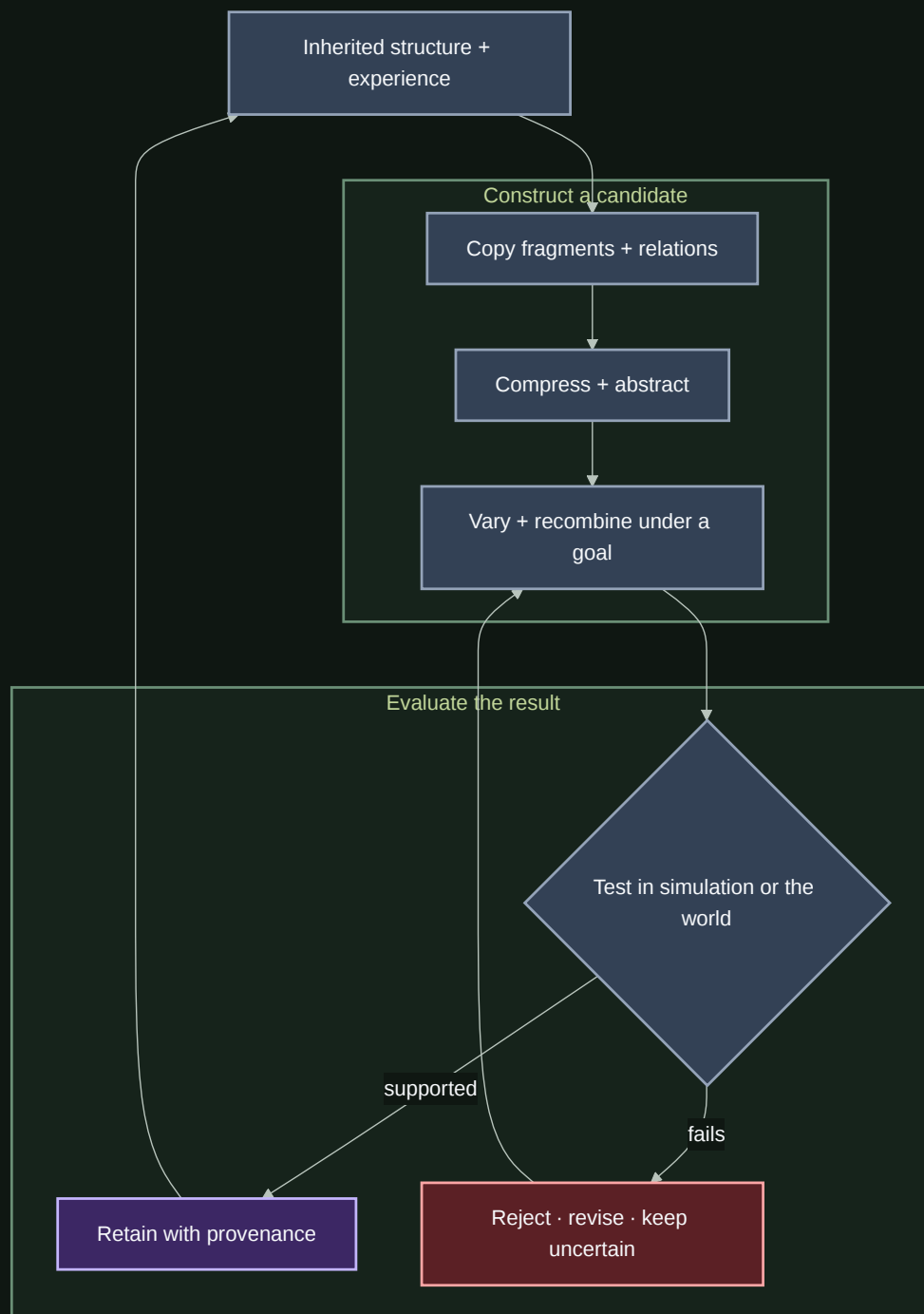


FIGURE 3 – Process diagram for 2. Adaptation. Key elements: Inherited structure + experience; Construct a candidate.

Editable source: [../assets/diagrams/generative-recombination.mmd](#).

Stochasticity can alter which variants are tried. Usefulness comes from the entire cycle: decomposition, recombination, intervention, comparison, and selection. The supporting evidence and unresolved distinctions are tracked in the endogenous generation audit and is falsified first through Candidate 004.

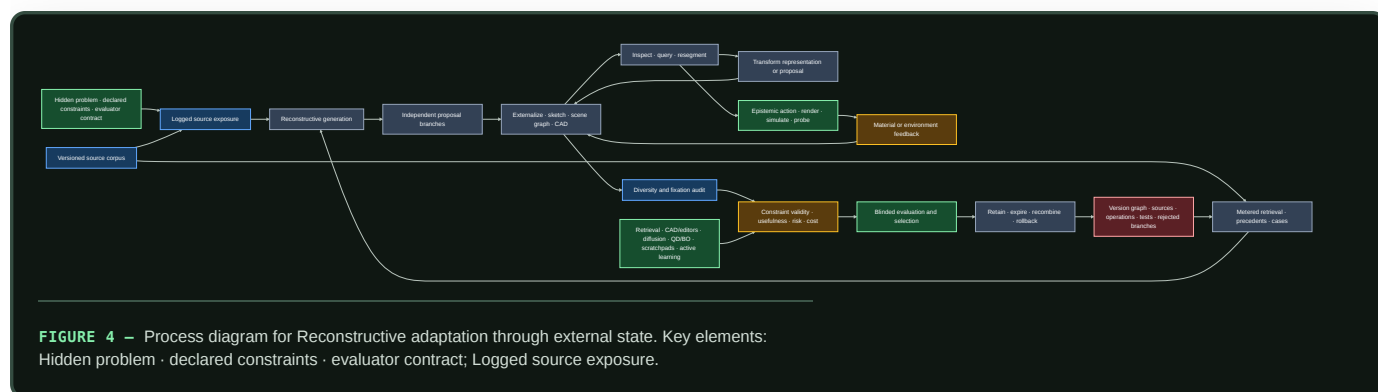
RECONSTRUCTIVE ADAPTATION THROUGH EXTERNAL STATE

Visual art and design cognition make the cycle more operational. Proposal generation is reconstructive: learned categories, exemplars, relations, styles, constraints, tools, and prior transformations shape what can be proposed (C-842–C-860). That does not reduce the process to nearest-neighbor retrieval. A system can externalize a partial state, inspect consequences that

were not explicit in its preceding plan, change representation, probe the world, and use the result to create a new branch. The novelty is relative to a declared history; the validity comes from constraints and observed outcomes.

The adaptation loop therefore distinguishes seven event types:

1. **Exposure:** which sources, examples, styles, and constraints were available before the task;
2. **Retrieval:** which versioned sources were deliberately requested during the task and at what cost;
3. **Reconstruction:** which fragments, relations, or abstractions formed a proposal branch;
4. **Externalization:** which editable representation—text, sketch, scene graph, CAD state, program, simulation, or prototype—made the proposal inspectable;
5. **Epistemic action:** which render, measurement, simulation, query, or material intervention was selected to reduce a decision-relevant uncertainty;
6. **Evaluation and selection:** which hard constraints, qualified evaluators, risk limits, and selection rule admitted or rejected a version; and
7. **Retention:** which source, parent, operation, test, rejected branch, dependency, expiry, and rollback records were preserved.



Editable source: [versioned-reconstructive-design.mmd](#).

Several separations are non-negotiable. More proposals are not necessarily better proposals; proposal diversity, constraint validity, selection regret, and realized outcome are different measurements. Similarity to an example can transfer a useful structural relation or cause fixation and negative transfer, so similarity and task effect are both reported. A provenance graph can record derivation without proving intention, authenticity, correctness, or originality. Aesthetic response is observer-, context-, expertise-, and history-qualified rather than a universal scalar.

Parallel branches are useful only if they remain genuinely independent before critique. A shared canvas can coordinate work and simultaneously propagate an early fixation. The comparison therefore includes private branches, delayed reveal, source substitution, example removal, evaluator swap, material swap, and representation conversion. Every alternative pays for retrieval, proposals, tools, simulations, human critique, physical material, waste, retention, and recovery.

Fixture F-002 tests this complete loop against a mature composition of retrieval, CAD and editors, generative sampling, explicit search and quality-diversity, scratchpads and shared workspaces, active learning, simulation, version control, provenance, and independent reranking. The mathematical contract keeps exposure, retrieval, relative novelty, reinterpretation, information gain, fixation, negative transfer, constraint margins, evaluator qualification, selection regret, lineage, reconstructability, and lifecycle cost separate. If that ordinary stack ties, the additional architecture is removed while the fixture remains as a negative result.

3. Maintenance

Maintenance receives a budgeted set of possible actions. For any memory, module, or route it may:

- defer while collecting more evidence;
- replay against related and conflicting cases;
- merge redundant records while retaining provenance;
- externalize mutable propositions to an inspectable store;
- reduce influence or plasticity;
- delete state after retention and safety checks;
- reopen mature structure in a reversible branch;
- change logical topology or physical placement; or
- promote stable computation into a cheaper representation.

The formal action model is defined in `../math/memory-lifecycle.md`. Structural changes remain shadowed and reversible until regression, calibration, and energy tests pass.

4. One event end to end

1. A sensor, user, or tool changes the current state.
2. The runtime predicts what matters and estimates uncertainty.
3. The selector chooses depth, modules, memory, precision, and possible intervention under the current budget.
4. Execution produces an outcome and physical telemetry.
5. The adaptation loop stores an attributable episode and may construct a bounded hypothesis or provisional module.
6. The maintenance loop later replays the episode against related and conflicting evidence.
7. A validated regularity may be consolidated, compiled, quantized, externalized, relocated, or used to change future routing.
8. Regression or fragility tests can reject the change and restore the previous state.

5. State ownership

STATE	PRIMARY OWNER	NORMAL WRITE PATH	REASON FOR SEPARATION
Current predictive state	runtime	every event	cheap, transient, task-specific
Attributable episodes	adaptation	observed outcome	preserves evidence before abstraction
Provisional hypotheses	adaptation	bounded generation and trials	permits novelty without global drift
Reusable skills and representations	slow model	validated consolidation	stable transfer across events
Mutable propositions	factual memory	sourced, versioned update	correction, attribution, and conflict
Hardened execution paths	compiled store	promotion pipeline	lower repeated interpretation cost
Lifecycle and fragility state	maintenance	replay, probes, regressions	repair remains separate from execution
Resource and risk policy	controller	calibrated policy update	local mechanisms cannot hide physical cost

Efficiency mechanism

Efficiency is an event-level constrained decision. At time t , the selector chooses an admissible action a from $\mathcal{A}_t$:

$$a_t^* = \arg \max_{a \in \mathcal{A}_t} \mathbb{E}[\Delta V(a) \mid s_t]$$

subject to

$$E(a) \leq B_t^E, \quad L(a) \leq B_t^L, \quad R(a) \leq \epsilon_t^R.$$

Here st is the predictive state; $\Delta V(a)$ is expected task-value improvement; $E(a)$ is energy in joules; $L(a)$ is latency in seconds; $R(a)$ is the declared risk measure; and B_t^E , B_t^L , and ϵ_t^R are event-specific limits. Measured cost after execution recalibrates the estimator.

Lifecycle comparison includes the work usually hidden outside inference:

$$E_{\text{life}} = E_{\text{run}} + E_{\text{memory}} + E_{\text{network}} + E_{\text{adapt}} + E_{\text{maint}} + E_{\text{recovery}}.$$

An architecture advances only if it improves the joint frontier of quality, risk, latency, and lifecycle energy against its strongest conventional baseline. Skipped FLOPs alone are not an efficiency result.

Evidence status

COMPONENT	CURRENT EVIDENCE	TRANSLATION STATUS
Conditional experts and early exit	C-003, C-004	implemented mechanisms; physical benefit remains system-specific
Predictive residual allocation	C-005–C-007, C-022	plausible composition; grounded experiment required
Fast/slow memory and replay	C-008–C-010, C-036–C-042	scoped evidence; lifecycle controller experimental
Structural maturation and reopening	C-012, C-043–C-045	biological interventions established; digital state machine speculative
External factual memory	C-014	usable mechanism; conflict and energy policy unresolved
Local resource and context control	C-046–C-051	biological observations established; proposed control planes experimental
Endogenous proposal generation	C-061–C-066	constituent observations established; integrated curriculum speculative
Versioned reconstructive design	C-842–C-860	constituent observations scoped; complete externalize–inspect–transform–evaluate loop unvalidated
Complete three-loop system	none	unvalidated synthesis

The engineering analogue audit maps the architecture to feedback control, value of information, caches, schedulers, change detection, adaptive routing, replicated state, and system identification. Each experiment must compare against the strongest relevant version of those methods.

Speculative extensions

- Receiver-specific fast and slow context channels under a hard bit budget.
- Modules bidding for compute by expected value improvement while a global controller enforces physical and safety constraints.
- Active fragility probes on shadows or replicas before topology changes.
- Capability-gap repair that admits complementary modules instead of restoring arbitrary capacity.
- Joint learning of logical routing, memory placement, and hardware locality.

Each remains outside the core architecture until its isolated contract beats the appropriate engineering baseline.

Failure modes

- **Dense control disguised as sparsity:** routing, broadcast, or monitoring moves nearly as much state as dense execution.
- **Controller conflict:** runtime, adaptation, maintenance, and resource loops oscillate or optimize incompatible signals.

- **False maturity:** a shortcut appears stable and is hardened before compositional and intervention tests.
- **Memory laundering:** unsupported episodes become slow parameters without provenance or contradiction checks.
- **Maintenance dominates cost:** replay, probes, migrations, and regression testing consume the savings of conditional execution.
- **Fixation propagates through shared state:** early examples or collaborators collapse independent branches before alternatives are tested.
- **Generation impersonates evaluation:** the proposal model's own score is reported as independent evidence of validity or usefulness.
- **Lineage impersonates truth:** a complete derivation record is treated as proof of correctness, intention, authenticity, or originality.
- **Renaming without mechanism:** a familiar cache, controller, or scheduler receives a biological label without a different causal operation.

Measurable predictions

The architecture earns implementation only if isolated experiments support the following predictions:

1. Conditional execution reduces measured data movement and wall energy at matched quality, calibration, and tail risk.
2. Fast attributable memory improves adaptation latency while slow-model regression remains bounded.
3. Reversible maturity gates reduce catastrophic drift without blocking necessary relearning or newcomer admission.
4. Structured recombination plus targeted intervention produces more valid, useful novelty than matched-budget stochastic sampling alone.
5. Editable externalization, inspection, representation change, and epistemic action improve hidden-constraint validity, selection regret, and transfer beyond the complete retrieval/editor/generation/search/tool stack without increasing fixation or lifecycle cost.
6. Maintenance improves retention and adaptability after replay, monitoring, migration, and recovery costs are counted.
7. At least one proposed mechanism beats the strongest ordinary controller, scheduler, cache, router, or system-identification baseline.

Until those tests pass, this is a working architecture: precise enough to reject, modular enough to simplify, and explicitly incomplete.

Biology is a launchpad, not a ceiling

Scope

This chapter defines how the project may borrow from living systems without turning a biological implementation into a specification. The brain is an existence proof for efficient adaptive intelligence, but neither evolution nor neural tissue optimized the same objective, hardware, scale, or reliability contract as an artificial system.

The working rule is:

Reproduce the useful constraint or computation, not the substrate accident.

Biological observation

Biological intelligence is shaped by slow, noisy, failure-prone components; metabolically expensive communication; physical growth; local chemical signals; and an inability to checkpoint, clone, or roll back a whole organism. Brains compensate with local processing, sparse activity, redundancy, specialized cell types, multiple timescales, and continual maintenance.

Those compensations are informative, but they are not automatically optimal on silicon. Digital systems offer different capabilities: fast switching and interconnect, exact copying, explicit addressing, external storage, reversible experiments, global synchronization when it is worth its cost, and precision that can be selected per operation. At the same time, digital systems still pay heavily for data movement, memory access, communication, cooling, and idle capacity. The relevant question is which constraints and useful computations survive a change of substrate.

The speed comparison must stay qualitative until a shared task and latency boundary exist. Axonal conduction, synaptic integration, transistor switching, GPU kernels, and cluster collectives are different operations. Selecting the fastest number from each domain would recreate the invalid comparison rejected by C-016.

Proposed AI translation

For each biological candidate, record five transformations:

- 1. **Observed function:** what the organism demonstrably achieves.
- 2. **Biological constraint:** which physical or evolutionary limit shaped it.
- 3. **Candidate invariant:** the computation that may survive a substrate change.
- 4. **Silicon implementation:** the least literal engineered mechanism that tests the invariant.
- 5. **Escape hatch:** the biological limitation that engineering should not inherit.

BIOLOGICAL PATTERN	CANDIDATE INVARIANT	SILICON-NATIVE ESCAPE HATCH
Mostly local signaling	Price communication and keep repeated work near its state	Permit fast global exchange when its measured value exceeds its traffic cost

BIOLOGICAL PATTERN	CANDIDATE INVARIANT	SILICON-NATIVE ESCAPE HATCH
Slow, noisy spikes	Event-driven, uncertainty-sensitive updates	Use dense vector arithmetic or exact digital state where it is cheaper
Synaptic and dendritic computation	Compute near stored state; route locally before global aggregation	Implement fused kernels, hierarchical memory, or programmable modules rather than literal morphology
Sleep and replay	Separate acquisition from protected integration	Consolidate asynchronously, continuously, or from exact checkpoints
Development and pruning	Explore with reversible capacity, then commit after evidence	Grow, clone, roll back, and reallocate modules without waiting for physical development
Multiple memory systems	Match update rate and provenance to information lifetime	Use databases, caches, logs, tools, and versioned weights unavailable to animals
Homeostasis and repair	Treat stability and maintenance as active control loops	Use telemetry, deterministic tests, redundancy, and replacement hardware

This transformation is a project decision, not evidence that any row will produce a gain. The candidate inventory is maintained in the [neuroscience opportunity map](#) and [comparative-biology map](#). Its relationship to existing AI is tracked in the [adoption matrix](#).

Efficiency mechanism

The approach preserves useful functions in the cheapest available form and uses digital capabilities—including exact copies, reversible branches, and direct addressing—where they improve the measured system.

For a biological candidate m , the engineering experiment compares at least three systems: a conventional baseline B , a literal or close biological translation L_m when meaningful, and a silicon-native abstraction S_m . Using the measurement contract from the [energy model](#), the candidate is interesting only if S_m improves the quality–risk–energy frontier:

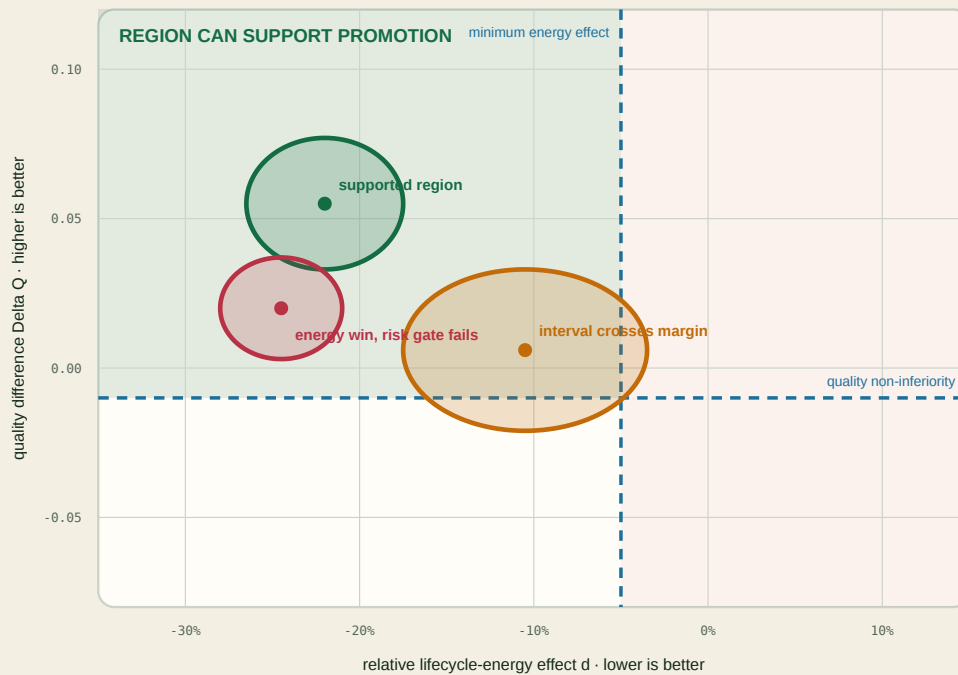
$$(Q, R, E, L)_{S_m} \succ (Q, R, E, L)_B,$$

where Q is task quality, R is the declared risk metric, E is energy within the declared boundary, and L is latency. The symbol $\succ$ means Pareto dominance under pre-registered tolerances; it does not collapse unlike units into a decorative score.

The decision geometry below is illustrative. Its points and simultaneous regions are not measurements; they show why a favorable mean on one coordinate cannot override uncertainty or a failed protected-outcome gate.

A mean improvement is not a promotion decision

$$d = (e_C - e_B) / e_B; \Delta Q = Q_C - Q_B$$



HARD GATES

latency risk support

supported region



interval crosses margin



energy win, risk gate fails



One failed gate keeps $D(p,n) = 0$.

FIGURE 5 — A favorable mean energy difference is insufficient: promotion requires the entire simultaneous uncertainty region to clear the energy and quality margins while every risk, latency, and support gate remains valid.

Evidence status

- Neural energy constraints are established only within the scope of C-001.
- Nonlinear dendritic subunits, homeostatic scaling, neuromodulated plasticity, and specialized inhibitory control are biological observations under C-017 through C-020; their proposed artificial abstractions are not thereby validated.
- Performance of the complete cross-substrate system remains speculative until matched engineering tests establish it.

Speculative extensions

- Learn when global communication is worth buying instead of banning it.
- Compile frequently reused local circuits into fast deterministic paths while retaining a slower plastic path for exceptions.
- Search jointly over algorithm, memory hierarchy, interconnect, precision, and physical embodiment rather than treating hardware as a final deployment detail.
- Import organizational principles from organisms without centralized brains, including plants, immune systems, cephalopod limbs, and adaptive transport networks.

Failure modes

- A metaphor is mistaken for a mechanism.
- “Biological” becomes an unearned synonym for efficient or intelligent.

- “Silicon is faster” is used to ignore memory movement, synchronization, or thermal limits.
- A negative result for one implementation is treated as falsifying the underlying biological abstraction—or vice versa.
- The project claims an idea is absent from AI after finding only that it is absent from mainstream foundation models.
- Evolutionary fitness is confused with task accuracy, truthfulness, or human values.

Measurable predictions

This framing earns its place if, across isolated mechanism experiments:

- silicon-native abstractions equal or outperform literal translations at matched quality and system boundary;
- retained biological constraints predict where energy or interference is saved;
- removing a proposed invariant removes the benefit even when the biological surface form remains; and
- at least one hybrid uses a capability unavailable to biology—such as exact rollback or external versioned memory—to improve continual learning without increasing the declared risk metric.

Cross-domain convergence and principle deduplication

Different sciences often describe the same constrained operation in different nouns. The architecture should pay for the operation once.

Scope

This chapter turns the project's open-world research policy into an architectural method. Findings may originate in neural tissue, plants, immune systems, animal collectives, ecosystems, control theory, databases, or materials. They enter the design only after their causal operation has been normalized and compared with the principles already present.

The canonical unit is therefore neither a paper nor an organism. It is a versioned problem–solution invariant with a stable **P-** identifier, scoped evidence, an engineering null model, and an experiment that can reject its AI translation. The detailed registry remains in the principle ledger; this chapter explains how its bundles compose into one system.

Biological observation

The current evidence corpus repeatedly encounters five pressures.

- 1. Scarcity:** more states, candidates, or possible actions exist than can be active at once. Examples include constrained cortical signaling (C-001), sparse insect odor codes (C-025), immune selection (C-028), and congestion-triggered reserve routes (C-035).
- 2. Locality:** a complete central description is too slow or expensive. Dendritic branches integrate locally (C-017); cephalopod limbs retain peripheral control (C-024); and collective systems can coordinate through sparse influence or shared environmental state (C-031, C-054).
- 3. Reversible change:** recent evidence must affect behavior before it deserves permanent structure. Eligibility traces (C-019), plant priming (C-026), selective replay (C-036), and retrieval-sensitive memory (C-039) all separate a temporary state from a later commitment decision.
- 4. Stability:** selection and learning create positive feedback that must be bounded. Synaptic scaling (C-018), feedback inhibition (C-025), reconstruction around a target organization (C-033), and ecological recovery dynamics (C-058, C-059) expose different parts of that control problem.
- 5. Repeated work:** a stable solution should migrate from expensive general search into topology, placement, representation, or material. Relevant observations include body–controller co-development (C-029), structured pruning (C-012), lower-precision representation (C-013), and activity-dependent resource placement (C-050).

These recurrences do not make the underlying mechanisms identical. They reveal where multiple fields expose the same engineering pressure and where one shared experiment can replace several renamed proposals.

Proposed AI translation

From domain language to a mechanism record

Every retained observation is rewritten as the following tuple:

$$M = \langle p, s, b, x, f, G, \tau, \rho, \phi \rangle,$$

where:

SYMBOL	FIELD	REQUIRED CONTENT
p	problem	the failure or objective faced by the observed system
s	source organization and scale	form/material, process/organism, ecosystem/collective, or an explicit cross-scale interaction
b	constrained budget	energy, bandwidth, material, time, risk, or capacity with declared units
x	sensed state	what the mechanism can actually observe
f	causal operation	the intervention-supported state transformation
G	information topology	local, hierarchical, broadcast, pairwise, or environment-mediated flow
τ	timescale	event cadence or duration in declared steps or seconds
ρ	reversibility	what can decay, reopen, roll back, or be reconstructed
ϕ	failure boundary	conditions under which the effect disappears, reverses, or becomes harmful

The tuple is a structured research record, not a numerical embedding. Its fields retain units and provenance. Two findings are candidates for one principle only when their problem, causal operation, information topology, and timescale agree at the abstraction needed by an experiment.

Source organization and scale are not optional metadata. A material structure, an organism-level control loop, and an ecosystem-level interaction can deliver a similar function through different causal paths, authority, and timescales; they do not merge until those differences are shown irrelevant to the discriminating experiment.

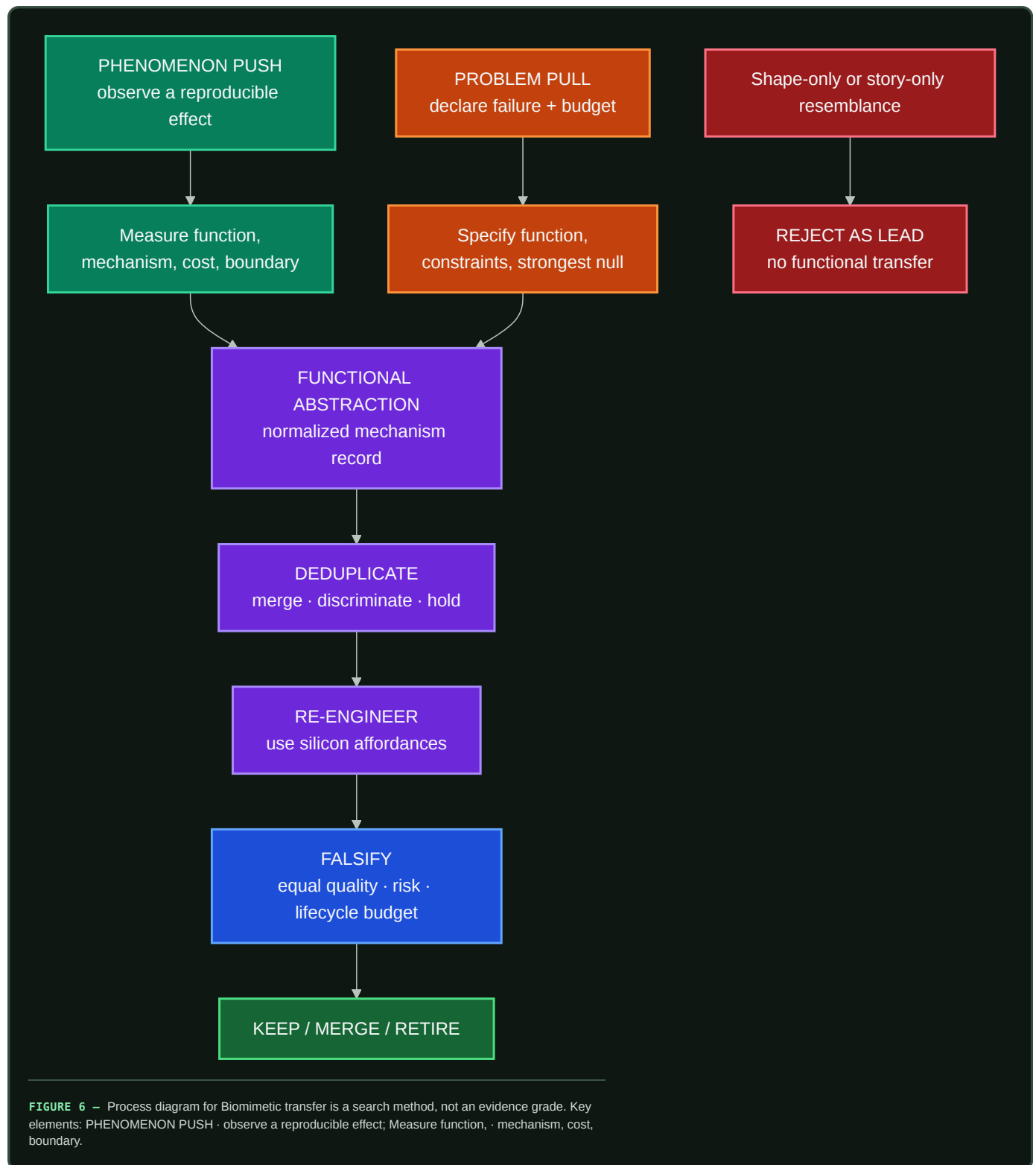
Biomimetic transfer is a search method, not an evidence grade

Biomimetics gives this project useful process language, but it does not create a shortcut through the evidence ledger. The search runs in both directions:

- 1. Phenomenon push** begins with a measured effect in a living system and asks which artificial problem shares its function, constraints, causal operation, and failure boundary.
- 2. Problem pull** begins with a measured artificial-system failure and asks which scientific fields contain a mechanism that solves the normalized problem under a comparable budget.

Phenomenon push broadens the established biology-push/solution-driven route; problem pull broadens technology-pull/problem-driven design. The broader names matter because this project searches physical, formal, social, and engineering sciences as well as biology.

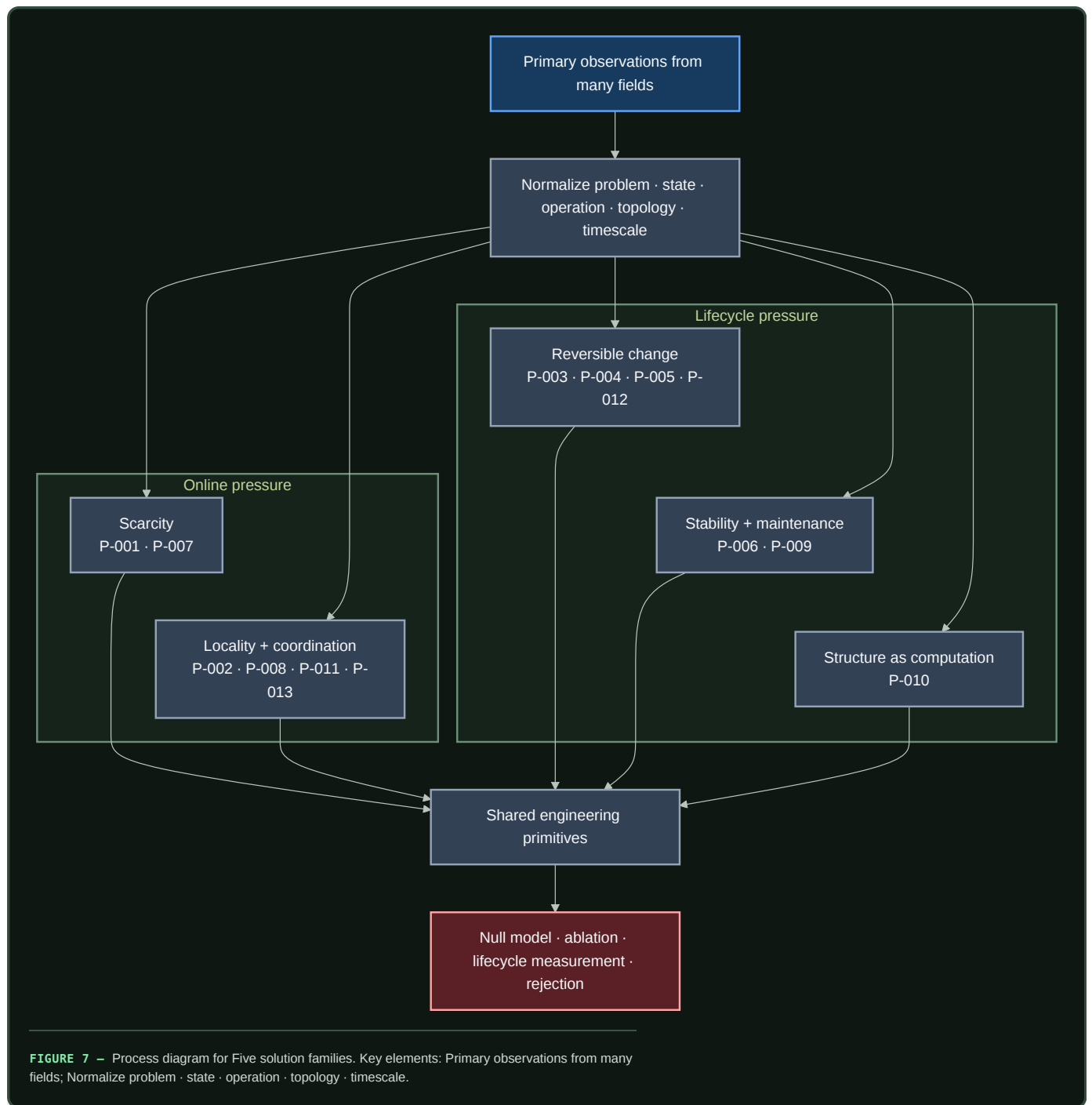
Both routes converge on M , then pass through deduplication, silicon-native redesign, a strong conventional null, and a falsifiable equal-budget test. A visual likeness without transferred function is rejected; a transferred function without a causal and cost model remains only a lead; and a validated source mechanism does not establish that its artificial translation works.



The public descriptions of biomimicry are useful for discovering biological strategies and unifying themes, while ISO 18458:2015 and VDI 6220 Part 2 provide more precise terminology and process frames. None is evidence that a particular mechanism transfers or improves efficiency. The dated method audit records the source hierarchy and the parts adopted here.

Five solution families

The thirteen current **P-** principles remain distinct, but they compose into five navigational families:



Editable source: `../assets/diagrams/recurring-solution-families.mmd`.

FAMILY	INCLUDED PRINCIPLES	SHARED ARTIFICIAL PRIMITIVE	IMPORTANT SEPARATION INSIDE THE FAMILY
scarcity	P-001, P-007	budgeted selection plus the option to buy more evidence	selecting a candidate is different from deciding whether uncertainty warrants more work
locality and coordination	P-002, P-008, P-011, P-013	stateful local modules with typed, priced communication paths	escalation, temporal binding, and shared workspaces solve different coordination failures
reversible change	P-003, P-004, P-005, P-012	versioned candidates, traces, memories, and topology with explicit promotion or decay	delayed credit, episodic content, candidate diversity, and graph mutation require different state
stability and maintenance	P-006, P-009	a slower controller that observes aggregate state and can throttle, repair, replay, or reopen	regulating a variable online is different from scheduling lifecycle work
structure as computation	P-010	migrate mature recurring work into topology, placement, precision, or compiled paths	each target has different invalidation, migration, and recovery costs

Deduplicate before architecture expansion

A new finding passes through six decisions:

1. **Scope the evidence.** Record the observed system, intervention, result, uncertainty, and boundary as a **C-** claim.
2. **Normalize the mechanism.** Fill every field of *M* without organism-themed names standing in for operations.
3. **Search the registry.** Compare against all existing **P-** records, including held candidates and negative evidence.
4. **Merge or discriminate.** Merge when the same control loop would be tested; otherwise name the smallest experiment that distinguishes the mechanisms.
5. **Name the strongest null.** A scheduler, cache, controller, estimator, database, routing rule, or standard learning method gets the same interface and budget.
6. **Change one canonical object.** Update a principle, primitive, equation, diagram, or experiment instead of adding another themed architecture.

This creates a many-to-one evidence graph:

many observations → scoped claims → fewer principles → shared primitives → decisive experiments

Recurring outcomes still need a construct firewall

Different fields often expose the same broad problem while assigning different causal content to it. Ecological succession describes ordered community change; an ecological priority effect requires an arrival-order intervention; continual-learning task order concerns sequential parameter updates; curriculum learning chooses an order; and a scheduler may merely decide when unchanged jobs execute. Calling all five “succession” would erase the experiment.

The history-conditioned modular-succession audit therefore reuses C-008, C-056, C-057, C-574, Fixture F-014, and Candidates 004 and 019 without creating another principle. Its residual test holds the task multiset, eligible module identity set, exogenous presentations, update ceilings, capacity, optimizer, evaluator and budget fixed, randomizes order, and separately cuts age, routed acceptance, capacity pre-emption, shared-state modification, facilitation and lock-in. If canonical replay, a non-learning scheduler, a mature continual-learning method, or a search-budget-charged curriculum explains the outcome, the biological transfer adds nothing.

New mechanisms can sharpen a principle without multiplying it

The developmental, regenerative, plant-plasticity, and multiscale-reduction depth passes add new failure boundaries without adding another universal principle. That is a useful result, not a failed search.

- Positional memory and repair separate **what state says should exist** from **which workers can currently rebuild it** (C-1506, C-1507). They sharpen P-005, P-009, and P-012 by making write authority, reconstruction capacity, and current context separately testable.
- Compensating sources, finite scaling, receiver geometry, delayed reinforcement, conditional redundancy, and local/global fields (C-1508–C-1515) refine locality, maintenance, and structural offloading. They do not imply that every duplicated route, boundary, or organizer is useful.
- Plant memory divides population-level digital switching, lifecycle reset, writing, trace maintenance, retrieval, systemic route identity, action-conditioned growth, and joint environmental context (C-1516–C-1525). These are discriminating state and protocol dimensions inside P-002, P-005, P-006, P-009, P-012, and P-013—not a plant-themed architecture.
- Projection memory, normal-hyperbolicity limits, macro-to-micro queries, and lift dependence (C-1526–C-1529) supply mathematical rejection tests for coarse state. They constrain when an existing runtime or memory primitive is valid; they are not biological claims at all.

The deduplicated output is therefore richer in contracts while remaining at thirteen P-series principles. Fixtures F-022, F-023, and F-024 preserve the distinctions so later implementations can test whether a proposed composition adds value beyond its strongest ordinary null within a declared workload and transfer boundary.

Composition without double counting

Principles can cooperate while remaining separately testable. For one event:

1. prediction-error allocation decides whether more evidence is valuable;
2. selective allocation chooses a module within the active budget;
3. local autonomy runs that module near its state;
4. a temporary trace records unresolved credit;
5. the maintenance plane later decides whether to replay, protect, merge, or forget it; and
6. structural offloading compiles only the recurring path that survives those tests.

Calling the entire chain “sparsity,” “memory,” or “homeostasis” would hide which operation produced a gain. Each experiment therefore removes one principle at a time while holding the other interfaces constant.

Interconnection is an intervention

A small or typed interface does not by itself make two modules dynamically modular. In transcriptional regulation, covalent-modification cycles and yeast synthetic circuits, attaching a downstream target can alter the upstream state or transient response (C-1550–C-1555). Substrate competition in *Drosophila* MAPK signalling supplies a distinct, potentially functional case (C-1559). The shared mechanism record is **connection-induced back-action**; the biochemical literature calls one important family retroactivity.

The domains do not collapse into one mechanism. Direct binding can sequester a signal, multiple substrates can compete for an enzyme, and nominally separate modules can instead couple through shared transcriptional, translational, CPU, memory, or scheduler resources (C-1558). The causal intervention therefore crosses connection state and shared-resource load independently. It also includes a genuinely nondestructive immutable-read control, because silicon can often copy or address state without the molecular consumption that motivated the analogy.

Insulation is bounded rather than automatically desirable. It must preserve both the producer and useful downstream service while exposing latency, staleness, copies, buffers, replicas, operations, maintenance and physical energy. Weak coupling may spend less while losing tracking or robustness (C-1556, C-1557); substrate competition can also be a useful integrator rather than a fault (C-1559). Fixture F-027 therefore compares suppressing, preserving and deliberately using back-action. If immutable messages, ordinary queues, resource isolation, explicit filters, stop-gradient paths or conventional controllers reach the same frontier, the translation is retired without adding a principle.

Passive self-organization is a mandatory null

Efficient-looking structure does not establish sensing, represented goals, or counterfactual action. Fracture sets, drainage networks, and river avulsions can arise from local stress, gravity, flow, conservation, thresholds, and stored geometry (C-232–C-242). The project calls a mechanism adaptive control only when it specifies:

1. a service variable external to the adaptation law;
2. observations with spatial/temporal support and latency;
3. a decision that could choose differently under different evidence;
4. authority and an actuator;
5. movement, reserve, monitoring, and recovery budgets;
6. persistent state and reset cost; and
7. a guarantee or tested failure envelope.

This criterion makes passive physics a stronger baseline. If local flow–structure feedback produces the same topology and service without a controller, the controller must justify its sensing, decision, switching, and maintenance cost. Connectivity is not throughput (C-234); material removed by “pruning” must appear as transport, storage, or output elsewhere (C-239);

and a topology change can reassign or destroy service rather than improve it (C-242). A closed-form branching law is therefore a regime-qualified physical null, not a portable architecture rule (C-1489).

Transported fields are not messages

Microbial signals, metabolites, fungal nutrients, plant hormones, process streams, queues, and network packets can all create a local field that later receivers use. The common systems object is not “communication in nature.” It is a transported, transformed, delayed, and locally decoded quantity whose physical path may also change future capacity (C-563, C-578, C-580).

For transported type k , use the local balance

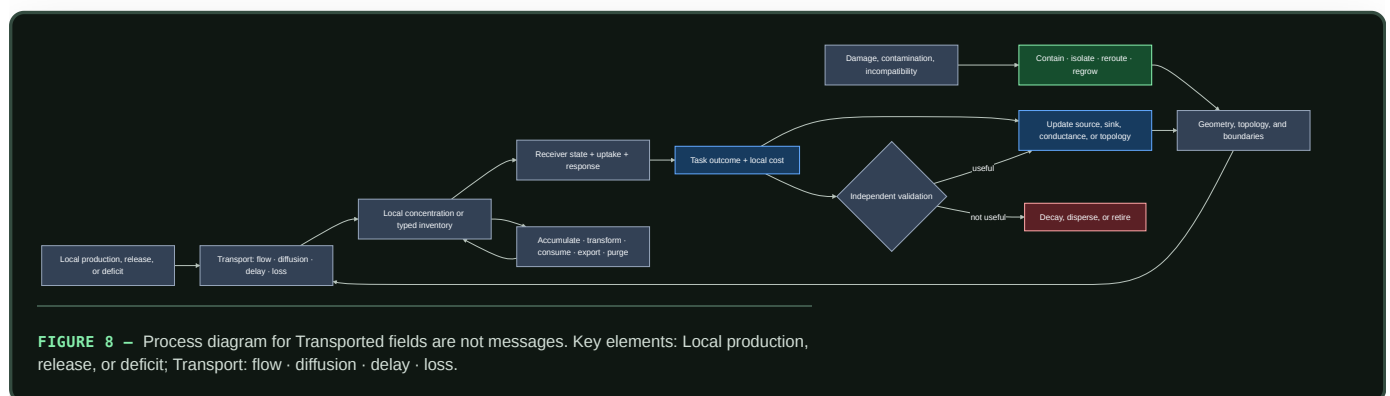
$$\frac{\partial c_k}{\partial t} = D_k \nabla^2 c_k - \mathbf{u} \cdot \nabla c_k + p_k - q_k - r_k c_k,$$

where concentration c_k is mol/m³, diffusion coefficient D_k is m²/s, velocity $\mathbf{u}$ is m/s, production and uptake rates p_k, q_k are mol/(m³ s), and first-order loss r_k is 1/s. Digital fields use an analogous typed byte or item ledger; they do not inherit molecular units or diffusion laws. Geometry, boundary conditions, sampling support, and receiver state are part of the contract. Countercurrent orientation is one finite exchanger arrangement inside that contract; it does not establish effectiveness or a generic bidirectional-computation benefit by itself (C-1490).

Several fields can be superposed at the same endpoints. If the observed outcome is

$$y = \sum_{g=1}^G h_g(c_g, \theta_g) + b + \varepsilon,$$

where G is the number of typed channels, g indexes one channel, c_g is its transported concentration or inventory, h_g is its receiver-dependent effect, θ_g is channel state, b is background, and ε is measurement error. Observing y or cutting every channel together does not identify any one h_g . Each claimed channel needs an independently targetable intervention or a justified identification model. Rillig et al.'s concurrent-fungal-network perspective makes this problem explicit for fungal guilds sharing plant endpoints; the project adopts it as an engineering inference, not as evidence that such guild combinations have a demonstrated beneficial function (DOI, C-217, C-584).



Editable source: [transported-field-contract.mmd](#).

The diagram deliberately spans several fields while retaining their differences:

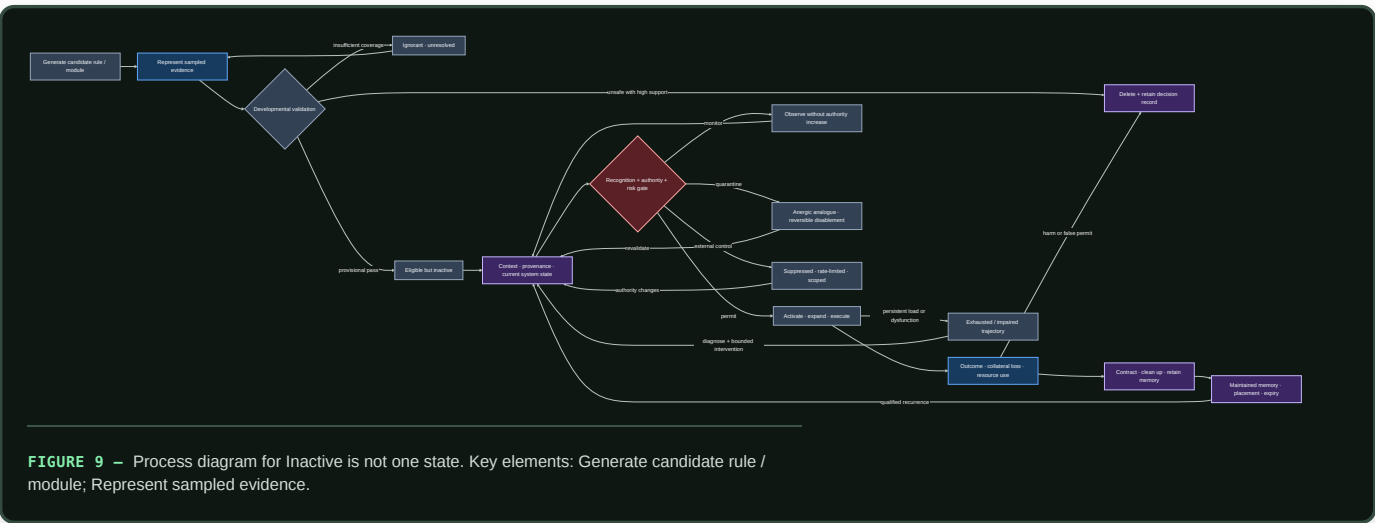
- microbial quorum output depends on production, confinement, flow, loss, and receptor state; it is not a vote count (C-563, C-575);
- cross-feeding can help, harm, or create a weakest-link dependency, so local transfer is not system utility (C-571);

- fungal flow can both deliver material and remodel conductance while rapid containment remains distinct from later bypass (C-577–C-581); and
- distinct symbiotic partners can exchange unequally under conflicting objectives (C-584); and
- concurrent typed channels sharing endpoints remain causally unresolved when the experiment observes or disables only their aggregate.

The held fixture belongs jointly to P-011/P-013 and existing Candidates 001 and 013 (C-585). It must beat typed pub/sub or queues, fixed-graph adaptive routing, backpressure/primal–dual allocation, and make-before-break state transfer. Charge production, transport, storage, cleanup, contamination, topology edits, failed delivery, reserve, and measurement. If geometry and cleanup do not affect the task, use the simpler digital primitive.

Inactive is not one state

The immune evidence in C-727–C-747 does not add another architecture principle. It contributes a stricter state contract. Representation, recognition, authentication, authorization, activation, suppression, deletion, impairment, contraction, memory, and recovery answer different questions and have different reversal costs.



Editable source: [typed-tolerance-lifecycle.mmd](#).

An activation bit cannot distinguish an absent module from an unobserved one, a policy-suppressed module from a damaged one, or post-response contraction from retained memory. That ambiguity makes reactivation unsafe and makes resource accounting wrong. Each transition therefore records evidence support, authority, expected useful effect, collateral-loss terms, resource budget, expiry, recovery target, and observed outcome.

Recognition supplies evidence; it does not grant authority. Deletion requires a higher support threshold than reversible monitoring or quarantine because a rare useful capability can be lost. Local copies pay maintenance, refresh, invalidation, and inconsistency costs. Recovery ends only when useful service and reserve are both restored.

The [typed lifecycle mathematics](#) compares this contract against a complete ordinary stack: typed state machines, calibrated risk and abstention, least privilege, constrained control, anomaly detection, evolutionary or ensemble search, replay, placement, and resource-aware scheduling. If the ordinary stack reproduces the decisions and frontier, the biological vocabulary is removed and the state contract remains.

Coordination without a privileged clock

Music cognition provides a combined stress regime for mechanisms that already exist elsewhere in the project: multiscale prediction, local timing control, transformation-aware memory, constrained generation, role protocols, repair, and cumulative transmission. The evidence in C-748–C-783 does not justify a music-specific principle.

FIGURE 10 — Process diagram for Coordination without a privileged clock. Key elements: Score · cue · literal task constraints; Phrase · motif · boundary state; PLL · Kalman · MPC · retrieval · typed protocol · rehearsal; Partner · role · repertoire memory; Correct · clarify · rehearse · update convention; Literal task · synchrony · response · quality · cost; Role-qualified local controller; Timed expressive action; Partner action; Delay · filter · loss · acoustical/physical path; Timestamped local observation + uncertainty; Local phase · tempo · drift estimate; Tempo drift · role switch · leader lock · new motif · turnover.

The fixture refuses several easy substitutions:

- similar mean tempo is not pairwise synchronization;
- stimulus-frequency neural power is not evidence of endogenous prediction;
- low onset error is not the correct phrase or role;
- familiar-partner performance is not co-adaptation when retrieval explains it;
- novelty is not validity, responsiveness, or usefulness; and
- within-group convention compression is not cross-group task success.

Preserved bits are not preserved meaning

FIGURE 11 – Process diagram for Preserved bits are not preserved meaning. Key elements: Payload · metadata · evidence links; Versioned preservation package.

20W · 33

This separates several failure modes that are otherwise mistaken for one success:

- fixity detects bit change; it does not establish truth or authenticity;
- provenance records lineage; it does not validate the assertion;
- a valid ontology constrains representation; it does not prove instance data;
- a citation edge is not endorsement or comparable impact;
- findability and interoperability do not establish correctness or safe reuse; and
- retaining records does not retain the tacit capability needed to interpret or act on them.

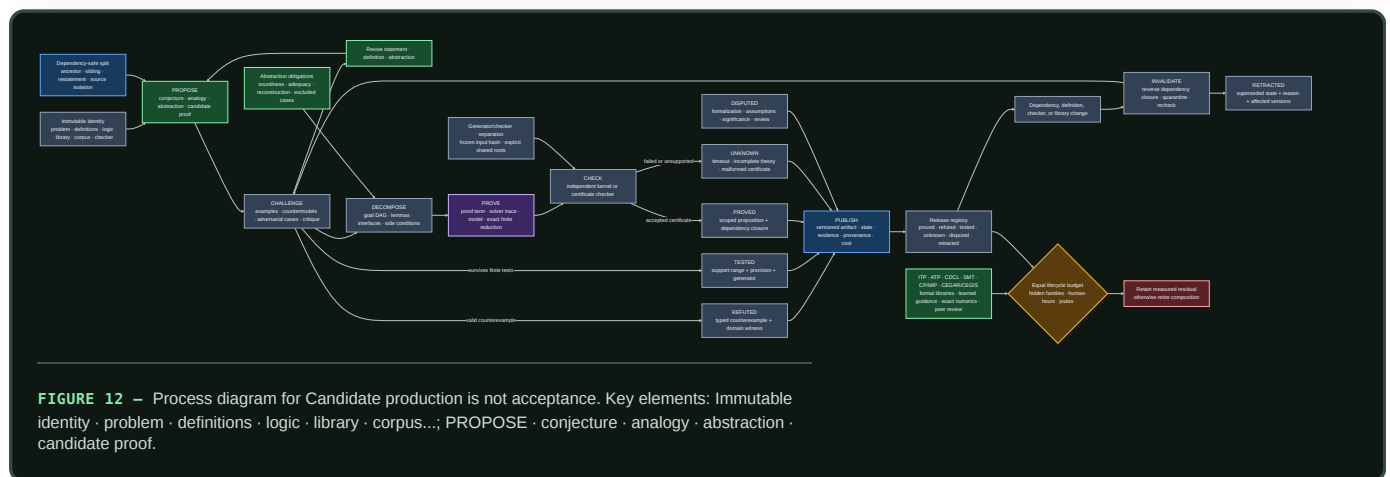
The query-registered preservation mathematics defines finite native-unit tolerances, evidence reachability, interpreter and dependency validity, migration regression, unregistered-query abstention, and complete lifecycle cost. The Candidate 017 track must beat ordinary OAIS/PREMIS-style packaging, versioned schema/vocabulary, retained-source recovery, and query-regression tests. If it does not, the ordinary preservation stack remains the implementation.

Candidate production is not acceptance

Mathematical practice supplies the sharpest version of a distinction already needed throughout this project: producing a candidate and earning acceptance are different operations. Conjectures, analogies, abstractions, proof sketches, retrieved lemmas, numerical patterns, solver traces, and learned tactic choices can all direct search. None of those origins establishes the resulting claim. The evidence and formal-system boundaries are tracked in C-861–C-880.

The transferable lifecycle is typed:

1. **Propose.** Bind a candidate to the exact problem, definitions, logic, axioms, library, corpus, proposal ancestry, and accessible evidence.
2. **Challenge.** Search examples, counterexamples, countermodels, boundary cases, definition changes, and adversarial encodings before commitment.
3. **Decompose.** Turn the target into a dependency DAG whose child artifacts reconstruct the parent goal through named interfaces and side conditions.
4. **Prove or refute.** Produce a proof term, checked finite reduction, solver certificate, admissible model, or exact counterexample—not only a verdict.
5. **Check.** Bind the artifact to the exact instance and library version, then run a small independent checker while publishing shared trust roots.
6. **Publish.** Retain `proved`, `refuted`, `tested`, `unknown`, `disputed`, and `retracted` as different states with evidence, provenance, and cost.
7. **Invalidate.** A definition, axiom, checker, dependency, or library change quarantines its reverse-dependency closure until rebuild or recheck.



Editable source: versioned-proof-discovery-lifecycle.mmd.

This lifecycle prevents five common substitutions:

- finite agreement is evidence for a tested range, not a universal result;
- a counterexample rejects the encoded universal claim but does not choose its unique repair;
- individually plausible lemmas do not close a parent goal unless their proof artifacts reconstruct it;
- a checker establishes the encoded proposition relative to its logic, definitions, axioms, libraries, preprocessing, and trusted implementation; it does not validate the informal intent or importance; and
- participant or model count does not establish independent checking when sources, parsers, libraries, prompts, or failure modes are shared.

Fixture F-004 compares the joined lifecycle with complete interactive and automated theorem provers, formal libraries, CDCL SAT, DPLL(T) SMT, constraint and model finding, CEGAR/CEGIS, learned premise/tactic guidance, exact numerical and exhaustive methods, and ordinary expert review. The **mathematical contract** defines immutable identities, leakage closures, abstraction obligations, proof-DAG reconstruction, certificate cost, typed states, invalidation, human effort, and joules. If the mature stack ties, the composition is retired; the state distinctions and negative result remain.

Efficiency mechanism

Deduplication saves two different resources.

First, it reduces research duplication. If plant priming, eligibility traces, and cache admission all motivate a temporary state before commitment, the project maintains one promotion interface and tests the domain-specific differences as variants. Papers and claims remain separate; architecture and instrumentation are reused.

Second, the resulting families define where the runtime should avoid repeated work:

- scarcity families reduce unnecessary activation and acquisition;
- locality families reduce movement and synchronization;
- reversible-change families prevent every event from rewriting durable state;
- maintenance families move repair and integration off the critical path; and
- structural offloading reduces the recurring cost of mature behavior.

For principle *j*, lifecycle acceptance uses the energy contract from chapter 80:

$$\Delta E_j(N) = N (E_{B,event} - E_{j,event}) - E_{j,introduce} - E_{j,maintain} - \mathbb{E}[E_{j,recover}],$$

where *N* is a dimensionless count of qualified events, event terms are joules per qualified event, and introduction, maintenance, and expected recovery are joules over the same observation horizon. A positive $\Delta E_j(N)$ is only an energy result; quality, risk, calibration, latency, and resilience must also remain inside their declared envelopes.

Evidence status

PROPOSITION	STATUS	BASIS
several scientific domains expose recurring scarcity, locality, memory, stability, and structural pressures	established within the current scoped corpus	claims ledger and domain inventory
candidate production and scoped formal acceptance are different artifact states	established or plausible within the audited formal systems and experiments	C-861–C-879; the joined lifecycle in C-880 remains plausible
the thirteen current principles are the correct deduplication	plausible working taxonomy	principle registry; boundaries remain revisionable

PROPOSITION	STATUS	BASIS
five families provide a useful navigation layer without erasing mechanism differences	proposed synthesis	must improve retrieval, experimental reuse, and reviewer agreement
recurrence across less-related domains predicts a useful artificial primitive	speculative	requires prospective tests against matched null models
the complete composition improves lifecycle efficiency	speculative	isolated and composed experiments have not yet established it

Speculative extensions

- Represent claims, principles, primitives, experiments, null models, and failures as a queryable versioned graph while keeping Markdown canonical.
- Track negative results as first-class edges so a rejected translation is not repeatedly rediscovered under another domain name.
- Estimate reviewer agreement on mechanism tuples before allowing a new principle ID.
- Search specifically for counterexamples: fields where the same pressure produces a different stable solution or where the recurring solution fails.
- Use contradictions between domains to generate new experiment regimes rather than averaging the difference away.
- Maintain silicon-native escape routes beside every principle: exact copying, direct addressing, typed storage, rollback, high-speed communication, and variable precision can change which operation is cheapest.

Failure modes

- **Naming duplication:** the same feedback loop appears repeatedly as a new organism-inspired component.
- **False convergence:** similar diagrams hide different sensed variables, causal operations, or timescales.
- **Overcompression:** a family becomes so broad that no ablation can isolate its mechanism.
- **Evidence laundering:** recurrence is treated as proof of independence, optimality, or transfer to AI.
- **Null-model neglect:** a familiar cache, scheduler, controller, or estimator is omitted because the biological story sounds novel.
- **Interface drift:** two variants share a principle ID while receiving different inputs, budgets, or evaluation envelopes.
- **Positive-result bias:** failed translations disappear, so the same proposal returns with a new metaphor.
- **Double-counted savings:** two principles claim the same avoided operation or compare against different baselines.
- **Verifier laundering:** checker acceptance is reported as truth about an informal statement without exposing formalization, axioms, dependencies, or the trusted implementation.
- **Finite-to-universal leakage:** a sampled numerical or test range is silently promoted from `tested` to `proved`.
- **Correlated checking:** producer and verifier share the parser, library, preprocessing, or defect that matters while being counted as independent.
- **Taxonomy lock-in:** stable IDs are mistaken for immutable scientific truth.

Measurable predictions

1. Independent reviewers given only normalized mechanism records agree on merge-versus-separate decisions more often than reviewers given titles and domain descriptions alone.

2. As domain coverage grows, the number of supporting claims per accepted principle rises faster than the number of principles; a near one-to-one ratio signals failed deduplication.
3. At least one proposed organism-specific mechanism is experimentally indistinguishable from an existing primitive at matched interface and cost and is removed rather than renamed.
4. Experiments built around shared principles reuse telemetry, null models, and failure regimes across more than one source domain.
5. Per-principle ablations attribute lifecycle savings to distinct avoided operations; overlapping savings disappear when all variants use one common baseline and boundary.
6. A principle promoted from recurrent evidence survives at least one regime derived from a domain outside the one that originally motivated its AI translation.
7. A typed propose–challenge–prove–check lifecycle reduces bad acceptance, localizes repair, and improves dependency invalidation beyond a generator plus checker alone after search, certificates, libraries, human effort, and joules are charged; otherwise Fixture F-004 retires the composition.

Structural growth, specialization, and conditional routing

Scope

Define how a modular system acquires new capacity without running, training, or retaining every possible module for every event. Growth is admitted only for a measured capability gap, new capacity earns traffic in probation, and mature capacity remains eligible for merge, reopening, or retirement.

The central object is a **capacity lifecycle**, not a continuously expanding pool. Birth, routing, specialization, placement, consolidation, and removal are separate decisions with separate costs.

Biological observation

Developing nervous systems generate and reorganize more cells and connections than remain in mature circuits. In the studied mouse retinogeniculate system, relative activity and complement signaling changed microglial engulfment and retention of developing inputs (C-043). This supports activity-sensitive structural refinement by a slower maintenance process, without specifying one general pruning rule.

Protection is not necessarily permanent. Targeted extracellular and receptor interventions reopened specific forms of adult visual-cortex plasticity (C-044, C-045). The resulting engineering states are candidate, consolidating, protected, reopened, and retiring. A module can be stable without becoming impossible to revise.

Other biological systems expose control operations that recur at different scales:

- germinal-center affinity maturation combines variation, selection, expansion, and later protection of useful lineages (C-028);
- *Physarum* and fungal networks couple use-dependent reinforcement, exploration, fusion, and contraction to changing flow (C-027, C-034);
- ants can open reserve routes under crowding before throughput falls (C-035); and
- activity can recruit local energy production, alter resource placement, and change local vascular supply in scoped neural preparations (C-049–C-051).

These mechanisms are not interchangeable. Together they impose a useful system constraint: variation and reserve capacity consume resources; selection needs an independent test; frequently used structure may stabilize; unused or duplicated structure must be able to leave the hot path.

A defined microbial-community experiment adds a practical design tactic: adding candidates selected for missing functions repaired the tested community better than merely restoring organism count (C-056). The inverse problem also matters. Functional redundancy was associated with poor newcomer engraftment in two small reanalyzed human microbiome datasets (C-057). Mature capacity can resist both harmful and beneficial entrants, so an artificial newcomer needs protected evaluation traffic rather than permission from incumbent routing logits alone.

Proposed AI translation

The capacity lifecycle

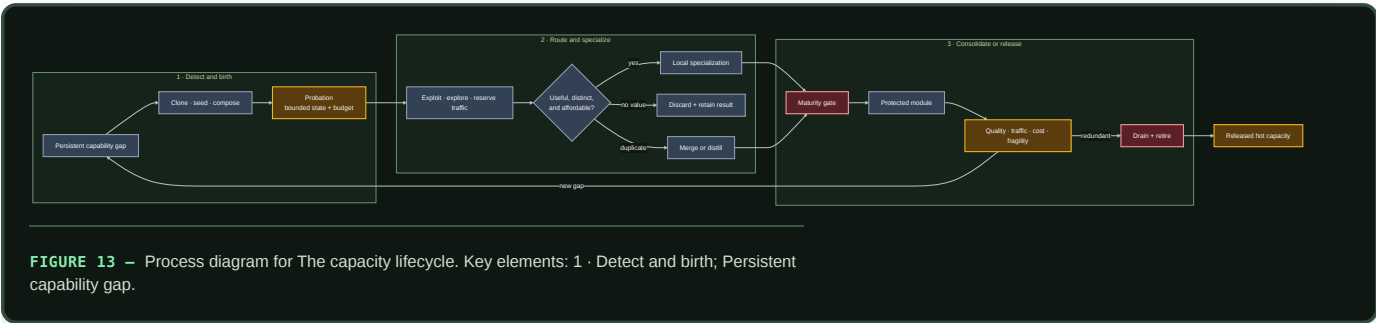


FIGURE 13 — Process diagram for The capacity lifecycle. Key elements: 1 · Detect and birth; Persistent capability gap.

Editable source: `../assets/diagrams/structural-growth-routing.mmd`.

The initial system contains modality encoders, predictive shared state, conditionally addressable experts, hierarchical routers, episodic and factual memory interfaces, and a declared reserve. Total addressable capacity may be large, but each event receives only a bounded route. Reserve capacity is stored, placed, and periodically tested; it is not free merely because it is inactive.

1. Detect a capability gap

Growth begins with a gap record, not a global loss spike. A valid record groups attributable episodes that existing routes fail in a consistent way and asks whether the failure is better explained by:

- missing capability;
- insufficient active compute or depth;
- missing context or memory;
- router error or capacity congestion;
- interference inside an existing module;
- distribution drift; or
- corrupted data, tools, or feedback.

The maintenance plane attempts the cheaper explanations first. New capacity is eligible only when the gap persists across resampling or recurrence, existing modules cannot absorb it inside their interference and resource bounds, and a candidate has a declared validation contract. This ordering prevents every hard example from becoming an expert.

2. Choose a birth operation

The proposal names both the new structure and what it is expected to repair:

BIRTH OPERATION	BEST INITIAL CONDITION	PRINCIPAL COST
Clone and diverge	one incumbent is close but suffers interference	copied parameters, optimizer state, later deduplication
Activate a seed	the gap is genuinely outside active coverage	stored reserve, cold-start training, placement
Compose a module	existing primitives are adequate but repeatedly coordinated	router depth, boundary traffic, compilation work
Reopen a protected module	a previously valid skill must change	regression risk, branch validation, rollback

Every birth creates a versioned provisional module. It cannot write the slow model, replace an incumbent, or claim permanent memory during probation. Initialization source, training episodes, expected role, resource ceiling, and discard path are recorded before it receives traffic.

Composing or admitting that module is itself a perturbation. The newcomer can pin shared activations, add router or queue pressure, alter normalization, create a gradient path, mutate shared cache state, or intentionally feed back. Before probation credit is assigned, compare the

incumbent producer under the same input and seed with the connection absent and present. Cross connection state independently from shared CPU, memory and scheduler load so ordinary contention is not mislabeled as connection-induced back-action (C-1550, C-1558). Immutable snapshots, bounded queues, resource isolation, stop-gradient adapters and explicit feedback remain the minimum nulls. The full qualification and retirement rules are in Fixture F-027.

Birth order is another intervention, not a background detail. A newcomer that arrives first can be older, receive more examples, pre-empt finite capacity, modify shared state, facilitate a later module, or create a costly lock-in. Those mechanisms predict different remedies. The same task multiset and eligible module identity set must therefore be replayed across randomized orders with equal exogenous presentations, task-local update ceilings, capacity, optimizer, evaluator and lifecycle budgets. Realized active, consolidated, merged, or retired state remains an outcome. Routed acceptance remains a measured mediator and is equalized only in the exposure-cut cells. Canonical replay, ordinary scheduling, replay/EWC/OGD, fixed and capacity-matched MoE, random curricula, optimized curricula with their search cost, population-based training and quality-diversity search are the required null stack. The exact factorial and kill rules are in the history-conditioned succession contract.

3. Give the candidate probation traffic

The router divides admitted work into three explicit budgets:

- **exploit traffic** goes to the strongest validated route;
- **exploration traffic** compares plausible candidates on informative events; and
- **reserve traffic** preserves failover and tests paths that would otherwise decay unnoticed.

A newcomer receives a capped share of gap-relevant episodes plus matched control episodes. The control traffic reveals whether it learned a capability or merely a narrow identifier for the failure cluster. Incumbents cannot reduce the evaluation share through their own confidence, but the maintenance plane can stop the trial for quality, risk, latency, memory, or energy violations.

For event x , use a dimensionless routing objective

$$\mathcal{L}_{\text{route}}(x) = \mathcal{L}_{\text{task}}(x) + \lambda_E \frac{\widehat{E}(x)}{E_0} + \lambda_B \frac{\widehat{B}(x)}{B_0} + \lambda_{\text{bal}} \mathcal{L}_{\text{balance}}(x) + \lambda_{\text{churn}} \mathcal{L}_{\text{churn}}(x),$$

where $\widehat{E}(x)$ is estimated joules/event, $\widehat{B}(x)$ is estimated bytes/event across named memory and network boundaries, and E_0 and B_0 are declared reference scales with the same units. All λ coefficients and losses are dimensionless. The physical joule and byte measurements remain separate reported outcomes; normalization does not turn them into task quality.

Balance keeps one expert from taking all traffic. Churn is applied only after a route has accumulated evidence of stable specialization; penalizing early movement would protect arbitrary initialization.

When module reports become strategic

Ordinary routing remains the default. Shadow-price feedback already represents scarcity under declared controller conditions (C-133), and auction or matching language adds nothing when the router can observe costs and every module shares the system objective. A market-like mechanism is in scope only when a persistent module holds decision-relevant private information, can improve its future traffic by misreporting, and faces a real opportunity consequence it cannot reset or evade.

That regime creates specific failures. Selection pressure on a visible metric can damage poorly measured substitute tasks (C-139); proper scoring needs an independently verified outcome and does not establish competence or causal contribution (C-140); peer agreement can reward shared error or collusion (C-141); and fixed-agent truthfulness does not solve false identities or a biased allocator (C-143).

Candidate 008 therefore begins with a cooperative applicability control that it should not beat. Only then does it introduce hidden costs, adaptive metric gaming, protected outcomes, entrants, identity resets, collusion, and allocator deviation. Withheld audits, lineage-bound consequences,

protected entrant traffic, and replayable commitments remain only if they improve external task, risk, energy, and latency outcomes after their evaluation and storage costs.

4. Measure specialization rather than naming it

A candidate becomes useful when it improves a defined region of behavior while remaining distinguishable from existing modules. Evidence includes:

- causal improvement when the candidate is admitted and regression when it is ablated;
- reduced interference on incumbents or protected history;
- consistent advantage on held-out gap and recurrence episodes;
- a stable but non-exclusive routing region;
- calibration and rare-case behavior inside its declared envelope; and
- physical cost that remains inside its allocation.

Low routing entropy alone is not specialization: a router can collapse onto a module for the wrong reason. High activation diversity alone is not useful: a pool can fragment one capability across many expensive duplicates. The test is complementary causal contribution at a measured lifecycle cost.

5. Hand off to merge, protection, or retirement

At the end of probation, the candidate has three normal outcomes:

- 1. Discard.** It adds no reliable capability. Preserve the negative result so the same proposal is not regenerated indefinitely.
- 2. Merge or distill.** It reproduces an incumbent or several modules have converged on one operation. Build a compact branch, rerun intervention and recurrence tests, then drain duplicates.
- 3. Consolidate.** It contributes a distinct reusable capability. Pass it to the maturity lifecycle for protection, reopening rules, structured pruning, and rollback.

A protected module remains monitored for traffic, unique contribution, physical placement, recovery, and newcomer exclusion. Persistent redundancy returns it to a merge-or-retire gate. Logical topology changes can be evaluated with Candidate 001, which charges reconfiguration, migration, reserve, controller, and recovery costs.

Conventional null models

The growth controller must beat ordinary ways of allocating or restructuring capacity, not only a frozen weak model:

ID	NULL MODEL	WHAT IT TESTS
N0	Capacity-matched dense monolith	whether modular growth is needed at all
N1	Fixed-capacity sparse MoE with tuned load balancing	whether conditional routing alone explains the gain
N2	Fixed modules plus adapters or low-rank updates	whether new structure beats ordinary parameter-efficient adaptation
N3	Periodic global architecture or topology optimization	whether a standard batch redesign explains local lifecycle control
N4	Usage- or magnitude-prune/retrain cycle	whether causal merge/retire gates add value
N5	Random valid birth, merge, and retirement under equal budgets	whether the lifecycle signals carry information
N6	Trace-aware oracle with future gap labels	unattainable ceiling; never a superiority baseline

Equalize initial capacity, maximum stored capacity, active work, optimizer updates, training examples, router information, tuning trials, migration bytes, validation work, and wall-clock opportunity. Charge candidate failures and discarded births. Otherwise growth buys more search while the null models are asked to solve the task in place. A stylized optimum such as a fixed branching exponent belongs in this null stack only inside its verified objective, transport law, and constraint regime (C-1489).

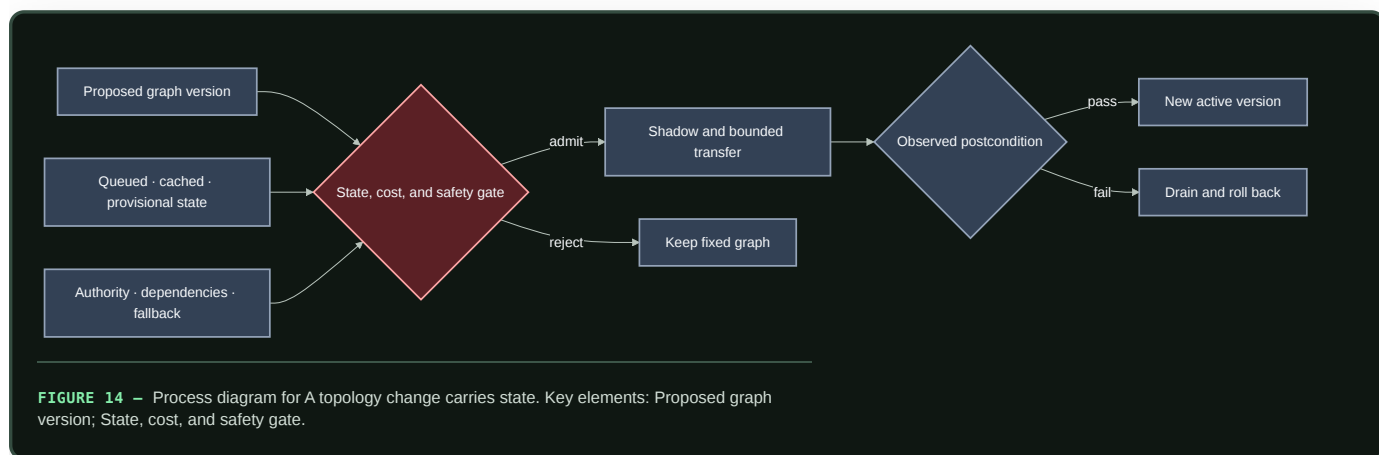
A topology change carries state

Changing an edge label is not the same as transferring a working system. The process-engineering audit makes the missing state visible: an installed path may hold inventory, energy, contamination, pending work, actuator authority, wear, calibration, maintenance obligations, and shared protection dependencies (C-501, C-510). Two endpoint configurations can both be feasible while the path between them is unsafe (C-512). Static branching relations do not remove that transition state (C-1489); fast allocation and slow structural adaptation must retain separate action, deficit, build, stranded-capacity, and recovery histories (C-1496).

For digital modules, use a typed transition inventory rather than pretending that bytes obey material conservation. For state class k ,

$$I_k(t_1) - I_k(t_0) = A_k - D_k - X_k + R_k,$$

where I_k is bytes present at a named boundary, A_k is admitted bytes, D_k is deliberately deleted bytes, X_k is exported bytes, and R_k is internally replicated bytes over $[t_0, t_1]$. Each term is bytes and carries a provenance, validity, and ownership version. This is an accounting contract, not a claim that information is physically conserved (C-517).



Editable source: [conservation-qualified-reconfiguration.mmd](#).

Candidate 001 therefore includes a physical stress track. It must beat fixed-graph adaptive control, multi-mode supervisory control, and offline redesign after installed reserve, transition state, flushing, downtime, maintenance, and rollback are charged. If those nulls tie it, “adaptive topology” describes an implementation choice rather than an efficiency mechanism (C-516).

Positional instruction, repair capacity, and structural fields

The developmental and regeneration residue sharpens the capacity lifecycle by separating operations that a generic “grow and repair” controller would hide. The claims are source-domain results; their engineering translations remain prospective and are tested only by the preimplementation F-022 fixture.

- 1. Keep repair instruction distinct from repair workforce.** In the studied planarian experiments, differentiated muscle carried position-control expression while neoblasts supplied replacement capacity (C-1507). A modular system should therefore test a two-plane design: surviving task modules retain bounded reconstruction constraints, while a replaceable worker pool executes them. A worker reserve is not positional knowledge, and retained instructions are not proof that enough capacity survives. DEV-T02 compares this separation with complete checkpoint-and-log restoration and requires abstention when surviving instruction support is insufficient.
- 2. Treat positional memory as writable, gated state.** Transient intervention rewrote persistent positional state during the scoped axolotl regeneration window, but not in comparable uninjured cells (C-1506). The translation is neither immutable metadata nor unrestricted self-editing: retain versioned local role state, open writes only inside an

authenticated repair episode, validate against independent service evidence, and preserve rollback. DEV-T01 makes valid, locally corrupted, and common-mode-corrupted memory separate cases so saved reconstruction work cannot excuse confident wrong-role repair.

3. **Separate source compensation, scale support, and receiver observability.** Opposed BMP-related sources compensated for scoped perturbations in *Xenopus* (C-1508); DEV-T03 asks whether reciprocal sources beat distributed robust control after delay, stability, and double-source detection are charged. Dpp-gradient scaling in the studied wing disc had an experimentally exposed finite envelope (C-1509); DEV-T04 therefore requires an explicit beyond-support disposition rather than a scale-invariance claim. In the cited micropattern assay, edge access and receptor orientation changed which signal a cell could observe (C-1510); DEV-T05 compares local boundary-relative inference with exact distributed graph distance. A compensating field can still be unreadable at the receiver, and a readable field can still be outside its calibrated size range.
4. **Consolidate paths and boundaries only after persistent evidence.** Tension recruited and stabilized Myosin II in the scoped tissue experiments (C-1511), motivating capped, decaying load-path reinforcement rather than permanent reinforcement after a burst; DEV-T06 includes load reversal and event-triggered backpressure as controls. An actively maintained Myosin-II barrier limited compartment mixing (C-1512), motivating delayed, reversible interface hardening tested against a mature soft boundary in DEV-T07. Apparent enhancer redundancy improved robustness only across particular perturbations in the studied loci (C-1513); DEV-T08 accordingly tests partially input-diverse small gates against full duplicate controllers and preserves common-mode corruption as a hostile null.
5. **Compose local patterning with coarse global context without conflating their roles.** The limb-pattern study supports a plausible, non-unique reaction–diffusion interpretation modulated by global gradients (C-1514). DEV-T09 therefore requires local-only and global-only ablations and compares the composition with central constrained assignment. Chick perturbations also support local-feedback symmetry breaking, while the proposed molecular model remains non-unique (C-1515); DEV-T10 tests temporary organizer election against a complete randomized leader-election protocol, including duplicate detection and re-election.

These distinctions add three gates to structural growth: verify that surviving instruction is sufficient before allocating repair workers; verify that the field is both inside its support envelope and observable by the receiving module; and verify that reinforcement or boundary writes amortize their write, maintenance, reopening, and fallback costs. Every F-022 track remains `NO_RESULT`; a written protocol does not establish an efficiency advantage.

Competence-gated structural transition

A signal is not a complete command. Developmental experiments show that concentration, duration, position, stage, and receiver state can change the response to the same extracellular input (C-539–C-549). Commitment can later be redirected, but reopening is an intervention with selection, resource, integrity, and safety costs rather than free reversal (C-550, C-551).

For module m and transition-contract version v , define the dimensionless admission predicate

$$G_{m,v}(t) = \mathbf{1}[q_m(t) \in \mathcal{C}_v, t \in W_v, e_m(t) \geq \theta_v, \mathbf{r}_m(t) \succeq \mathbf{r}_v^{\min}, R_{m,v}(t) = 1] .$$

$q_m(t)$ is typed local competence state; $\mathcal{C}_v$ is the permitted state region; t is seconds; W_v is a declared time interval in seconds; e_m and evidence threshold θ_v are dimensionless; resource vectors $\mathbf{r}_m$ and $\mathbf{r}_v^{\min}$ are compared componentwise in their native units; and $R_{m,v}$ is a dimensionless predicate that a tested rollback path exists. The signal is eligible to create a provisional transition only when $G_{m,v} = 1$.

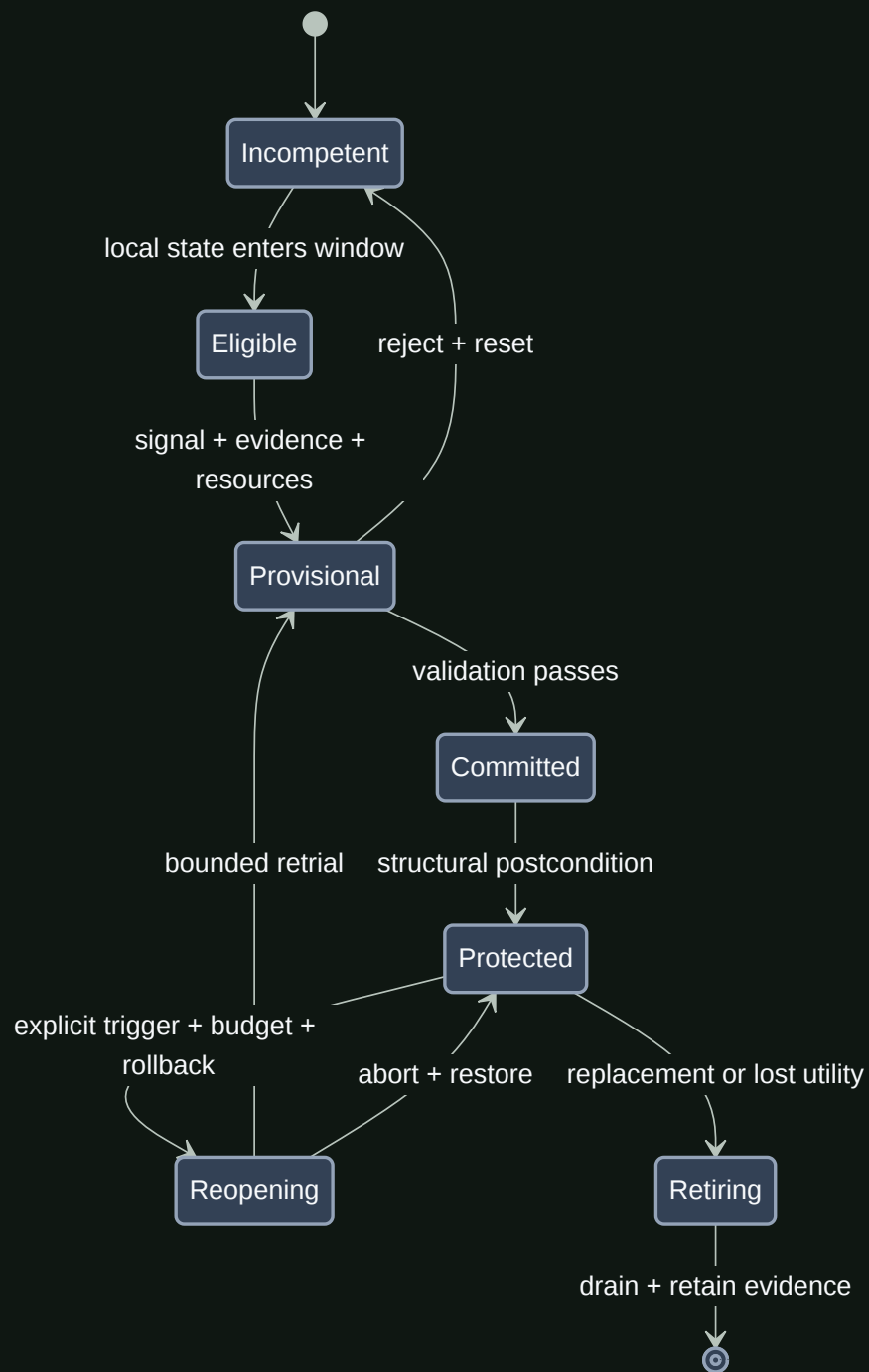


FIGURE 15 – State-transition diagram for Competence gated structural transition. Key elements: Incompetent; Eligible.

Editable source: competence-gated-transition.mmd.

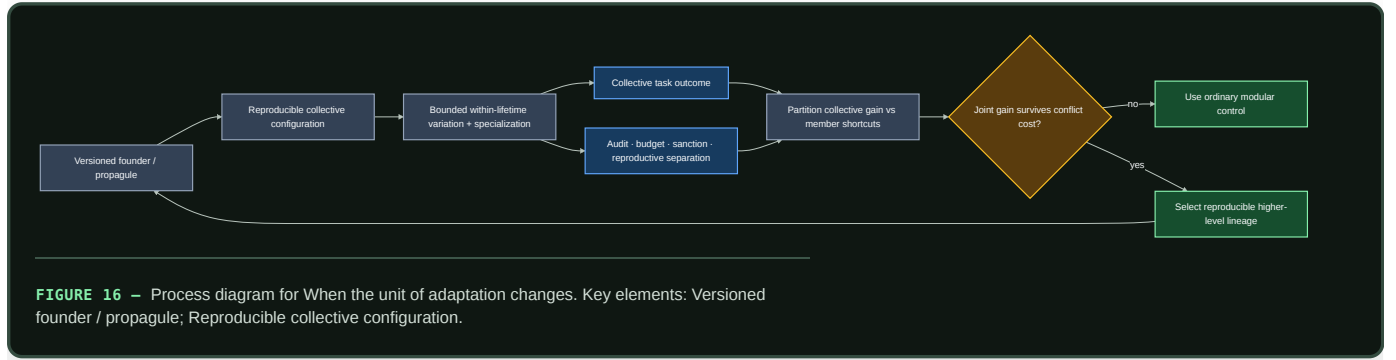
The fixture splits responsibilities instead of adding another candidate:

- Candidate 002 tests low-bandwidth signal decoding by versioned receivers;
- Candidate 009 owns admissibility, authority, dependencies, and invalidation;
- Candidate 010 owns provisional evidence before commitment and reset;
- Candidate 006 owns structural write, health, shadow, and physical reopening; and
- Candidate 014 owns support, stage, vintage, and the observations defining competence.

Reject the composition if an ordinary gate, schedule, staged verifier, and versioned migration procedure matches quality, interference, rollback, availability, bytes, latency, and joules. The fixture is useful only if the interaction among signal, receiver history, finite window, commitment, and reopening predicts failures those separate controls miss (C-562).

When the unit of adaptation changes

A cooperating set of modules is not automatically a higher-level unit. The claim becomes testable only after defining its boundary, child-configuration event, inherited state, within-lifetime state, descendant relation, collective performance, member-level incentives, and conflict-control cost. This separates aggregation from individuality and a demographic bottleneck from a reproducible founder boundary (C-282–C-295).



Editable source: [conflict-bounded-unit-transition.mmd](#).

For collective lineage k , let Z_k be task-native collective performance and W_k its dimensionless admitted-descendant weight. The selection accounting is

$$\Delta \bar{Z} = \frac{\text{Cov}(W_k, Z_k)}{\bar{W}} + \frac{\mathbb{E}[W_k \Delta Z_k]}{\bar{W}}.$$

The first term is change among declared collectives and the second is transmission change. A nested partition separately reports member-level shortcuts. The identity is accounting, not causal proof; group and descendant definitions are preregistered. Candidate 016 must beat typed modular systems, permissions and tests, clean versioning, external evaluation, routed experts, and ensemble/population selection after enforcement, false-sanction, interface, founder, reserve, and recovery costs.

Efficiency mechanism

For experts $i = 1 \dots n$, let $g_i(x) \in \{0, 1\}$ be the dimensionless event gate, $C_i(x)$ be executed operations/event under a declared precision, and $C_{\text{router}}(x)$ use the same convention:

$$C_{\text{active}}(x) = C_{\text{router}}(x) + \sum_{i=1}^n g_i(x) C_i(x).$$

This separates addressable parameter capacity from executed work, as sparse mixture-of-experts systems demonstrate in specific implementations (C-003). It does not price parameter reads, dispatch, all-to-all communication, imbalance, cold starts, or maintenance.

The lifecycle energy per served event is therefore

$$\bar{E}_{\text{capacity}} = E_{\text{route}} + \sum_{i=1}^n g_i E_i + E_{\text{comm}} + \frac{E_{\text{birth}} + E_{\text{train}} + E_{\text{place}} + E_{\text{validate}} + E_{\text{merge/retire}} + \mathbb{E}[E_{\text{recovery}}]}{N_{\text{served}}},$$

where every E term is in joules, E_i is expert execution energy/event, and N_{servd} is the number of events over the comparison horizon. Report stored parameter and optimizer bytes, bytes moved/event, latency, utilization, quality, calibration, and risk alongside energy.

Growth is efficient only if conditional execution saves more than candidate search, idle reserve, placement, validation, and later contraction consume. Lottery-ticket results show that competitive sparse subnetworks can exist in tested settings (C-012); they do not establish that this lifecycle discovers them or realizes energy savings on a target system.

Evidence status

ELEMENT	STATUS	ROLE IN THIS CHAPTER
conditional expert routing (C-003)	established in published systems	capacity and active compute can be separated in suitable implementations
competitive sparse subnetworks (C-012)	established in tested settings	staged selection and pruning are viable operators
use-dependent biological topology (C-027, C-034)	established in scoped organisms/models	motivates reinforcement, decay, exploration, and contraction tests
diversity, selection, and protection (C-028)	established in the cited immune experiment/model	motivates a bounded candidate lifecycle
congestion-triggered reserve use (C-035)	established in the cited ant setup/model	reserve paths should be priced and tested before overload
developmental refinement and reopening (C-043–C-045)	established in scoped neural preparations	supports distinct candidate, protected, and reopened states
local resource demand and placement (C-049–C-051)	established in scoped neural preparations	physical supply and movement belong in routing cost
shadow prices, matching, proper scoring, metric pressure, identity, and allocator credibility (C-133–C-143)	established under scoped economic models and experiments	ordinary routing remains the null; contestable allocation is conditional on a measured strategic-information problem
audit-backed contestable allocation (C-144)	speculative systems composition	Candidate 008 must lose its distinction when modules are cooperative and directly observable
capability-gap repair (C-056)	established for the defined mouse community and challenge	motivates selecting additions by missing function
functional redundancy and engraftment (C-057)	plausible association	motivates protected newcomer evaluation and a lock-in test
higher-level heredity and conflict accounting (C-282–C-295)	scoped population-genetic and evolutionary results; artificial composition speculative	Candidate 016 tests whether a collective becomes a useful adaptation unit beyond ordinary modular lifecycle controls
complete grow–route–specialize lifecycle	speculative synthesis	requires comparison with N0–N6

Speculative extensions

- Let modules request a birth trial with a compact capability-gap certificate; maintenance allocates the trial, not the requesting module.
- Maintain seed modules at several parameter and precision scales so a new role need not start from the largest available structure.
- Learn placement jointly with specialization only after migration cost and rollback are measurable.
- Use recurrence-aware cold storage: retire a module from hot execution while retaining enough checkpoint and routing evidence to restore it if its regime returns.
- Allow two candidates to share an encoder or memory interface while keeping their update authority and resource accounts separate.
- Test local birth/retirement against periodic global architecture optimization under recurrent rather than one-way task sequences.

Failure modes

FAILURE	OBSERVABLE SIGNATURE	REQUIRED RESPONSE OR ABLATION
Router collapse	one module takes most traffic; queue tails or overflow rise	fixed-capacity MoE and stronger load-balancing baseline
Candidate inflation	birth rate and stored bytes rise without held-out gap closure	cap trials; compare no-growth and random-birth nulls
Fragmentation	many modules show overlapping ablation effects and high boundary traffic	merge/distill branch with causal coverage tests
Incumbent lock-in	a superior newcomer cannot acquire evaluation traffic	reserved probation share; compare router-logit admission

FAILURE	OBSERVABLE SIGNATURE	REQUIRED RESPONSE OR ABLATION
Premature localization	modality-specific routes lose cross-modal transfer	shared-module and monolithic controls on compositional tests
Reserve starvation	no path remains for faults or new regimes	price and enforce declared reserve capacity
Reconfiguration thrash	repeated births, moves, merges, or retirements dominate cost	hysteresis and slower maintenance epochs
Stranded capacity	cold modules occupy memory but never serve, fail over, or restore	cold-storage, deletion, and restore-value comparison
Cosmetic sparsity	active gates fall while loaded bytes, communication, or joules do not	hardware trace and dense-kernel ablation
Maintenance inversion	search, validation, migration, and rollback cost exceeds runtime saving	full lifecycle equation and fixed-structure nulls
Rare-role deletion	average quality holds while rare or safety-critical cases regress	protected recurrence suite and reconstructable checkpoint
False higher-level unit	aggregate reward rises but collective inheritance is transient or member shortcuts dominate	preregister descendant relation; selection partition; ordinary modular and external-evaluator nulls

Measurable predictions

1. Capability-gap-driven births close held-out failure clusters with fewer admitted candidates and lower lifecycle energy than random birth, periodic fixed growth, and capacity-matched adapter baselines.
2. A protected probation share lets genuinely better newcomers establish causal value faster than incumbent-logit admission without increasing harmful promotions at the same validation budget.
3. Successful specialization produces a stable, non-exclusive routing region and positive unique ablation value; routing entropy alone predicts promotion less reliably.
4. Conditional growth reduces executed operations and bytes moved/event relative to a capacity-matched monolith while preserving quality, calibration, and rare-case performance.
5. Causal merge/retire gates preserve recurring and intervention capability better than usage- or magnitude-only pruning at matched hot capacity.
6. Reserve capacity improves recovery after faults or returning regimes enough to justify its stored bytes, idle energy, and periodic test traffic.
7. Joint routing and placement lowers communication energy only after migration, cold-start, and rollback work are included.
8. A collective lifecycle advances only when cost-adjusted between-collective selection and inherited capability persist under member shortcuts and turnover beyond ordinary lifecycle governance.
9. The complete lifecycle advances only if it improves the quality–risk–latency– energy– adaptability frontier over N0–N5; a parameter-count or FLOP reduction alone does not satisfy the prediction.

Multimodal sensorimotor grounding

Scope

Grounding means that a model's state is constrained by temporally ordered observation, action, and consequence—not that every concept must be reducible to pixels or motor commands. This chapter defines the trajectory record from which such state can be learned, how language attaches to it, and the tests that separate action-conditioned structure from passive multimodal correlation.

The central unit is a versioned episode containing asynchronous sensor events, commanded and realized interventions, exogenous events, and language with provenance. The model must predict and act under missing modalities, timing error, partial observability, and finite sensing energy. A representation is treated as grounded only to the extent that it supports held-out prediction, intervention, control, and reference at a declared uncertainty.

Biological observation

Biological perception is embedded in a closed loop. During active whisker exploration, measured contact mechanics predicted primary sensory-neuron firing better than whisker angle alone in the scoped mouse experiment (C-022). The result establishes that self-generated sensing mechanics can matter to the sensory code; it does not make embodiment necessary for every capability.

Other evidence constrains parts of the translation:

- sparse coding can learn localized structure from natural images in the model studied under C-002;
- latent target prediction can learn semantic image representations without pixel reconstruction under C-006;
- children in a controlled toy study targeted actions toward unresolved causal structure under C-062;
- pedagogical demonstration narrowed subsequent exploration in the scoped toy experiments under C-063; and
- causal transparency changed copying of irrelevant demonstrated actions in a comparative puzzle-box experiment under C-064.

Together these findings motivate aligned experience, active evidence collection, and separation of imitation from outcome learning. They do not yet establish the project-level hypothesis that embodied multimodal training yields more general physical concepts than comparable text-centric training. That claim remains speculative under C-007.

Proposed AI translation

Canonical trajectory record

One episode is

$$\tau = (\nu, \mathcal{O}, \mathcal{A}, \mathcal{X}, \mathcal{Y}),$$

where τ is an episode record and all five components are immutable after publication; corrections create a new version.

SYMBOL	MEANING	UNIT OR TYPE
ν	episode metadata	identifiers and versioned configuration
$\mathcal{O}$	asynchronous observation-event set	timestamped records
$\mathcal{A}$	commanded and realized action-event set	timestamped records
$\mathcal{X}$	exogenous event set	timestamped records
$\mathcal{Y}$	language-event set	timestamped text or token records

Metadata ν contains episode and environment identifiers; train, development, or test split; actor or policy version; simulator or device version; clock domains; calibration versions; random seed when applicable; and parent episode for a counterfactual branch. Frames from one episode never cross data splits independently.

An observation event from modality r is

$$or,k = (r, t_{r,k}^{\text{cap}}, t_{r,k}^{\text{recv}}, vr,k, mr,k, qr,k, pr,k),$$

where:

- r is a modality identifier such as vision, audio, touch, proprioception, force, temperature, or depth;
- k is an event index;
- $t_{r,k}^{\text{cap}}$ is physical capture time in seconds;
- $t_{r,k}^{\text{recv}}$ is the time in seconds at which the learner could first access the event;
- vr,k is the sensor value with the sensor's declared unit, shape, encoding, and quantization;
- mr,k is a dimensionless missingness/status code;
- qr,k is a quality record with named quantities and units, such as exposure time in seconds or signal-to-noise ratio in decibels; and
- pr,k is a provenance pointer to source, calibration, transformation, and checksum records.

Capture and receipt time are both required. Offline alignment uses capture time; a deployable causal policy may use an event at time t only when $t_{r,k}^{\text{recv}} \leq t$. This prevents a delayed sensor packet from becoming accidental future information.

An action event is

$$\alpha_j = (t_j^{\text{cmd}}, t_j^{\text{on}}, t_j^{\text{off}}, a_j, \tilde{a}_j, \pi_j, p_j),$$

where t_j^{cmd} , t_j^{on} , and t_j^{off} are command, realized onset, and realized offset times in seconds; a_j is the commanded action; $\tilde{a}_j$ is the measured realized action; π_j is the behavior-policy probability for a discrete action (dimensionless) or probability density for a continuous action (in reciprocal action-volume units) when it is known; and p_j is its provenance pointer. Each action component declares its physical unit—for example newtons, newton-metres, metres/second, radians, or a discrete tool identifier. Command and realization are never silently substituted for one another.

An exogenous event $x_l \in \mathcal{X}$ records event and receipt timestamps t_l^{event} and t_l^{recv} in seconds, event type, measured magnitude and unit, and provenance. Examples include another agent's action, an uncommanded collision, lighting change, or simulator reset. Known simulator state s_t^* may be stored for evaluation in its declared physical units, but it is oracle information and is withheld from learned systems unless a baseline explicitly receives it.

A language event is

$$y_n = (t_n^{\text{start}}, t_n^{\text{end}}, w_n, c_n, r_n, p_n),$$

where the two timestamps are seconds, w_n is text or a token sequence, c_n identifies speaker and communicative role, r_n is a set of referent pointers to objects, events, trajectory intervals, or external records, and p_n is source provenance. A retrospective caption, an instruction available before action, a question, a report from another agent, and model-generated text have different roles even when their words match.

Observation is a versioned contract

A sensor record is not the latent state. Propagation, foreground, exposure, instrument response, background, calibration, reconstruction, threshold, and selection intervene before a downstream model sees it (C-218–C-221). For latent target θ , nuisance state η , measurement y , and response version H_v ,

$$y = H_v(\theta, \eta) + \epsilon,$$

where y and H_v have matching declared units and ϵ follows a declared noise and background model. If the analyzed record was selected by event $S = 1$, inference is conditional on that event:

$$p(\theta, \eta \mid y, S = 1, v) \propto p(y, S = 1 \mid \theta, \eta, v) p(\theta, \eta \mid v).$$

The trajectory record therefore carries response and calibration version, exposure, selection state, detection power for non-detections, reconstruction choice, and data vintage. Cross-modal fusion also retains association uncertainty and shared dependencies; a different sensor type is not evidence of conditional independence (C-224).

Candidate 014 tests whether propagating this contract with each claim adds value beyond a complete typed, calibrated, lineage-aware, selection-aware, and simulation-checked evidence stack. The candidate must abstain on response null spaces and exact degeneracies rather than buy more observations that cannot identify the missing direction (C-229).

Derived observations inherit an interface and a support

A ratio, difference, derivative, normalized innovation, or latent coordinate is an observation operator, not a neutral rewrite of the raw signal. The operator record must name its input interface, reference state, causal window, positive-domain and saturation support, units, timestamp, calibration and any absolute information it removes. The statistic can legitimately change across the external-input, internal-state, downstream-readout and action interfaces: the Wnt/ β -catenin result is an internal fold readout (C-1543), while the two NF- κ B studies address different upstream and downstream arrows (C-1544, C-1545).

An invariant observation can make a transformed scale unidentifiable from that interface without proving that the information is universally destroyed. Adding a calibrated absolute observation or changing the output map can change recoverability (C-1549). The trajectory record therefore retains both the derived value and the smallest raw or calibrated side channel required by registered downstream targets.

Missingness is observed state, not a zero tensor

For modality r and decision time t , define $M_t^{(r)} \in \{0, 1, 2, 3, 4\}$ as a dimensionless status:

CODE	MEANING
0	present and inside calibration range
1	absent by experimental design
2	unavailable because of sensor or transport failure
3	present but occluded, saturated, clipped, or outside calibration
4	status unknown

The raw value is not imputed before status is retained. If an imputed value is needed, the model receives both the imputation and $M_t^{(r)}$. Training dropout is labeled as synthetic missingness and is sampled in contiguous outages as well as isolated frames. Confirmatory tests include natural missingness, failure, corruption, and combinations not seen during training.

Missingness may itself be informative—for example a tactile sensor becomes available only after contact. Evaluation therefore distinguishes performance gained from legitimate availability structure from shortcut prediction of a label or environment identifier.

Temporal alignment contract

Each modality has a calibrated clock offset δ_r in seconds and residual jitter scale σ_r in seconds. The corrected capture time is

$$\bar{t}_{r,k}^{\text{cap}} = t_{r,k}^{\text{cap}} - \delta_r.$$

Alignment does not mean forcing all modalities onto one frame rate. Encoders consume timestamped events or aggregate them inside declared causal windows. For a decision at time t , the available history is

$$\mathcal{H}_t = \{o_{r,k} : t_{r,k}^{\text{recv}} \leq t\} \cup \{a_j^{\leq t} : t_j^{\text{cmd}} \leq t\} \cup \{x_l : t_l^{\text{recv}} \leq t\} \cup \{y_n : t_n^{\text{end}} \leq t\},$$

where $\mathcal{H}_t$ is a set of records, not a unit-bearing scalar, and $a_j^{\leq t}$ contains only action fields measured by time t ; a future realized offset or actuator trace is not revealed with the earlier command. The equation admits a language event after its end time; a streaming system may instead add explicitly timestamped token prefixes. Outcome targets may use later events, but the state used to choose an action may not.

Alignment robustness is tested by adding known per-modality offsets and jitter, dropping clock-synchronization messages, and withholding one calibration version. A mechanism that works only at exact simulator step boundaries has not learned a robust temporal relation.

Observation, intervention, and consequence

An action token is useful only when its consequence can be distinguished from background change. Training retains four separable cases:

1. passive observation with no controlled action;
2. policy-chosen action with known or estimated behavior probability π_j ;
3. randomized or scripted intervention with known assignment; and
4. paired simulator branch from the same saved initial state with one changed action.

Only the third and fourth directly support an interventional comparison. Policy-chosen logs are confounded by the policy's state and require a declared system-identification, propensity, or model-based assumption. Exogenous events remain separate from realized action so the model cannot credit itself for an outside cause.

For a horizon $\Delta > 0$ seconds, an action-conditioned predictor estimates

$$p\theta(z_{t+\Delta} \mid z_t, \tilde{a}_{[t,t+\Delta)}, M^{\leq t}),$$

where $z_t \in \mathbb{R}^d$ is a dimensionless latent state of width d , θ is the parameter vector, $\tilde{a}[t, t+\Delta)$ is the realized action sequence over the interval with per-component physical units, and $M_{\leq t}$ is missingness history. The predictor is not called causal merely because action is an input. Causal interpretation depends on the data case and evaluation intervention.

This realized-action form is used for training and retrospective scoring. A prospective rollout conditions on a candidate command a and integrates over the calibrated distribution of actuator delay and realized action; it does not assume that a command is executed exactly.

Predictive state and action loop

The observation encoder forms

$$z_t = E_\phi(\mathcal{H}_t),$$

where E_ϕ is an encoder with dimensionless parameters ϕ and z_t is dimensionless. Separate heads predict future latent state, selected sensor targets, task outcomes, and calibrated uncertainty. Pixel or waveform reconstruction is used only when the task requires that detail; normalized latent prediction follows the narrower evidence in C-006.

For target-event set $\mathcal{G}$, a probabilistic prediction loss is

$$\mathcal{L}_{\text{pred}} = -\frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \log p_\theta(v_g \mid \mathcal{H}_{t_g}, \tilde{a}[t_g, t_g + \Delta_g)),$$

where $|\mathcal{G}|$ is a target count, v_g is a declared target under a declared quantization or likelihood on dimensionless normalized coordinates, t_g is target context time in seconds, Δ_g is its prediction horizon in seconds, and $\mathcal{L}_{\text{pred}}$ is nats/target. A normalized latent-distance loss may be reported in dimensionless units, but it is not added to negative log likelihood without a declared conversion weight.

At a decision, the system may act on the world, request a modality, ask a question, consult sourced memory, or wait. Let $b \in \mathcal{B}_t$ denote one such acquisition or intervention choice. A conventional one-step value-of-information policy chooses

$$b_t^* = \arg \max_{b \in \mathcal{B}_t} [V(b \mid \mathcal{H}_t) - \lambda_E E(b) - \lambda_T T(b)],$$

where V is expected task-utility improvement in declared utility units, $E(b)$ is full sensor, communication, and compute energy in joules, $T(b)$ is added latency in seconds, λ_E has units utility/joule, and λ_T has units utility/second. `wait` is an explicit zero-acquisition option. This is the strongest conventional null for claims that prediction error should route sensing or compute, as described under P-007 and in the engineering analogue audit.

Controlled observability and plant binding

The observation operator is partly controlled. Sensor pose, locomotor microstructure, emission timing, sampling density, and contact can change which state directions are observable (C-589–C-592). At the same time, body, tool, attachment, payload, compliance, and contact change how commands become motion and sensory consequences (C-599, C-603–C-605).

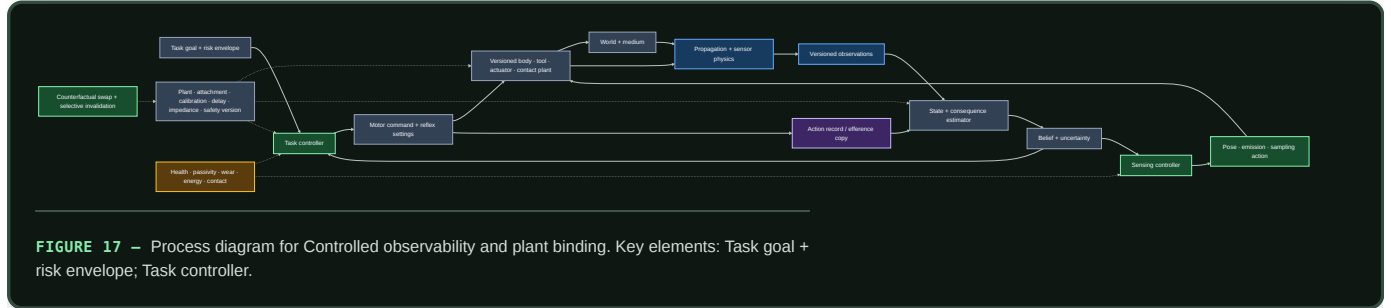
Separate the immediate task controller from the sensing controller:

$$u_t = \pi_{\text{task}}(b_t, \Pi_v), \quad a_t^{\text{sense}} = \pi_{\text{sense}}(b_t, \mathcal{O}_v, \Pi_v),$$

where b_t is a dimensionless belief state, task action u_t retains its physical units, sensing action a_t^{sense} retains pose, emission, sampling, or motion units, $\mathcal{O}_v$ is observation-contract version v , and Π_v is the controller–plant binding

$$\Pi_v = (B, T, A, C, S, J, Z, D, E),$$

with body/tool identity B , task T , attachment/payload A , contact model C , sensor/actuator calibration S , task Jacobian J , impedance/passivity envelope Z , delay model D , and safety/authority envelope E . Each field is typed, versioned, and linked to the controller, estimator, data, and tests that depend on it.



Editable source: [controlled-observability-plant-binding.mmd](#).

Two held tests share this record:

1. **Controlled observability.** Keep the task policy fixed while a sensing policy maintains information through pose, emission, or sampling. It must beat fixed excitation, random acquisition, one-step expected value of information, active SLAM, observability MPC, and dual control at equal action, time, risk, compute, and energy (C-596).
2. **Counterfactual plant binding.** Swap one body, tool, attachment, payload, sensor, delay, contact, or impedance field at a time. Dependencies should predict which estimator/controller state is safe to reuse and which must be invalidated. It must beat reset/retraining, unversioned adaptation, model banks, gain scheduling, operational-space impedance control, and adaptive MPC (C-606).

This composition avoids two attribution errors. Passive mechanics and compliance are credited before a controller (C-597); an action-correlated internal signal is not assumed to be a complete accurate forward model (C-601).

Opportunity- and history-qualified adaptive action

Comparative cognition adds a sharper condition to the observation contract: the same apparatus does not create the same task when bodies, sensors, prior training, rewards, demonstrators, handlers, or feasible actions differ. The evidence in C-804–C-841 spans tool manufacture, causal transfer, future preparation, event memory, uncertainty control, imitation, teaching, exploration, negative transfer, and central-versus-peripheral control. Its durable contribution is not a species ranking. It is a requirement to record the opportunity that made an outcome possible.

Extend each episode with

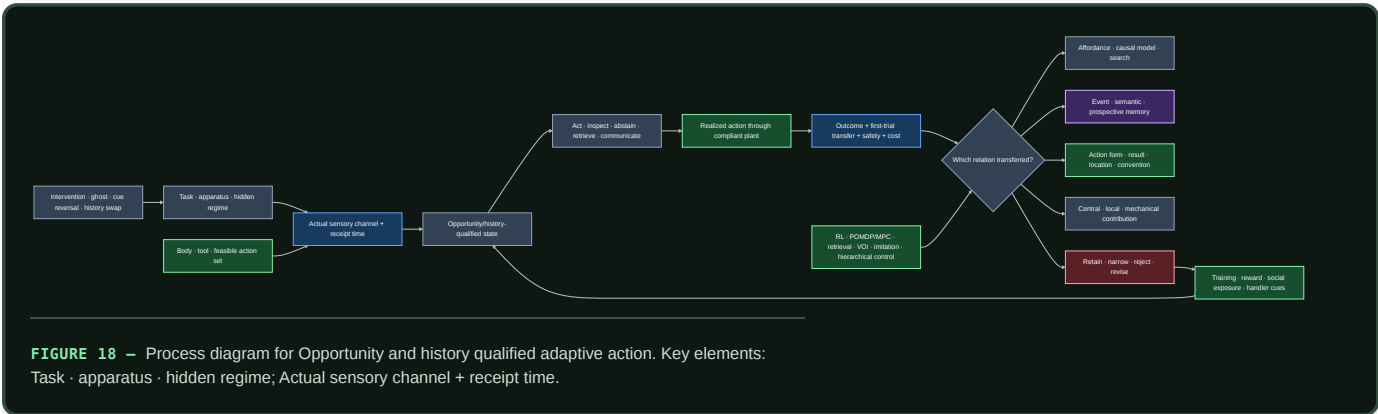
$$\Omega = (X, S, A, H, R, C, \tau, U),$$

where X is task and apparatus state, S is the sensory information actually available, A is the feasible realized-action set, H is prior training and social exposure, R is reward and stopping state, C is the intervention and control set, τ is elapsed time in seconds, and U is the sampling unit. The detailed definitions and units are maintained in the opportunity/history mathematics.

This record prevents several category errors:

1. **Terminal success is not transfer.** First-trial response after a material, geometry, cue, relation, body, or history change is retained separately from later test-time learning.

2. **One copied outcome is not one copied process.** Demonstrations are decomposed into action form, trajectory, end state, location, demonstrator, and functional relation. Ghost and result-only controls identify which component changed the learner.
3. **A future-directed action is not a unique memory mechanism.** Value tables, successor representations, planning, and retrieval remain nulls for delayed selection and resource reservation.
4. **Adaptive checking is not a unique uncertainty mechanism.** Public difficulty cues, response strength, calibrated confidence, selective prediction, and ordinary value of information must be beaten.
5. **Local control is not independent control.** Body mechanics, local sensing, peripheral feedback, central goal selection, override latency, and message traffic are measured separately.
6. **Repeated trials are not independent agents.** Episode, learner, dyad, group, site, rearing or training cohort, and population remain distinct sampling levels.

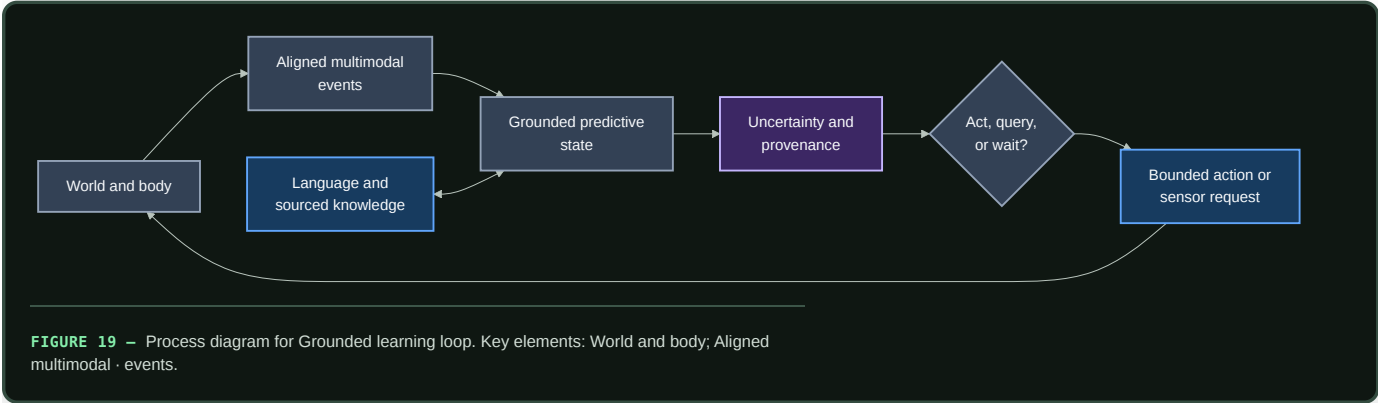


Editable source: `opportunity-history-qualified-action.mmd`.

Fixture F-003 turns the contract into eight separable tracks: manufacture transfer, causal inversion, future resource reservation, event memory, costed uncertainty, copy-target decomposition, social acquisition, and local/central/mechanical control. It compares the full composition with learned affordances and search, model-free and model-based RL, POMDP/MPC, semantic and episodic retrieval, selective prediction and value of information, imitation learning, central replay, and hierarchical control under one lifecycle boundary.

The fixture earns architectural credit only when opportunity qualification changes a conclusion, exposes a reproducible confound missed by the existing stack, or produces robust transfer beyond the complete nulls. Otherwise the episode contract remains useful instrumentation and the additional mechanism is retired.

Grounded learning loop



Editable source: `../assets/diagrams/grounded-learning-loop.mmd`.

Language attaches to grounded state

Language serves at least four different functions:

1. **reference**: name an object, relation, action, event, or trajectory span;
2. **instruction and query**: change the agent's task or request information;
3. **compression and communication**: summarize learned regularity for another time, module, or agent; and
4. **testimony**: introduce facts and abstractions not available through direct sensorimotor experience.

Reference is trained from the r_n links in language events, including negative and ambiguous referents. Instructions are available only from their receipt time and are treated as interventions that may change action. Testimony retains source, version, trust policy, and retrieval trace; it is not forced into a fictional physical observation. A concept can therefore combine direct experience, action–outcome evidence, and sourced symbolic knowledge without erasing which support came from where.

The staged design learns predictive physical state before or alongside a small language adapter, then compares late attachment, joint training, and language-first training under equal data and capacity budgets. Physical probes use paraphrases and unseen names; language probes include relations with no direct sensory referent. Demonstrated action and demonstrated outcome are separate targets, following the distinction motivated by C-064.

Uncertainty and provenance contract

Each prediction exposes a distribution or calibrated interval at the level at which action is chosen. Report negative log likelihood in nats/target, Brier score for declared categorical events, interval coverage in percent, and calibration error with binning fixed before evaluation. Calibration is broken out by modality status, prediction horizon, intervention type, environment, and linguistic versus direct evidence.

A single confidence scalar does not replace modality-specific uncertainty. The system records whether uncertainty arose from noisy observation, missing modality, disagreement among plausible dynamics, unseen composition, ambiguous language, or untrusted testimony when the estimator supports that distinction. If it does not, the output remains an undifferentiated predictive uncertainty.

Every training and evaluation event resolves through $p_{r,k}$, p_j , or p_n to:

- raw or generated source identifier and checksum;
- simulator, device, actor, and policy version;
- capture and receipt clock domains;
- calibration and transformation chain;
- missingness and corruption operations;
- split assignment and counterfactual parent; and
- license or access policy where applicable.

Answers and plans can then cite direct observation, intervened outcome, retrieved testimony, or model rollout separately. Provenance does not make a prediction correct; it makes its evidential path inspectable and reversible.

Strongest conventional baselines and nulls

ID	BASELINE	QUESTION IT ANSWERS
B0	Text-only model with matched language tokens and capacity	Does language alone explain the reported transfer?
B1	Passive multimodal encoder on the same recorded sensor events	Does temporal multimodality help without action conditioning?

ID	BASELINE	QUESTION IT ANSWERS
B2	Synchronized masked/contrastive multimodal predictor	Does ordinary cross-modal prediction explain the gain?
B3	Standard action-conditioned latent state-space/world model	Does the proposed data and uncertainty contract add value beyond a conventional world model?
B4	Behavioral cloning and action-prediction model	Is copying the logged policy enough without predicting consequences?
B5	Calibrated linear/nonlinear system-identification model on tasks where its assumptions apply	Is learned grounding better than an established dynamics estimator?
B6	Random, entropy-seeking, and one-step EVSI acquisition policies	Does active sensing beat ordinary exploration and information purchasing?
B7	Oracle-state predictor/controller	How much error comes from partial observation rather than dynamics or control?
B8	Intact architecture with time, action, modality, or language links shuffled	Does performance depend on the claimed alignment and causal structure?

B7 is an upper bound, not a superiority baseline. Comparisons match trajectory seeds, raw sensor exposure where logically possible, model capacity, optimizer updates, action count, sensor acquisitions, context bytes, latency ceiling, and lifecycle energy. The active-data question and the action-conditioning question are evaluated separately: hiding action tokens in the same log is not the same experiment as allowing an agent to collect a different log.

Staged experiment

Stage 0 — Data and clock qualification

Build a deterministic replay harness before training a world model. It must reproduce event order, verify checksums and split isolation, expose capture and receipt time, reconstruct commanded versus realized actions, and audit every missingness code. Inject known offsets, jitter, packet delay, dropped intervals, and actuator lag. Reject a dataset version when these perturbations cannot be recovered or bounded from its metadata.

Stage 1 — Passive action-conditioned prediction

Use a controlled compositional environment with vision, audio, proprioception, touch/force, and optional depth. Objects vary independently in shape, color, mass, surface friction, and containment. Actions include look, approach, push, lift, rotate, drop, occlude, and tool contact. The held-out split contains:

- unseen combinations of known object properties;
- new visual textures and camera placement;
- changed mass or friction under familiar appearance;
- action delays and partial actuator failure; and
- one-modality and multi-modality outages.

Training, development, and test split by episode seed, object combination, and environment configuration—not by frames. Simulator state s_t^* is stored for evaluation and B7 only.

Train B1–B5 and the candidate predictive state on the same logged trajectories. Compare intact realized actions with actions masked, time-shifted, and permuted within episode. Primary outcomes are future-state and sensor NLL in nats/target, physical-property probe error in declared units, held-out intervention error, control success percent with a frozen small controller, calibration, bytes, and joules/qualified episode.

Stage 2 — Paired interventions and active acquisition

From a saved initial state, run paired branches differing in one action or sensor request while holding the environment seed fixed. Separately allow each active policy the same maximum world actions, sensor acquisitions, wall time, and energy. Compare random exploration, entropy seeking, one-step EVSI, and the candidate policy.

Measure prediction change in the correct physical direction, intervention effect error, successful disambiguations/action, downstream task utility, unsafe interventions, sensor byte, action energy, total joules, and latency. This stage tests the open issue in C-022 and the targeted-exploration translation from C-062.

Stage 3 — Language attachment

Attach language using trajectory-span and referent links. Compare late attachment to a frozen physical core, joint training from initialization, and language-first pretraining followed by grounded data. Equalize language tokens, grounded episodes, trainable parameter count, optimizer updates, and tuning budget.

Evaluate new names for known referents, paraphrased instructions, ambiguous reference, imitation versus outcome emulation, physical questions about unseen compositions, testimony-only facts with changed source versions, and removal of language at control time. Report whether language improves sample efficiency without becoming the only route to physical success.

Stage 4 — Asynchronous physical transfer

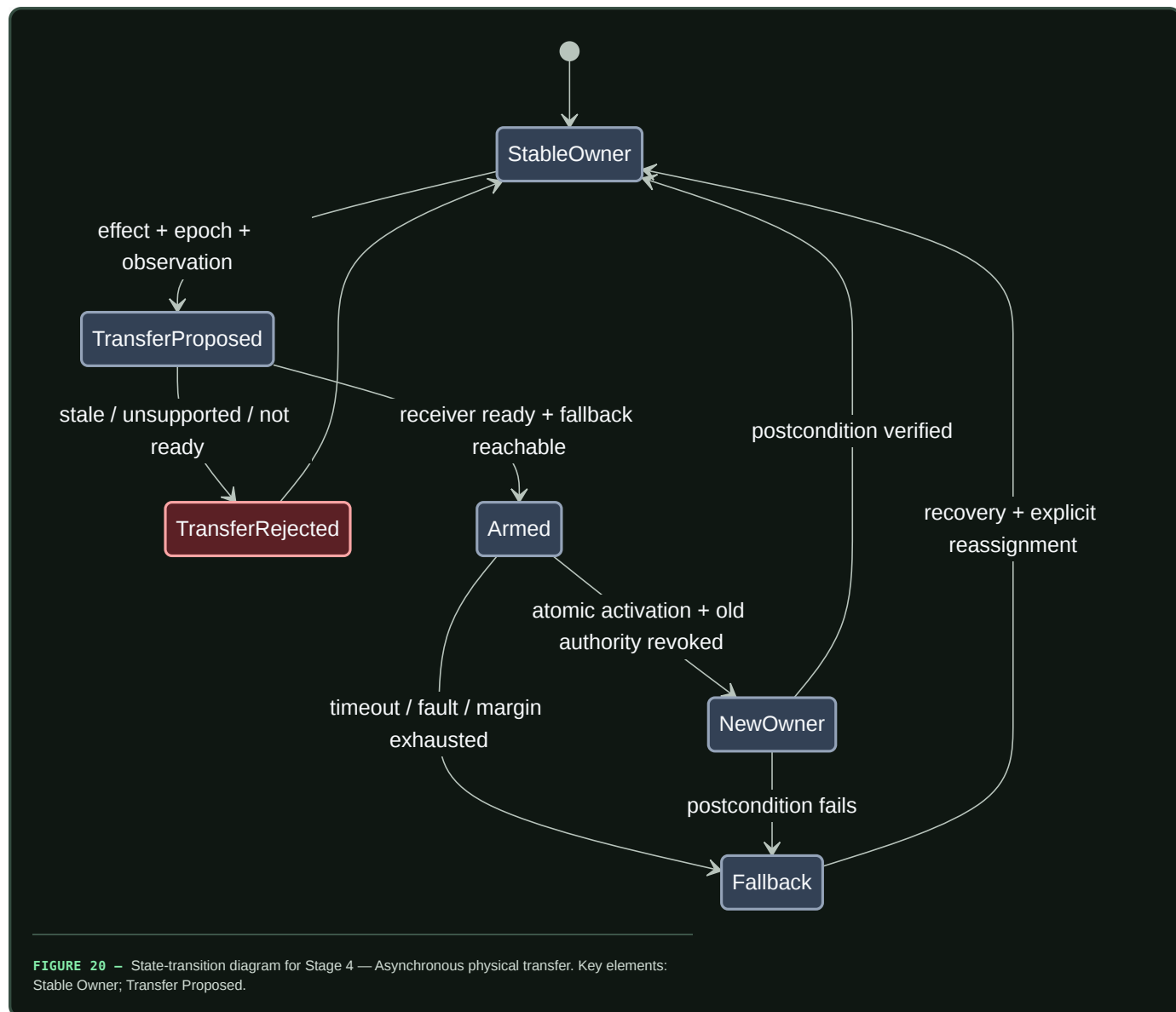
Move the frozen comparison to one named real sensorimotor platform. Preserve native clocks, packet delay, actuator mismatch, calibration changes, occlusion, and sensor failure. Recalibrate energy at device and node boundaries using the energy evaluation contract. Simulation-to-real transfer, recovery, and negative results are reported by intervention and missingness stratum rather than one average.

The real platform adds two contracts that simulation accuracy cannot supply. First, estimation is qualified by integrity. An illustrative position protection level is

$$PL = K_{\text{int}}\sigma_{\text{pos}}, \quad \text{usable only when } PL \leq AL,$$

where PL , position standard uncertainty σ_{pos} , and alert limit AL are in metres, while K_{int} is dimensionless and tied to a declared fault and risk allocation. A learned confidence score does not become integrity without bias bounds, dependence, coverage, and time-to-alert (C-446).

Second, a controller handoff must preserve exclusive effective authority while the world keeps moving. The transfer binds the controlled effect, epoch, outgoing and incoming controller, actual and pending mode, observation basis, headroom, time-to-boundary, faults, outstanding commands, fallback reachable set, expiry, acknowledgement, atomic activation, and observed postcondition.



Editable source: [asynchronous-authority-transfer.mmd](#).

This is a refinement of Candidate 012, not a new autonomy layer. It must beat mode annunciation, interlocks, epoch leases, runtime assurance, explicit arbitration, Candidate 014, and the recoverable-initiative contract under split-brain, stale-command, no-owner, delayed-link, common-mode, and failed- postcondition injections (C-447–C-461).

Efficiency mechanism

Grounding adds sensors and interaction, so its efficiency claim is conditional: the learned state must avoid more future sensing, compute, data movement, or failed action than it costs to acquire and maintain.

The candidate levers are:

- **event-driven encoding:** update modality state on timestamped change rather than resampling every channel at the highest rate;
- **latent prediction:** predict task-relevant state without reconstructing every high-bandwidth detail;
- **local preprocessing:** reduce raw sensor traffic near the source while retaining calibration and uncertainty metadata;

- **conditional acquisition:** power, transmit, or process another modality only when its expected utility exceeds energy and latency cost;
- **temporal reuse:** carry forward stable state with an uncertainty increase instead of re-encoding unchanged observations; and
- **language compression:** communicate grounded referents and relations when sending the underlying sensor history is unnecessary.

Every experiment reports sensor-on time in seconds, sensor energy in joules, raw and processed bytes by boundary, acquisitions/episode, world actions/episode, encoder and world-model operations, latency in seconds, controller energy, maintenance/calibration energy, and total lifecycle joules/qualified episode. A lower model FLOP count is not a grounding-efficiency result when sensor collection or interaction dominates.

The relevant efficiency comparison is a frontier over task utility, risk, latency, energy, and physical interaction. One-step EVSI is the minimum null for conditional acquisition. A passive model trained on a larger recorded dataset is an additional null when interaction energy, not online autonomy, is the proposed benefit.

Evidence status

- Sparse natural-image representation under C-002: **established for the cited model and images.**
- Joint-embedding image prediction under C-006: **established for the cited image experiments.**
- Active sensor mechanics under C-022: **established for the measured whisker task.**
- Targeted exploration under C-062, pedagogical narrowing under C-063, and imitation/emulation difference under C-064: **established for the scoped behavioral tasks.**
- A generally useful predictive-coding abstraction under C-005: **plausible**, not a complete grounding theory.
- Robust general physical concepts from the integrated multimodal curriculum under C-007: **speculative.**
- Navigation integrity, protection-envelope scope, fault-stage separation, degraded/fallback distinctions, and assurance boundaries under C-445–C-460: **established or plausible within the audited vehicle domains.**
- Validated asynchronous authority transfer under C-461: **speculative.**
- Opportunity, history, causal-transfer, event-memory, social-learning, and distributed-control observations under C-804–C-840: **established, plausible, or disputed only within their declared tasks**; the integrated action contract under C-841 is **speculative.**

Speculative extensions

Controllability curriculum

Order early tasks by what an agent can reliably change and observe: persistence and contact, motion and occlusion, material response, tools and other agents, then linguistic abstraction. Curriculum order becomes an ablation rather than an assumed developmental law.

Learned morphology and sensor placement

Jointly adapt sensor placement, body parameters, and control only after the fixed-body grounding loop is understood. The simulation evidence under C-029 motivates the question, while lifecycle energy and reality-gap robustness remain required outcomes.

Socially supplied interventions

Treat demonstration, correction, question answering, and pedagogy as actions that change the learner’s evidence policy. The system can preserve uncertainty after a demonstration and deliberately test affordances not shown, rather than assuming instruction is exhaustive.

Counterfactual branch memory

Store compact paired branches from matched initial states as high-value causal records. Promotion into long-term memory depends on reuse and provenance, not on intervention novelty alone.

Failure modes

SIGNATURE	INTERPRETATION AND REQUIRED RESPONSE
Action shuffling leaves held-out intervention error unchanged	the representation uses passive correlation; remove the action-grounding claim
Timestamp shuffling or large clock offsets cause no degradation	static identifiers or leakage dominate; audit split and alignment
Exact simulator timing wins but calibrated jitter collapses performance	temporal relation is brittle; retain time uncertainty or revise the encoder
Missingness code alone predicts task labels	acquisition policy or dataset artifacts leak the answer; rebalance and report the shortcut
One modality receives nearly all gradient or attention and its removal collapses all tasks	shared state has become single-modality state; enforce and test modality-specific competence
Commanded action predicts outcomes but realized action does not	actuator or logging shortcut; train and evaluate on realized intervention
B3 or B5 matches the candidate	conventional world modeling or system identification explains the result; use the simpler baseline
One-step EVSI matches active acquisition	no new sensing principle is supported; retain EVSI as the controller
Language removal destroys physical control learned before language	linguistic shortcut or catastrophic overwriting; separate adapters and replay physical probes
Fluent reports disagree with action-conditioned predictions	language head is not constrained by world state; expose uncertainty and provenance by source
Counterfactual performance is high only when oracle state leaks into input	the partial-observation problem remains unsolved; remove oracle fields and rerun
Active training wins only because it observed more frames or spent more action energy	data or resource exposure explains the result; compare at matched events, actions, and joules
Sensor and interaction energy erase compute savings	narrow the result to representation quality; do not claim system efficiency
Calibration fails specifically under missing modalities or novel interventions	uncertainty is not deployment-valid; block autonomous escalation in those strata
Low state-estimation error coexists with integrity alerts that are late or absent	average accuracy is not safe-to-use evidence; bind fault hypotheses, alert limit, and deadline
Handoff produces overlapping owners, a control gap, or stale queued effects	transfer is not atomic or command expiry is incomplete; revoke, fall back, and preserve the trace
Raw task success changes after body, cue, training, reward, handler, or site changes	the opportunity record is incomplete; stratify and rerun before attributing a mechanism
A socially exposed policy reaches the same end state but changes no novel action form	result, affordance, or location learning explains the effect; remove the imitation claim
A local controller wins until passive mechanics, local hardware, or communication are charged	the control-locus claim is misattributed; keep the cheaper complete baseline

Measurable predictions

H-G1 — Action-conditioned consequence

With identical logged sensor events, intact realized actions will improve held-out intervention NLL and physical-effect error over masked, shifted, and permuted actions. If only reconstruction or in-distribution probe accuracy improves, the result does not support action grounding.

H-G2 — Temporal alignment

Models trained with explicit capture/receipt time and calibrated clock uncertainty will degrade more gradually under held-out offset, jitter, and packet delay than fixed-step concatenation. The comparison reports error as a function of injected milliseconds and identifies the offset at which the quality envelope is crossed.

H-G3 — Missing modalities

Typed missingness plus modality-specific predictive heads will preserve a better quality–calibration frontier under contiguous sensor outages than zero imputation or independent encoders fused only at the output. Performance is reported separately for absence, failure, occlusion/corruption, and unknown status.

H-G4 — Active evidence collection

At the same action, acquisition, latency, and joule budgets, a learned active policy will require fewer interventions to identify held-out mass, friction, containment, or causal relations than random and entropy-seeking exploration. It supports a new mechanism only if it also improves beyond calibrated one-step EVSI; otherwise EVSI becomes the accepted implementation of P-007.

H-G5 — Compositional physical transfer

Action-conditioned predictive state will yield lower sample count to a fixed control-success threshold on unseen property combinations and dynamics than text-only, passive multimodal, and behavioral-cloning baselines of matched capacity. The threshold, maximum episodes, and unsuccessful runs are fixed before evaluation.

H-G6 — Language attachment

Late or joint language attachment with explicit referent and provenance links will improve new-name, paraphrase, and testimony-version tests without reducing pre-language physical probes beyond a preregistered equivalence margin. If language-first training alone matches intervention transfer at equal grounded episodes, the additional grounding curriculum has not earned its cost.

H-G7 — Calibrated abstention

Prediction uncertainty will increase under unseen interventions, multi-sensor failure, and ambiguous reference, and calibrated abstention or acquisition will reduce high-cost physical errors at a declared false-abstention rate. Average confidence without stratum calibration does not satisfy this prediction.

H-G8 — Lifecycle efficiency

Conditional sensing, event-driven updates, and latent prediction will reduce sensor byte, processed byte, and node joules/qualified episode at matched task utility, risk, and latency relative to always-on sensing and fixed-rate encoding. The claim is rejected if the confidence interval for lifecycle energy includes the preregistered equivalence margin after collection, calibration, language, maintenance, and failed-action costs are included.

H-G9 — Integrity-qualified authority transfer

Under delayed links, boundary misclassification, degraded sensors, outstanding commands, common-mode faults, and failed postconditions, a validated transfer record will reduce double-owner, no-owner, stale-command, and unsafe-boundary time beyond mode annunciation, interlocks, epoch leases, runtime assurance, ordinary arbitration, and the HCI/observation contracts at equal sensing, reserve, communication, training, review, latency, and joule budgets. A tie merges its useful fields into those conventional mechanisms.

H-G10 — Opportunity-qualified transfer

At equal episodes, interventions, action feasibility, search, storage, communication, human effort, and lifecycle joules, explicit opportunity and history state will improve first-trial functional transfer and calibration under held material, causal-relation, body/tool, acquisition-history, and social-channel changes relative to an otherwise identical learner. The claim is rejected if the gain disappears after actual sensor cues, feasible actions, prior training, reward schedule, handler/site, or later test-time learning are modeled, or if the complete affordance/search, planning, retrieval, VOI, imitation, replay, and hierarchical-control stack matches the result.

Promotion requires H-G1 plus at least one transfer result from H-G3–H-G6, valid uncertainty under H-G7, and no lifecycle contradiction under H-G8. Stage-4 authority-transfer claims additionally require H-G9. A comparative-cognition-derived transfer claim additionally requires H-G10. A positive result remains scoped to its trajectory schema, intervention family, environment shift, hardware boundary, and evaluated language role.

Representative adaptive performance

Scope

This chapter turns the sports expertise, adaptive performance, and team coordination audit into a system-wide performance contract. A result is interpretable only when it travels with the information that was available, the actions that were feasible, the history that produced the policy, the opponent and team it faced, the feedback it received, its current resource and damage state, and the selection process that determined which systems were observed.

The contract constrains:

1. sensing and observation provenance;
2. routing and allocation under deadlines;
3. action and closed-loop control;
4. memory of practice, opponents, damage, and recovery;
5. curriculum and online adaptation;
6. assurance, degradation, and staged return;
7. coordination and communication; and
8. population-level evaluation and lifecycle accounting.

It connects the sensorimotor grounding chapter, sparse predictive computation, memory and consolidation, system synthesis, and the energy model. Its executable specification is Fixture F-006, with notation and derivations in the mathematical contract.

Biological observation

Perception is coupled to a possible response

Experts can extract predictive information from early opponent kinematics, but the useful cue depends on the task, feature, opponent, prior, viewing window, and response mode (C-926–C-930). Predicting a label, moving a joystick, initiating a partial movement, and intercepting a physical event under its real deadline are therefore different experiments. Temporal or spatial occlusion can locate when or where information becomes useful; it does not identify a unique representation.

Training transfers more reliably when relevant information and action coupling are preserved, but “representative” is not synonymous with visually realistic (C-931–C-932). Timing, feasible action, opponent response, consequence, and resource state can matter even when surface appearance changes little.

Practice is not one outcome

Practice variability can improve learning when it explores a relevant task dimension at a useful challenge level; the effect is not universal, and larger perturbations do not automatically help (C-933–C-936). Accumulated practice correlates with expertise, yet amount, type, opportunity, selection, survival, access, and retrospective classification are entangled (C-937–C-939).

This creates three separate questions:

1. Does performance improve while repetition, instruction, and feedback are present?
2. Does the change remain after a declared delay without that scaffold?
3. Does it survive a change in cue, opponent, body, task, rule, deadline, or resource state?

The answers are practice performance, delayed retention, and transfer. They must not be substituted for one another.

Performance follows a frontier, not a single score

Movement speed and accuracy trade off under task-specific conditions, and throughput can appear stable while speed and error move in opposite directions (C-940–C-941). Risk, physical work, latency, and recovery add further axes. Optimizing only accuracy can hide slower action; optimizing only latency can hide errors or unsafe events; optimizing only completed work can hide depletion.

Fatigue changes the state from which an action is controlled. Opponent behavior, remaining work, feedback, recovery, and sleep can alter output, but effects remain task- and outcome-specific (C-942–C-948). Internal load, external load, readiness, damage, and injury risk are different constructs. A workload ratio or a single measurement cannot certify an individual system's safe capability (C-949–C-950).

Recovery and return are staged and reversible

Return to participation, return to the full task, and sustainable recovery of prior performance form a staged continuum. Time and functional evidence can carry nonredundant information, while an apparently high function score can be misleading after early return (C-951–C-953). Each promotion exposes the recovering system to more load and therefore creates new evidence. Deterioration must permit regression to a safer stage.

Teams coordinate through information and adjustment

Joint practice can increase predictive knowledge, and shared displays can change later team state. Yet synchrony is only an observable; similar movement can follow common input without reciprocal coordination (C-954–C-957). A coordination claim needs perturbation, lagged compensation, reduced task error, turnover, and cross-play. A shared-information claim additionally needs records of messages, observations, beliefs, and their discrepancies.

Deceptive behavior exploits an observer's cue policy. Expertise can reduce some susceptibility without removing it, and confidence may rise while the judgment becomes wrong (C-958–C-960). Feedback effects are likewise conditional on content, timing, task, autonomy, and whether the outcome is immediate performance or later learning (C-961–C-962).

Selection changes the population it measures

Relative age, maturity, present capability, coaching access, playing time, opponent quality, and later observation are coupled by selection. A talent system can amplify early differences and then mistake the resulting population for evidence that its original ranking was correct (C-963–C-965). Excluded systems have missing counterfactual careers precisely because the selection policy withheld the experience needed to produce them.

Finally, metabolic estimates and wearable estimates do not equal complete system energy. Human attention, facilities, sensors, equipment, computation, recovery, failed trials, maintenance, medical or safety work, and displaced opportunity alter the efficiency comparison (C-966–C-968). The transferable result is therefore an evaluation contract (C-969).

Proposed AI translation

Carry the state that makes a comparison valid

For agent i in episode e , preserve

$$\mathcal{K}_{i,e} = (X_e, O_{i,e}, A_{i,e}, H_{i,e}, M_e, F_{i,e}, R_{i,e}, D_{i,e}, G_{i,e}, C_e, U_e, B_e),$$

where:

- X_e is physical state, task, rules, deadline, consequence, and hidden regime;
- $O_{i,e}$ is information actually received, including source, support, latency in seconds, occlusion, noise, loss, and calibration;
- $A_{i,e}$ is the feasible action set under the current body or actuator, equipment, authority, rate, range, and safety limits;
- $H_{i,e}$ is timestamped practice, feedback, opponent, teammate, damage, exclusion, reward, and opportunity history;
- M_e is opponent and teammate composition, role, policy history, turnover, and communication topology;
- $F_{i,e}$ is feedback identity, information in bits, delay in seconds, and provider;
- $R_{i,e}$ is the resource and fatigue vector in its native units;
- $D_{i,e}$ is damage or fault state, diagnostic uncertainty, protected envelope, and return stage;
- $G_{i,e}$ is the selection and opportunity policy;
- C_e is the intervention, paired control, counterfactual seed, and stopping rule;
- U_e is the sampling unit, such as action, episode, agent, dyad, team, site, or cohort; and
- B_e is the complete ceiling in events, bytes, seconds, person-hours, joules, damage, unsafe events, replacements, and opportunity.

For method m and literal outcome k , the comparison target is

$$Q_{m,k}(\mathcal{K}) = \mathbb{E}[Y_k \mid do(m), \mathcal{K}],$$

where Y_k uses the registered unit for outcome k . If a baseline and a new method receive different observations, actions, histories, opponents, feedback, resources, or selection opportunities, $Q_{m,k} - Q_{b,k}$ does not isolate method m from baseline b .

Treat representativeness as an inspectable vector

For training distribution P_{tr} and target distribution P_{te} , record

$$\mathbf{d}_{\text{rep}} = (dO, dA, dT, dM, dF, dR, dD, dG),$$

where the components compare actual observation O , feasible action A , deadline and consequence T , teammate/opponent state M , feedback F , resource state R , damage and return state D , and selection policy G . Each d_z is a declared divergence between the corresponding training and target distributions: dimensionless for a statistical divergence or in the declared ground-cost unit for optimal transport. The components remain visible; one weighted “realism” number would conceal which relation transferred.

Keep outcomes separate

OUTCOME	MINIMUM EVIDENCE	COMMON FALSE PROXY
anticipation	proper score by cue time and opponent; calibration; commitment latency	expert label or reaction time
physical interception	success, onset, trajectory, endpoint error, safety	video or joystick judgment
cue use	randomized removal, neutralization, conflict, or timing intervention	gaze or saliency
practice	acquisition curve by attempt and exposure time	end-of-practice score as retained learning
retention	delayed scaffold-free test	immediate post-practice score

OUTCOME	MINIMUM EVIDENCE	COMMON FALSE PROXY
transfer	first target trial and full source-to-target matrix	later target learning
exploration	action and outcome information, feasible coverage, later utility, cost	raw variance or entropy
adaptability	perturbation loss, recovery time, overshoot, recurrence, damage	stationary accuracy
pacing/readiness	output trajectory, state calibration, admissible actions, abstention	elapsed workload or one score
staged return	false promotion/withholding, dwell, recurrence, rollback, availability	calendar time or one test
coordination	perturbation-conditioned compensation, task stability, cross-play, repair	synchrony or proximity
shared information	messages, observations, predictive beliefs, discrepancies	similar behavior
deception	matched causal contrast, calibration, exploitability, regret, adaptation	confidence or surprise
talent prediction	prospective out-of-cohort calibration and counterfactual opportunity	current rank or selected-cohort accuracy
complete efficiency	protected outcomes plus all resource and harm axes	device energy or success/trial

Separate useful exploration from noise

For action variable A , reached-outcome variable Z , and method m , measure

$$\mathcal{X}_m = (H_m(A), H_m(Z), I_m(A; Z), K_m, Q_m^{\text{tr}}, \mathbf{C}_m),$$

where both entropies H_m and mutual information I_m are in bits, K_m is the dimensionless fraction of the registered feasible region covered, Q_m^{tr} is later transfer in the task's literal unit, and $\mathbf{C}_m$ is the complete cost vector. More action entropy is useful only when it improves outcome information, transfer, or later task value at an acceptable cost.

An adaptive curriculum therefore needs the initial skill and history stratum, the perturbed task dimension, perturbation dose, retry cost, feedback bits, criterion exposure, and delayed tests. This directly constrains Candidate 004 and the memory boundary in the consolidation chapter.

Route and act through resource-qualified state

The runtime policy should expose the state on which pacing and authority depend:

$$a_t = \pi_m(o_{\leq t}, \hat{r}_t, \hat{d}_t, s_t, \hat{\pi}_{\text{opp},t}, \hat{b}_{\text{team},t}, f_t),$$

where a_t is the commanded action or power target in its native unit, $o_{\leq t}$ is causally received observation history, $\hat{r}_t$ is the estimated resource/fatigue vector, $\hat{d}_t$ the estimated damage state, s_t remaining work in metres, seconds, events, or joules, $\hat{\pi}_{\text{opp},t}$ the opponent-policy estimate, $\hat{b}_{\text{team},t}$ the teammate-state estimate, and f_t available feedback. Every estimate and channel receives a separate ablation.

This state conditions sensing, sparse routing, action authority, memory access, and recovery. It sharpens Candidate 002, Candidate 006, Candidate 007, and Candidate 012.

Compare a frontier, not a winner

For one task, retain

$$\mathcal{P}_m = (T_m, \epsilon_m, p_m^{\text{unsafe}}, E_m^{\text{life}}, H_m^{\text{human}}),$$

where T_m is latency in seconds, ϵ_m task error in its declared physical or task unit, p_m^{unsafe} dimensionless unsafe-event probability, E_m^{life} lifecycle energy in joules, and H_m^{human} role-stratified effort in person-hours. A method is more efficient only through a preregistered utility or a Pareto improvement with non-inferiority on protected outcomes.

Make readiness an action envelope

At decision time t , admissible actions are

$$\mathcal{A}_t^{\text{ready}}(\alpha) = \{a \in A_t : \Pr(Z_{t:t+h} \in \mathcal{Z}_{\text{safe}} \mid a, \mathcal{I}_t) \geq 1 - \alpha\},$$

where A_t is the currently feasible action set, $Z_{t:t+h}$ the multidomain outcome vector over horizon h in hours, $\mathcal{Z}_{\text{safe}}$ the registered safe envelope, $\mathcal{I}_t$ information available at time t , and α the dimensionless tolerated risk. An empty envelope triggers abstention or escalation.

Let $g_t \in \{0, 1, 2, 3, 4\}$ denote protected, modified, controlled, full-load, and adversarial operation. Promotion requires the next stage's envelope; a violation requires that $g_{t+1} < g_t$ remain possible. This turns staged return into a concrete test for Candidate 009 rather than a one-time health classification.

Distinguish coordination from shared input

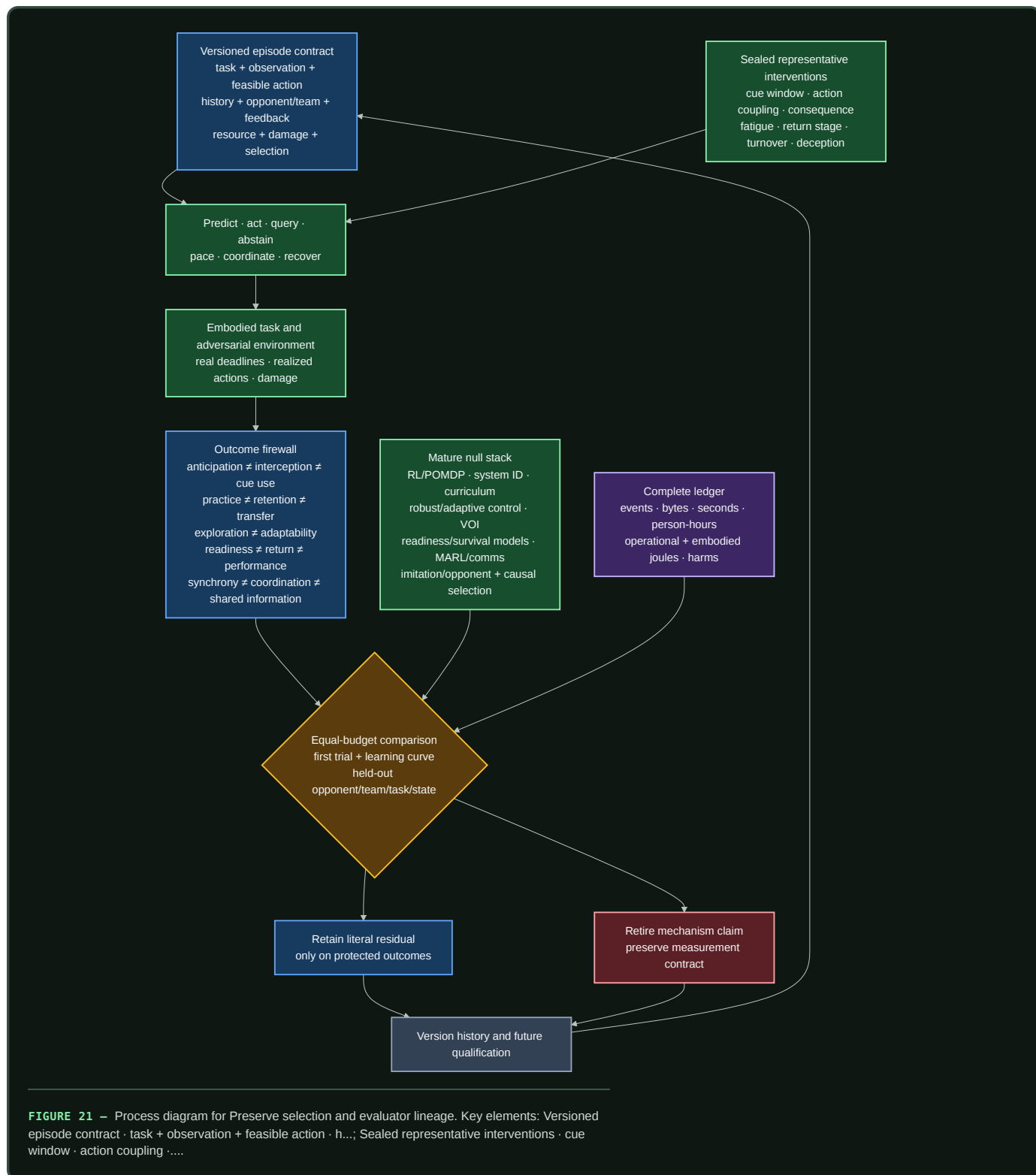
For agent contributions $u_i(t)$ and $u_j(t)$, apply perturbation $do(\eta_i)$ to agent i and estimate

$$\Gamma_{ij}(\ell) = \text{Cov}(\Delta u_i(t), \Delta u_j(t + \ell) \mid do(\eta_i), X_t),$$

where lag ℓ is in seconds, X_t is task state, u_i and u_j retain their native contribution units, η_i is a registered intervention in the unit of i 's action or state, and Γ_{ij} has the product unit of the two contributions. A useful response must also reduce error in the protected task variable without increasing risk. Teammate turnover, role reassignment, message ablation, common-input controls, and never-co-trained cross-play separate reciprocal adjustment from synchrony. The nephron-coupling boundary adds the same requirement: coherence alone is not direct coupling, information transfer, or functional benefit (C-1495).

Preserve selection and evaluator lineage

Population evaluation carries the policy that granted training, observation, feedback, compute, role, and survival opportunity. Selected and rejected systems remain in the analysis, with censoring and missing outcomes explicit. Additional opportunity near selection thresholds is randomized when admissible or handled with a declared causal design. This extends Candidate 019 and the observation lineage in Candidate 014.



Editable source: representative-resource-qualified-performance.mmd.

Efficiency mechanism

The contract can improve efficiency through five measurable effects:

1. **Less proxy optimization.** Coupled perception/action tests prevent spending training and inference resources on a symbolic score that does not improve the physical or operational task.

- 2. Higher-value variation.** History-qualified curricula direct perturbations toward task-relevant uncertainty instead of buying undirected entropy.
- 3. Resource-aware allocation.** Routing, sensing, and authority respond to remaining work, fatigue, damage, opponent state, and recovery rather than treating every episode as fresh.
- 4. Reversible exposure.** Staged return seeks a better availability--recurrence frontier than either permanent exclusion or an immediate full-load restart.
- 5. Complete comparison.** Population selection, human effort, failure, recovery, and embodied resources are charged before a claimed saving is accepted.

Lifecycle energy for method m is

$$E_m^{\text{life}} = E_m^{\text{train}} + E_m^{\text{infer}} + E_m^{\text{sense}} + E_m^{\text{act}} + E_m^{\text{comm}} + E_m^{\text{facility}} + E_m^{\text{recover}} + E_m^{\text{maint}} + E_m^{\text{emb}},$$

where every term is in joules under one declared service interval. The terms are training, inference, sensing, actuation, communication, facility, recovery, maintenance, and amortized embodied energy. Human design, demonstration, coaching, labeling, tuning, monitoring, repair, safety, and medical effort are reported separately in person-hours. A lower device-energy reading does not establish a lifecycle saving.

Evidence status

COMPONENT	EVIDENCE BOUNDARY	ARCHITECTURAL STATUS
early cue use and anticipation	task-, cue-, opponent-, and response-specific human studies; C-926–C-930	established scoped observations; transferable artificial mechanism unassigned
representative information/action coupling	transfer studies distinguish action coupling from visual similarity; C-931–C-932	plausible evaluation constraint
variability and practice	benefits depend on task dimension, dose, history, and outcome; C-933–C-939	established heterogeneity; adaptive curriculum residual speculative
speed--accuracy frontier	scoped task relations; C-940–C-941	established need for multi-axis reporting
fatigue, pacing, load, and readiness	outcome-specific interventions and measurement critiques; C-942–C-950	resource state is necessary metadata; learned controller untested
staged return	staged clinical framework and nonredundant criteria; C-951–C-953	plausible assurance translation; ordinary staged rollout remains the null
coordination and shared information	practice, shared display, synchrony, and coordination studies; C-954–C-957	perturbation and cross-play are evaluation requirements
deception and feedback	bounded effects under declared tasks; C-958–C-962	adversarial calibration requirement
selection bias and talent prediction	relative-age, maturity, and prospective-validity evidence; C-963–C-965	prospective causal evaluation requirement
complete efficiency	system-boundary and measurement evidence; C-966–C-969	required accounting contract; net advantage unknown

The proposed composition remains a benchmark target until it beats the complete ordinary stack in F-006. A complete stack includes calibrated prediction and retrieval; model-free and model-based RL; POMDP/MPC; system identification; domain randomization and curriculum; robust/adaptive control; value of information and selective prediction; workload, readiness, survival, canary, and rollback models; multi-agent control and explicit communication; imitation and opponent models; and causal selection models.

Speculative extensions

Resource-conditioned sparse routing

Routing could condition on expected remaining work, resource uncertainty, damage, and safe fallback rather than only token or task features. It must beat a conventional estimator plus constrained control at equal sensing, compute, reserve, and failure allowance.

Retention-aware curriculum control

A curriculum controller could value a perturbation by delayed retention and held-out transfer rather than practice loss. It must beat fixed augmentation, domain randomization, active learning, Bayesian experimental design, novelty, quality-diversity search, and automatic curriculum at equal attempt, perturbation, feedback, evaluator, time, and energy budgets.

Deception-calibrated opponent memory

Opponent memory could retain policy versions, cue conflicts, confidence, change points, and abstention value. It must improve calibration and adaptation on new deceptive opponents beyond Bayesian opponent models, fictitious play, self-play, recurrent policies, retrieval, and conformal abstention.

Turnover-resilient communication

Teams could compress communication after shared practice while retaining typed repair, acknowledgements, and belief discrepancy when membership changes. Lower message volume counts only if cross-play, repair latency, task quality, and safety survive.

Counterfactual development policies

Population management could compare selection with broad-development or threshold-lottery policies and explicitly model opportunity-mediated outcomes. It must predict later capability in new cohorts and recover false negatives without hiding attrition or development cost.

Failure modes

FAILURE	OBSERVABLE SIGNATURE	REJECTION OR CONTAINMENT RULE
symbolic proxy win	label accuracy rises while interception, deadline, or safety does not	require coupled action and protected physical outcomes
realism scalar	transfer is attributed to a surface label without factorial manipulation	publish the representative-distance vector and causal source--target matrix
practice/learning substitution	end-of-practice score is reported as retention or transfer	freeze delayed tests and first target trial before updates
useless variability	action entropy rises without outcome information or later utility	retire the schedule or use the simpler fixed/null curriculum
state-blind pacing	policy depends on elapsed work and fails unseen resource pathways	compare with calibrated state estimation and constrained control
single-score readiness	one sensor or aggregate certifies broad capability	use action-specific calibrated envelopes with abstention
irreversible promotion	a damaged system cannot regress after deterioration	require monitored stage rollback and recurrence accounting
synchrony attribution	common-input correlation is called coordination	perturb one agent; require reciprocal compensation and lower task error
brittle team convention	co-trained teams work but turnover and cross-play fail	retain typed protocol, repair, and never-co-trained evaluation
deception overconfidence	confidence rises as calibration, regret, or safety worsens	add genuine/deceptive causal controls and calibrated abstention
selection self-fulfillment	selected systems receive more opportunity and later validate the selector	follow rejected units and estimate opportunity-mediated effects prospectively
survivor-only evaluation	dropout, failure, exclusion, or missing follow-up disappears	retain every assigned unit, censoring event, and stopping decision
budget leakage	human work, failed trials, recovery, facility, or embodied energy is omitted	withhold efficiency claim until the complete ledger closes
vocabulary-only residual	ordinary control, curriculum, or inference matches the result	keep the conventional mechanism and retire the added label

Measurable predictions

F-006 contains the full protocols and hard retirement rules. The chapter-level commitments are:

ID	INTERVENTION AND COMPARATOR	MEASUREMENTS	PREDICTION AND FAILURE BOUNDARY
RAP-01	cue-window/channel interventions versus calibrated sequence prediction, retrieval, Bayesian cue model, and POMDP	bits/event, calibration, commitment ms, endpoint m, unsafe events, J/event	residual must transfer to held-out deceptive opponents and full interception; a symbolic-only gain fails

ID	INTERVENTION AND COMPARATOR	MEASUREMENTS	PREDICTION AND FAILURE BOUNDARY
RAP-02	factorial information/action/deadline/consequence/resource changes versus domain randomization, system ID, robust optimization, curriculum, and equal-sample fine-tuning	complete source–target matrix, regret, calibration, adaptation samples, safety, lifecycle cost	a named factor must predict held-out transfer beyond distribution distance; “more realistic” alone fails
RAP-03	history-qualified adaptive variability versus fixed schedules, augmentation, uncertainty-directed curriculum, and random search	practice curve, delayed retention, first-trial near/far transfer, information, failures, J/run	improve transfer at equal criterion exposure and perturbation dose; entropy without utility fails
RAP-04	resource-aware pacing versus state-space readiness, system ID, MPC, robust/adaptive control, risk-sensitive RL, and fixed reserve	task value, output trajectory, state error, admissibility, recovery, recurrence, unsafe events, J	improve the frontier under unseen resource pathways; equality retains the ordinary estimator/controller
RAP-05	reversible staged gate versus time-only, single-score, survival, canary, runtime-assurance, and hand-authored staged rollout	false promotion/withholding, dwell h, recurrence, rollback, availability, review person-hours, J	improve risk--availability on new fault classes and safely regress after deterioration
RAP-06	perturbation and turnover test versus centralized/decentralized control, shared display, protocol, MARL, and common-input controller	task error, lagged compensation, belief bits/event, messages, repair s, cross-play, J	require useful reciprocal compensation and never-co-trained cross-play; synchrony alone fails
RAP-07	deceptive-policy changes versus Bayesian opponent model, fictitious play, imitation, self-play, retrieval, robust and conformal prediction	opponent bits/action, calibration, exploitability, regret, abstention, adaptation s	improve calibrated adaptation at equal interactions, memory, search, and latency
RAP-08	prospective selection policy versus adjusted regression, causal selection, survival, threshold lottery, and broad development	new-cohort calibration, capability, false-negative recovery, opportunity h, attrition, harm, person-hours, J	improve sustainable capability without manufacturing validity through unequal opportunity
RAP-09	complete F-006 composition versus the strongest compatible conventional stack	all protected outcomes, events, steps, bytes, seconds, person-hours, lifecycle J, damage, withheld opportunity	require non-inferiority on protected outcomes and a preregistered Pareto resource gain across tasks, histories, opponents, states, sites, models, and hardware

Mechanism ablations are selective: removing actual-channel state should damage cue interventions; removing resource state should damage pacing and recovery; removing reversible stages should damage recurrence or availability; removing team-belief state should damage turnover and cross-play; removing selection lineage should damage prospective calibration. If every ablation merely reduces capacity and degrades every outcome, the proposed modules have not isolated their claimed roles.

Active acoustic inference

Scope

Sound is not a ready-made sequence of objects. A receiver obtains pressure over time through a medium, room, body, aperture, transducer, clock, calibration state, and preprocessing chain. Sources overlap; direct paths mix with delayed copies; motion changes the operator; an emitted probe changes both the evidence and the physical scene. The system must therefore infer sources and task state from a versioned, action-dependent measurement.

This chapter turns the acoustics, hearing, and auditory-scene-analysis audit into an engineering contract. The detailed equations live in operator-qualified acoustic inference, and Fixture F-009 tests the contract across nine hostile tracks. The editable systems diagram is kept in operator-qualified-active-acoustic-inference.mmd.

The chapter connects four existing parts of the project:

1. **sensorimotor grounding**, because head, body, receiver, and emitter actions change what can be known;
2. **operator-qualified sensing**, because acoustic inference is conditional on a physical forward operator;
3. **sparse predictive compute**, because multirate and event-driven representations can reduce traffic only under a complete task-and-energy test; and
4. the **energy model**, because sensing, emission, motion, communication, maintenance, and embodied hardware remain inside the service boundary.

The intended output is a calibrated decision, retained acoustic artifact, or safe action—not one universal “hearing” score. Detection, discrimination, intelligibility, grouping, separation, dereverberation, localization, ranging, uncertainty, and lifecycle efficiency remain distinct outcomes.

Biological observation

The ear is a level-dependent, multirate measurement system

Acoustic level has meaning only with a reference pressure, frequency and time weighting, integration interval, sensor position, calibration, and uncertainty (C-1054). Room quantities likewise depend on a declared procedure, frequency band, source-receiver geometry, and spatial and temporal support (C-1055). These constraints are part of the biological stimulus, not bookkeeping added after a model runs.

The cochlea implements frequency-dependent mechanical filtering whose response near characteristic frequency is compressively nonlinear. Outer-hair-cell motility contributes causally to that amplification, while psychophysical filter estimates depend on centre frequency and fitting assumptions (C-1056–C-1058). Temporal coding also has regimes: envelope sensitivity is band-limited, auditory-nerve phase locking degrades with frequency, and downstream circuitry can sharpen synchrony in scoped low-frequency conditions (C-1059–C-1061).

“Cochlear,” “spiking,” and “neuromorphic” are therefore insufficient mechanism descriptions. A transferable front end must state its filter shapes, bandwidths, sample rates, level dependence, compression history, saturation, thresholds, refractory periods, timestamp precision, and calibration state. FFT, wavelet, mel, ERB, gammatone, gammachirp, modulation, and learned filters remain competing implementations; cochlear-like filtering is already conventional signal processing (C-1095).

Spatial hearing combines bounded cues with movement

Coincidence-sensitive circuits can encode interaural time differences, but the useful cue changes with frequency content, head and pinna transfer functions, room response, and source spectrum. Interaural level differences and spectral cues contribute on different bands and geometries (C-1062–C-1063). Head motion can disambiguate otherwise similar observations (C-1064); it is an information-acquisition action, not nuisance augmentation.

Early-arriving spatial information can dominate later copies, while spatial separation can improve recognition under masking (C-1065–C-1066). That does not make reverberation irrelevant. Stable spectrotemporal context and prior room exposure can alter later judgments; monaural context can compensate for some reverberant speech distortion; direct-to-reverberant energy contributes to distance judgments (C-1070–C-1073). The useful state therefore includes direct, early, and late components plus their time-varying relation to source and receiver motion.

Auditory objects are inferred under masking

Energetic masking concerns overlap at the receiver. Informational masking also depends on whether target and masker are perceptually segregated, and selective listening changes strongly with ear and source configuration (C-1067–C-1068). Temporal coherence across frequency channels is a plausible grouping mechanism, not a settled universal decomposition rule (C-1069).

This makes four outcomes non-interchangeable:

- **detection:** whether task evidence is present;
- **grouping:** which events belong together and how uncertain that assignment is;
- **separation:** recovery of one or more waveforms under an explicit reference and permutation contract; and
- **causal identity:** which physical source produced a recovered component.

A separator can improve a waveform metric while switching identities, damaging localization, or changing perceptually important structure. Published separation metrics can hide severe waveform changes (C-1094), so no one metric is allowed to stand in for the scene.

Sparse activity and efferent control have narrow causal interpretations

Awake auditory cortex can represent sounds with low population activity, and sparse-coding objectives can learn cochlea-like kernels in some sound ensembles. The learned code depends on the training ensemble (C-1074–C-1076). Sparse activity is evidence about representation, not a direct measurement of system joules. Sensors, clocks, memory traffic, event routing, decoding, idle power, and training may dominate (C-1097).

Efferent activation can causally reduce cochlear gain and improve masked-tone auditory-nerve responses in the studied preparations (C-1077). A general human speech-in-noise benefit from medial olivocochlear strength remains disputed (C-1078). An AI analogue earns the name only when an intervention on task-conditioned front-end gain changes a registered external endpoint after distortion, recovery time, calibration, and energy are charged. An attention map or smaller hidden activation is not that evidence.

Echolocation is active observation with physical consequences

Echolocators vary call timing with spatial task difficulty, adapt emission level with distance, steer or reshape the acoustic field of view, and coordinate head or pinna motion with emission (C-1079–C-1082). Human experts can use self-generated echoes, but the demonstrated capability includes extensive experience and neural reorganization (C-1083). Spectral jamming avoidance in bats is context-dependent rather than universal (C-1084).

The transferable loop is emit, receive, infer, move or re-aim, and decide. Its benefit is conditional on waveform ambiguity, target reflectivity, range, clutter, association, other emitters, detectability, acoustic exposure, and the energy required to produce and receive the probe (C-1085, C-1096).

Proposed AI translation

Preserve the episode, operator, and action identity

For episode e , retain the contract

$$\mathcal{A}_e = (X_e, S_e, E_e, R_e, H_e, O_e, C_e, T_e, U_e, B_e),$$

where:

SYMBOL	MEANING AND REQUIRED UNITS
X_e	geometry and boundaries [m], medium, temperature [K], relative humidity [1], flow [m/s], occupancy, and time-varying material state
S_e	target and interfering sources, waveforms [Pa], positions [m], velocities [m/s], directivity, identities, roles, and emission times [s]
E_e	commanded and realized active emissions: waveform, spectrum [Hz], duration [s], aim [degree or rad], acoustic energy [J], and exposure
R_e	receiver, body, and array pose; aperture [m]; morphology or manifold; transducer response; gain; health; saturation; and feasible motion
H_e	prior exposure, room and source history, adaptation, training, feedback, previous actions, and their timestamps [s]
O_e	versioned observation operator: impulse responses, sample rate [sample/s], precision [bit], clock, latency [s], preprocessing, support, and missingness
C_e	calibration identity, reference pressure and distance, covariance, traceability, method, interval, and data vintage
T_e	literal target construct, decision deadline [s], and registered loss or utility in its native unit
U_e	independent waveform, event, source, room, receiver, body, array, device, site, or population unit
B_e	componentwise limits on evidence, time, memory, communication, action, exposure, risk, human work, and lifecycle energy

Calibration identity can become part of the learned task when a model depends on device, room, or preprocessing state (C-1098). Train-test splits must group clean sources, convolved variants, neighboring windows, impulse-response and renderer lineages, HRTF or array families, devices, speakers or listeners, and derived labels. Capture time and receipt time remain distinct, and only causally received observations may affect a decision.

Keep pressure, level, exposure, and energy separate

For pressure $p(t)$ in pascals over interval T in seconds,

$$p_{\text{rms}} = \sqrt{\frac{1}{T} \int_0^T p^2(t) dt}, \quad L_p = 20 \log_{10} \left(\frac{p_{\text{rms}}}{p_0} \right),$$

where p_{rms} is RMS pressure [Pa], reference pressure $p_0 = 20 \mu\text{Pa}$ in air, and L_p is sound-pressure level [dB re $20 \mu\text{Pa}$]. Frequency weighting, time weighting, band, position, orientation, and calibration must be explicit. Sound exposure $\int p^2(t) dt$ has units $\text{Pa}^2 \text{s}$; it is not acoustic emission energy [J], electrical energy [J], or a complete biological risk model.

For source pressure $x_s(t)$, received pressure $y_m(t)$, and noise $n_m(t)$, all in pascals, use the time-varying mixture

$$y_m(t) = \sum_{s=1}^{N_s} \int h_{m,s}(t, \tau) x_s(t - \tau) d\tau + n_m(t),$$

where m is receiver-channel index [1], s is source index [1], source count N_s is dimensionless, t and delay τ are in seconds, and propagation kernel $h_{m,s}(t, \tau)$ is in reciprocal seconds. A fixed convolution is only a special case; moving sources, receivers, boundaries, media, and emitters make the operator time-varying.

Build a qualified multirate front end

The front end may combine:

1. an FFT, wavelet, constant-Q, mel, ERB, gammatone, gammachirp, modulation, or learned analysis bank;
2. level- and history-dependent compression with declared attack, release, saturation, and distortion;
3. masking and source-count hypotheses rather than a single global noise scalar;
4. dense samples or sparse onset, offset, change, phase, or threshold events with timestamp and refractory identity; and
5. a decoder or downstream task that is charged for reconstructing information discarded by the encoder.

Band selection and update rate should follow the causal time scales needed by the task. Fine timing cannot be assumed above the supported band; weak sustained signals and rare hazards cannot be discarded merely because an event count falls.

Fuse spatial evidence without erasing ambiguity

For path-length difference Δr [m] and sound speed c_{snd} [m/s], interchannel delay is

$$\Delta t = \frac{\Delta r}{c_{\text{snd}}},$$

where Δt is in seconds. For left and right RMS pressures p_L, p_R [Pa],

$$\text{ILD} = 20 \log_{10} \left(\frac{p_L}{p_R} \right)$$

is interaural level difference [dB]. Both cues are band-, source-, room-, body-, and calibration-qualified. Spectral cues and head or receiver motion remain separate intervention axes. A correlation surface under reverberation may be multimodal; reducing it to one delay without calibrated uncertainty loses the ambiguity the controller needs.

The learned path must be compared with generalized cross-correlation, matched filtering, steered-response power, delay-and-sum, robust and minimum-variance beamforming, MUSIC or ESPRIT, Bayesian delay estimation, and source tracking. GCC, MVDR, and subspace localization are mature nulls, not discoveries created by a biological label (C-1086–C-1088).

Separate scene inference from metric substitution

Each output head declares one literal construct and uncertainty model. The minimum firewall is:

CONSTRUCT	REQUIRED MEASUREMENT	INVALID SUBSTITUTE
detection	hit and false-alarm rates, criterion, calibration, and dimensionless d'	recognition accuracy or mixture SNR
intelligibility or identification	confusion matrix and words, phonemes, or information units under a protocol	SI-SDR, loudness, or localization
grouping	event-to-source assignment, count, switch rate, ambiguity, and identity	separated waveform or attention weight
separation	SI-SDR plus waveform, perceptual, downstream-task, and identity measures	localization or intelligibility alone
dereverberation	direct, early, and late waveform/operator measures plus downstream effect	a "dry" label or one decay scalar

CONSTRUCT	REQUIRED MEASUREMENT	INVALID SUBSTITUTE
localization	azimuth and elevation error [degree], front-back error, coverage, and calibration	beam width, identity, or range
ranging	bias, absolute error, and interval coverage [m]	echo detection or direction
sparse timing	timestamp error [s], misses, false events, information, bytes, task value, and joules	event count or active fraction
active sensing	randomized emission or motion contrast, decision value, latency [s], detectability, exposure, risk, and energy [J]	action-success correlation
uncertainty	proper score, interval coverage, risk-coverage, and abstention value by held-out operator	confidence magnitude

Independent-component methods, time-frequency sparsity methods, deep clustering, permutation-invariant training, and time-domain separation are mature but assumption-bound baselines (C-1089–C-1093). The strongest compatible composition—not an isolated weak separator—is the decisive null.

Close the action loop

At decision time t , policy π_q for method q selects

$$(a_t^{\text{emit}}, a_t^{\text{body}}, a_t^{\text{sensor}}, a_t^{\text{gain}}, a_t^{\text{task}}) = \pi_q(\mathcal{H}_t, \widehat{X}_t, \widehat{O}_t, \widehat{U}_t, \mathcal{A}_t^{\text{safe}}, B_t),$$

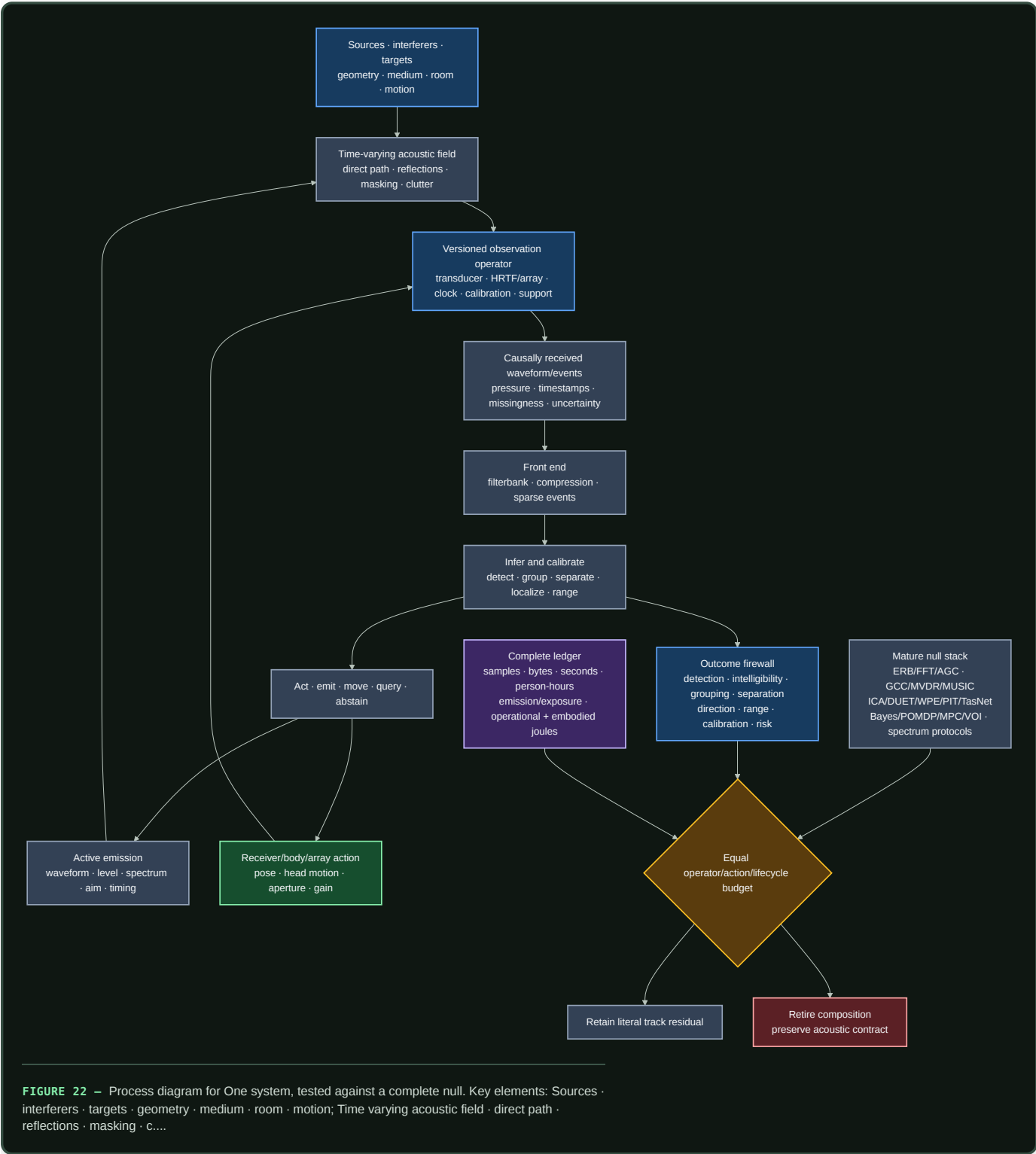
where $\mathcal{H}_t$ is causally received history, $\widehat{X}_t$ is estimated scene state, $\widehat{O}_t$ is operator and calibration state, $\widehat{U}_t$ is uncertainty, $\mathcal{A}_t^{\text{safe}}$ is the feasible action set, and B_t is remaining componentwise budget. Emission actions set waveform, spectrum [Hz], level [referenced dB], aim [degree or rad], and time [s]; body actions set pose and motion [m, rad, s]; sensor actions select aperture [m], channels, and sample rate [sample/s]; gain is dimensionless or declared in dB; the task action uses its native unit.

For monostatic echo emission and receipt times t_{emit} and t_{recv} [s], nominal range is

$$\hat{r} = \frac{c_{\text{snd}}}{2}(t_{\text{recv}} - t_{\text{emit}}),$$

where $\hat{r}$ is range [m]. The interval must propagate clock error, target motion, refraction, multipath, ringing, waveform ambiguity, association, and detector threshold. Multiple emitters add collision, authentication, fairness, adversarial imitation, and jamming state. The proposal must beat TDMA, FDMA, CDMA, random access, listen-before-talk, power control, coded probes, cancellation, authentication, and game-theoretic allocation at equal spectrum, power, messages, and exposure.

One system, tested against a complete null



Editable source: operator-qualified-active-acoustic-inference.mmd.

The contract is owned by existing components:

- Candidate 002 — Multiscale context broadcast for multirate context;
- Candidate 006 — Reversible physical skill for reusable analog or beamforming paths;
- Candidate 007 — Endogenous observation surveillance for action-dependent observation;
- Candidate 009 — Graded assurance envelopes for calibration, uncertainty, exposure, and fallback;

- Candidate 012 — Latency-qualified authority for travel time, evidence age, and decision authority; and
- Candidate 014 — Versioned observation contract for operator and calibration identity.

Efficiency mechanism

The proposed composition can improve efficiency through four literal pathways:

1. **Multirate analysis.** Slow envelopes and long context need not update at the rate required for fine timing. Savings must appear in samples, memory traffic, latency, and measured joules without degrading a protected band.
2. **Conditional events.** Onsets, offsets, changes, or uncertainty triggers can suppress redundant transport and downstream work. The decoder, timestamping, missed sustained signals, false events, and idle event hardware remain in the comparison.
3. **Selective physical action.** A head turn, receiver move, gain change, or acoustic probe can resolve an ambiguity more cheaply than processing another passive interval. Its motion, emission, exposure, delay, detectability, and interference costs must be lower than the decision value it creates.
4. **Reusable calibrated structure.** Stable filter, beamforming, correlation, or control paths can move to DSP, FPGA, ASIC, analog, or neuromorphic hardware when reuse amortizes conversion, characterization, drift monitoring, repair, and replacement.

For one registered service interval, lifecycle energy is

$$E_q^{\text{life}} = E_q^{\text{data}} + E_q^{\text{train}} + E_q^{\text{emit}} + E_q^{\text{sense}} + E_q^{\text{infer}} + E_q^{\text{move}} + E_q^{\text{comm}} + E_q^{\text{store}} + E_q^{\text{cal}} + E_q^{\text{maint}} + E_q^{\text{emb}},$$

where method q has data-acquisition, training, acoustic-emission, sensing, inference, motion, communication, storage, calibration, maintenance, and amortized embodied-energy terms, each measured in joules over the same service interval. The complete ledger also reports:

- pressure samples, sparse events, optimization steps, queries, and bytes;
- wall time [s], task latency [s], peak memory [byte], and network traffic [byte];
- acoustic exposure [$\text{Pa}^2 \text{s}$], unsafe-event count [1], and harm or detectability in a registered task-specific unit;
- emitted acoustic energy [J], electrical input [J], sensing [J], motion [J], compute [J], cooling [J], and facility allocation [J];
- recording, labeling, listening tests, calibration, tuning, monitoring, maintenance, and repair [person-hour]; and
- transducers, arrays, processors, actuators, replacements, wear, embodied hardware, and opportunity cost.

An arm that exceeds any binding budget is infeasible. Resources released by an ablation remain unused. Lower event count, lower FLOPs, quieter emission, or a faster accelerator is not an efficiency result unless accepted task service improves on the complete protected vector.

Evidence status

The stable claim ledger mirrors all 46 audit-local claims without promoting the audit itself into evidence:

TOPIC	STABLE CLAIMS	STATUS	CONSEQUENCE
calibrated acoustic and room measurement	C-1054–C-1055	established	reference, support, geometry, operator, and uncertainty are mandatory

TOPIC	STABLE CLAIMS	STATUS	CONSEQUENCE
cochlear mechanics, filters, timing, binaural cues, motion, precedence, masking, and selective listening	C-1056–C-1068	established	supplies scoped mechanisms and limits, not one front-end prescription
temporal coherence as a grouping mechanism	C-1069	plausible	test against spatial, statistical, and learned grouping nulls
context, room adaptation, distance, and low population activity	C-1070–C-1074	established	motivates explicit context and sparsity tests while keeping outcomes separate
sparse coding learns cochlea-like kernels	C-1075	plausible	representation result; no direct hardware-energy claim
ensemble dependence and causal efferent gain	C-1076–C-1077	established	qualify learned codes by data and test gain through intervention
general human speech-in-noise benefit from efferent strength	C-1078	disputed	do not use as a broad task-benefit premise
adaptive biosonar, field steering, coordinated motion, and expert human echo use	C-1079–C-1083	established	motivates costed emission-reception-action loops
universal spectral jamming avoidance	C-1084	disputed	use multi-emitter and adversarial controls rather than a universal law
waveform ambiguity and mature GCC, MVDR, MUSIC, ICA, sparse, deep, PIT, and time-domain separation nulls	C-1085–C-1093	established	compare with a complete competitive DSP and learned stack
metric failure, conventional cochlear-like DSP, and physical cost of active sensing	C-1094–C-1096	established	enforce the metric firewall and complete action ledger
sparse activity is not system energy; calibration identity can be learned task state	C-1097–C-1098	plausible	measure system energy and hold out operator versions
transferable residual is the operator- and action-qualified contract	C-1099	speculative	preserve the benchmark; retire the composition if the mature null matches it

No row establishes that the complete proposed system beats the conventional stack. The evidence supports the physical constraints, biological phenomena, and engineering nulls. The composition remains an experiment.

Speculative extensions

Operator-conditioned auditory state

Learn a compact state that jointly represents room response, body or array transfer, clock uncertainty, source hypotheses, and front-end state. It must predict held-out calibration probes and remain identifiable under scene changes. Compare it with explicit system identification, adaptive filtering, room banks, and Bayesian state-space models.

Counterfactual acoustic action

Train the policy to predict not only what a probe or head turn might reveal but also its exposure, detectability, collision, motion, and lifecycle cost. Use paired simulator seeds and physical replay so action value is estimated against the same latent scene, not easier episodes selected after acting.

Cross-timescale front-end control

Let fast protection and compression, intermediate event routing, and slower room or task adaptation share state without sharing unrestricted authority. Test whether the joint controller beats independently tuned AGC, PCEN, adaptive filters, selective prediction, and recurrent context after coordination cost.

Population-qualified acoustic hardware

Route recurring transforms among digital, analog, array, and event-driven front ends by measured device identity, calibration envelope, workload reuse, and remaining service life. Hold out hardware classes, lots, transducers, and array geometries. A nominal-device gain does not establish fleet-level benefit.

Strategic multi-emitter sensing

Model emitters as cooperative, selfish, deceptive, imitating, colluding, or turnover-prone agents. Joint waveform and scheduling policies must retain source authentication and calibrated uncertainty under never-coordinated population cross-play. Ordinary spectrum protocols remain the null.

Failure modes

SIGNATURE	INTERPRETATION AND REQUIRED RESPONSE
an unreferenced decibel value or unspecified integration interval drives the result	the stimulus and risk boundary are undefined; invalidate the comparison
train and test share a clean source, impulse response, renderer, HRTF, array family, device, or neighboring window	the result is an inverse crime or lineage leak; rebuild grouped splits
a fixed room convolution succeeds but moving-source, moving-receiver, or changing-boundary trials fail	the observation operator was misspecified; restrict the validity envelope
a cochlear label wins only against a weak mel front end	the mechanism was not isolated; restore FFT, wavelet, ERB, gammatone, learned, compression, and DSP nulls
sparse events lower active fraction but not bytes or joules	representation sparsity did not create system efficiency; retire the energy claim
event gating misses weak sustained signals, fine timing, or rare hazards	the threshold policy violates the protected outcome; retain a dense or hybrid path
localization collapses a multimodal GCC or beam surface to one confident point	estimator uncertainty is miscalibrated; preserve alternatives or abstain
apparent spatial super-resolution depends on source-count priors or training overlap outside aperture support	prior information is being reported as measured resolution; narrow the claim
SI-SDR rises while intelligibility, localization, identity, calibration, or downstream value worsens	metric optimization substituted for the task; fail the separation track
an efferent analogue changes gain or hidden activation without an external endpoint	no causal task benefit was demonstrated; remove the biological interpretation
active sensing wins only after extra samples, emission, exposure, motion, latency, or detectability	additional opportunity explains the result; match the full action ledger
echo range ignores clock drift, target motion, multipath, ringing, waveform ambiguity, or association	the interval is invalid; propagate those uncertainties or abstain
frequency shifting or timing changes fail against new emitters, collusion, authentication faults, or equal-power cross-play	the jamming policy is context-bound; retain conventional coordination
a learned path beats one beamformer or separator but not their strongest compatible composition	the baseline is incomplete; B7 in F-009 remains decisive
confidence fails in held-out rooms, bodies, arrays, clocks, calibration versions, sites, or hardware	the system learned operator identity without transferable calibration; restrict or retrain
an efficiency result omits recording, labels, calibration, failed trials, cooling, maintenance, replacement, or embodied hardware	the accounting boundary is incomplete; recompute the lifecycle result

Measurable predictions

The nine F-009 tracks convert the chapter into falsifiable predictions:

- 1. Level-dependent front end.** A qualified nonlinear multirate front end improves registered detection or discrimination across held-out levels, bands, histories, sensors, and domains beyond the strongest fixed and learned filterbank-plus-compression stack at equal bytes, latency, and joules.
- 2. Sparse timing.** Event encoding preserves protected timing, weak sustained sources, and rare hazards while reducing measured memory traffic and service energy beyond VAD, codecs, sparse convolution, and duty-cycled DSP.
- 3. Active emission and reception.** Joint waveform, emission, and receiver action creates positive causal decision value after exposure, detectability, motion, interference, calibration, and lifecycle energy are charged.
- 4. Localization under operator shift.** Calibrated direction and range remain better than GCC, Bayesian delay, SRP, MVDR, MUSIC, and tracking baselines on held-out HRTFs, arrays, bodies, apertures, rooms, spectra, faults, and clocks.
- 5. Reverberation adaptation.** Context state improves literal waveform, intelligibility, event identity, distance, or calibration outcomes in new rooms beyond WPE, Wiener filtering, explicit operator estimation, and equal-sample adaptation without confusing compensation with recovery.
- 6. Grouping and separation.** The model improves event-to-source grouping, source count, separation, causal identity, localization, uncertainty, and downstream value across new sources, rooms, counts, and cue conflicts beyond the complete ICA, IVA, DUET, NMF, WPE,

PIT, deep-clustering, TasNet, and neural-beamforming stack.

7. **Efferent-like gain.** Randomized intervention on task-conditioned gain improves a protected external endpoint after distortion, false suppression, recovery, calibration, and energy are included; internal suppression alone predicts no pass.
8. **Multi-emitter interference.** Adaptive emission coordination improves ranging or detection, association, throughput, fairness, calibration, and exploitability under unseen cooperative and adversarial populations beyond mature spectrum-access and authentication protocols at equal power and messages.
9. **Complete lifecycle.** Any surviving benefit replicates across at least two source families, rooms or media, arrays or bodies, operator versions, scene types, model families, sites, and hardware classes after every acquisition, action, person-hour, risk, operational-energy, and embodied-energy term is charged.

Preregister protected outcomes, non-inferiority margins, material-improvement thresholds, 95% intervals, multiplicity control, tail-failure ceilings, and stopping rules. Randomize and analyze at the unit receiving the intervention; adjacent samples and overlapping windows are not independent sources.

The hard decision is simple: if the complete conventional composition matches the proposal at equal operator, action, calibration, uncertainty, exposure, human-work, and lifecycle budgets—or if no ablation isolates value beyond that stack—retire the architectural residual. Keep the measurement, operator, action, metric, and accounting contract.

Operator-qualified sensing and physical inference

Scope

This chapter defines what an adaptive system is allowed to claim from a physical measurement. A sensor does not deliver a scene, object, or fact. It delivers a finite observation produced by an aperture, illumination pattern, medium, detector, clock, calibration state, acquisition policy, and noise process. Inference adds assumptions and prior information.

The architectural consequence is precise: decoded tensors may be convenient runtime inputs, but they are not self-describing evidence. The system must keep enough of the measurement operator, uncertainty, validity envelope, and lineage to know what the observation could resolve, what it could not resolve, and when another measurement is worth buying.

This chapter is operationalized by:

1. the optics, photonics, and inverse-sensing audit;
2. the operator-qualified optical-inference mathematics; and
3. Fixture F-007, which tests the full contract against inverse-method, active-sensing, calibration, control, digital-accelerator, and passive-optics baselines.

Biological observation

Biological sensing is already physical inference. An eye has finite aperture, spectral sensitivity, sampling density, integration time, dynamic range, blind regions, motion, adaptation, and a body that can change viewpoint. A useful percept can therefore depend on both received evidence and prior structure. Movement can reveal a surface that one view leaves ambiguous; longer exposure can buy photons while losing temporal resolution; adaptation can extend useful operation while changing the response function.

The transferable observation is not a particular visual anatomy. It is the closed coupling among:

- a bounded physical measurement channel;
- an internal estimate that remains conditional on that channel;
- actions that change future observability;
- calibration and adaptation over multiple timescales; and
- task-specific decisions made before every latent detail is known.

Optics makes those constraints measurable. Finite apertures, null spaces, photon statistics, ambiguity classes, saturation, and drift are established properties of physical sensing (C-970–C-988). They prevent a fluent reconstruction from silently becoming stronger evidence than the acquisition supplied.

Proposed AI translation

Preserve the evidence-producing operator

For latent physical state x_e in episode e , let acquisition t produce

$$y_{e,t} = g_{e,t}(\mathcal{H}_{\nu_{e,t}}(a_{e,t}, c_{e,t})x_e) + n_{e,t},$$

where:

- $y_{e,t}$ is the raw observation in detector counts [count] or another declared sensor unit;
- $a_{e,t}$ is the acquisition action, such as viewpoint, exposure, wavelength, or illumination pattern;
- $c_{e,t}$ is the calibrated parameter vector in declared native units;
- $\nu_{e,t}$ is the immutable operator-version identifier;
- $\mathcal{H}_{\nu_{e,t}}$ is the physical forward operator;
- $g_{e,t}$ is detector conversion, clipping, and readout response; and
- $n_{e,t}$ is only the residual noise represented by an explicit likelihood.

The observation record travels with operator version, calibration covariance, saturation/dead-time mask, capture and receipt times, preprocessing lineage, and validity envelope. A compact representation can replace the raw record only for a registered family of future queries and only while reconstruction, uncertainty, and provenance obligations remain satisfied. This connects the versioned observation contract, semantic compaction, and value-aware retention. A friction coefficient is one interface-specific instance: its parties, motion, load, environment, history, operator, support, and evidence age travel with the estimate (C-1497).

Separate measured information from prior-supported reconstruction

For a linearized operator $H = U\Sigma V^*$, a component along v_j with singular value $\sigma_j = 0$ lies in the measurement null space. A decoder may still propose a plausible value for that component, but the value comes from a prior, another measurement, or a convention—not from this observation.

Every output therefore carries three distinguishable uncertainty sources:

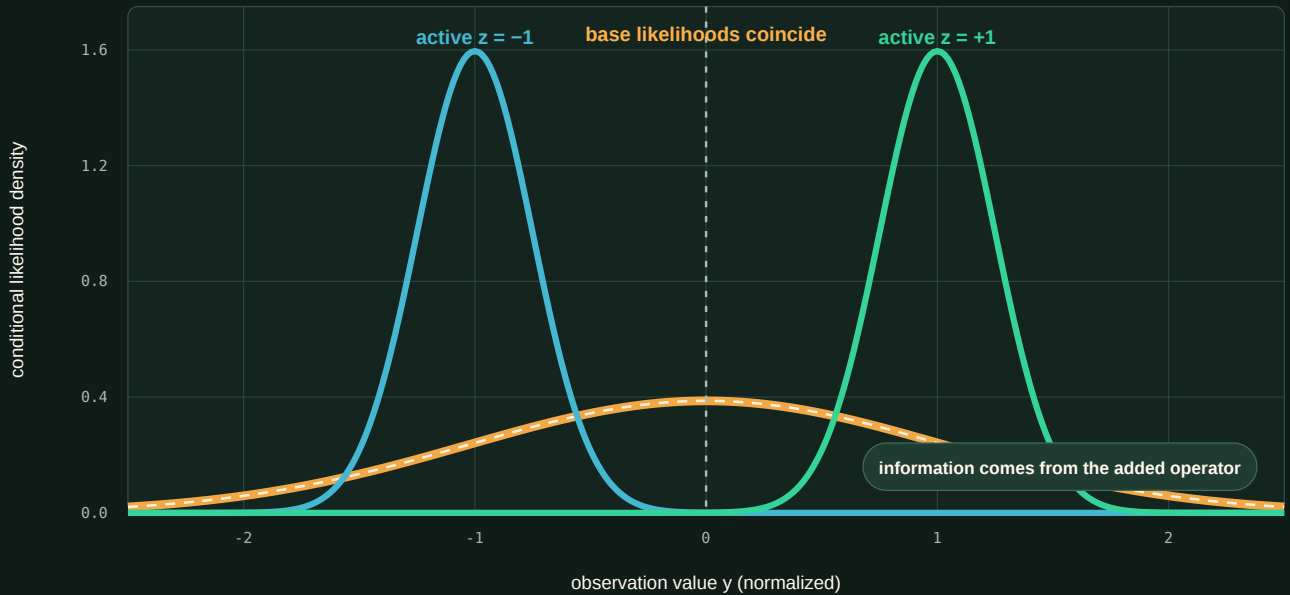
SOURCE	WHAT VARIES	APPROPRIATE RESPONSE
measurement noise	photon arrivals, read noise, background, quantization	propagate the likelihood; change exposure or sensor when valuable
operator uncertainty	calibration, alignment, drift, response, timing	monitor residuals; recalibrate, downgrade, reroute, or abstain
prior or model uncertainty	training support, regularizer, latent family, task shift	expose support dependence; acquire discriminating evidence or retain alternatives

The separation matters under compressed sensing, phase retrieval, computational super-resolution, blind calibration, and learned reconstruction (C-972, C-976–C-981, C-985). Pixel count, sharpness, or confidence cannot substitute for newly identified physical information.

The existing Fixture F-007 likelihood plot makes that distinction explicit. It is an analytical identifiability example, not an empirical superiority result.

Identical likelihood versus added evidence

$$p(y_b | z=+1) = p(y_b | z=-1); y_a = z + N(0, \sigma_a^2)$$



Fixture configuration · no empirical or superiority result

FIGURE 23 — Under the base operator, the two latent states have identical likelihoods; the active measurement separates them, and a decoder cannot turn null-space ambiguity into new measurement evidence without changing the operator or supplying a prior.

Buy another measurement only when it changes the decision frontier

Let b_t be the current belief, a a safe acquisition action, d a downstream decision, and $U(d, \theta)$ task utility for uncertain state θ . The expected value of information is

$$\text{EVI}(a | b_t) = \mathbb{E}_y \left[\max_d \mathbb{E}[U(d, \theta) | b_t, a, y] \right] - \max_d \mathbb{E}[U(d, \theta) | b_t].$$

No scalar acquisition price is assumed. The controller compares EVI against a cost vector containing at least photons [count], energy [J], latency [s], dose or disturbance in its task-specific unit, actuator wear [cycles], and risk on a declared scale. An action is admissible only inside its safety and authority envelope.

This converts active perception from “collect more data” into a resource allocation problem. The active path must beat fixed acquisition, greedy value of information, Bayesian experiment design, POMDP planning, and model-predictive control at equal opportunity and cost (C-975).

Monitor validity instead of trusting calibration indefinitely

Calibration is versioned state, not a one-time property of a device. Reference channels and task-independent residuals monitor alignment, gain, timing, temperature, background, saturation, and component aging. A threshold crossing does not identify the cause; it changes what action is permitted.

A validity transition can trigger, in order:

1. a qualified reduction in confidence or supported query set;
2. a new reference or calibration acquisition;
3. rerouting to a different sensor or computational path;
4. a digital or conservative fallback;

- 5. reset, repair, or replacement; and
- 6. abstention when none of those paths restores the evidence contract.

Blind self-calibration is tested for identifiability, and task residuals are not allowed to conflate scene shift with device drift (C-984–C-989).

Route a transform to the substrate that actually makes it cheap

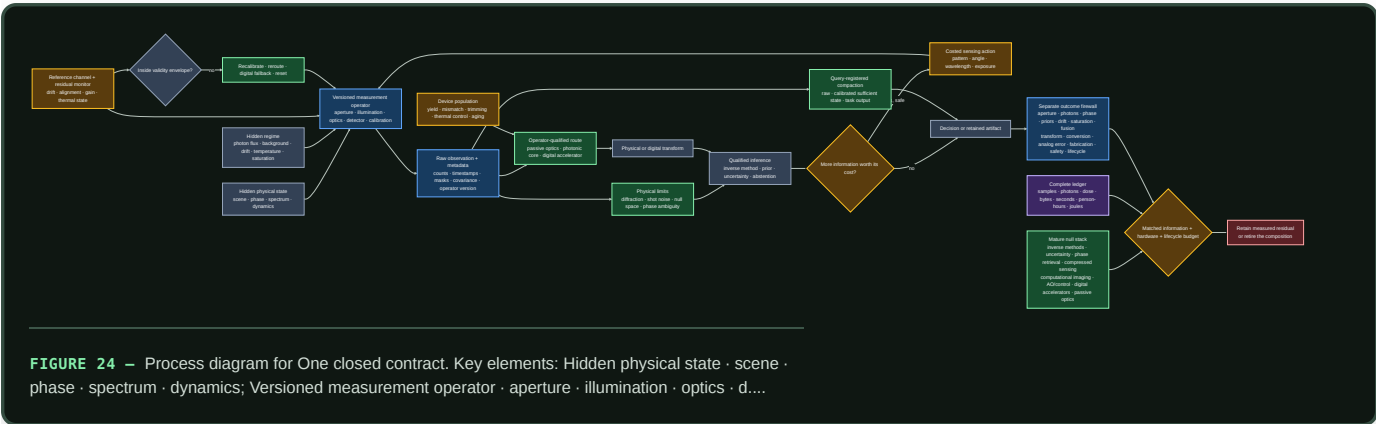
Passive and active optical hardware can execute physically matched linear transforms, sometimes before an observation becomes a large digital tensor. That is valuable when the input is already optical, the transform is reusable, conversion can be avoided, and required precision fits the device envelope.

It is not a general preference for an optical path. The route record declares:

- input locality and format;
- transform identity, reuse count, sparsity, and required precision;
- source, modulator, detector, ADC/DAC, control, thermal, and host work;
- device-specific calibration, mismatch, yield, drift, and age;
- accepted-output latency and quality; and
- fallback and migration cost.

Routing then compares passive optics, a photonic core, a digital accelerator, and hybrid compositions on the same workload and service boundary. Optical propagation is credited only for work it actually displaces (C-989–C-999).

One closed contract



Editable source: operator-qualified-physical-inference.mmd.

Efficiency mechanism

The contract permits four efficiency gains, each with a matching way to fail:

1. **Acquire selectively.** Spend photons, time, and motion only where another observation changes an accepted decision. It fails when the acquisition controller costs more than fixed sensing or shifts risk outside the ledger.
2. **Transform before expansion.** Use a physical operator to filter, aggregate, or project local optical information before high-volume digital movement. It fails when conversion, source, control, or recalibration erases the saving.
3. **Retain the sufficient level.** Store raw evidence, calibrated sufficient state, or task output according to registered future queries and recovery obligations. It fails when later queries expose discarded information.

4. **Route by validity and reuse.** Amortize a stable transform on hardware that suits its precision and repetition. It fails under workload shift, fabrication spread, thermal control, low utilization, or short lifetime.

For each accepted service unit,

$$E_{\text{service}} = E_{\text{source}} + E_{\text{mod}} + E_{\text{prop}} + E_{\text{detect}} + E_{\text{ADC}} + E_{\text{DAC}} + E_{\text{control}} + E_{\text{digital}} + E_{\text{thermal}} + E_{\text{facility}} + E_{\text{embodied}},$$

where every E term is energy [J] measured over the same accepted-output boundary. Embodied energy includes fabrication, packaging, yield loss, replacement, and end-of-life treatment amortized over accepted lifetime service. The comparison also reports quality, calibration, latency, risk, photons, bytes moved, and human maintenance effort; joules alone cannot hide a worse sensing contract.

Evidence status

INGREDIENT	STABLE CLAIMS	STATUS AND ARCHITECTURAL USE
operator, aperture, null space, photon and precision limits	C-970–C-974	established physical constraints; mandatory measurement metadata
active illumination and structural priors	C-975–C-980	established scoped mechanisms; advantage remains task- and prior-qualified
multiplexing, coded acquisition, and adaptive correction	C-981–C-984	established tradeoffs; sensorless objective validity remains plausible
blind calibration, drift, saturation, and fusion	C-985–C-988	established constraints on identifiability and valid combination
physical transforms and avoided conversion	C-989–C-990	physical execution established; end-to-end benefit workload-dependent
system energy, conversion, and analog error	C-991–C-993	system-boundary constraints established; core-only efficiency claims disputed
fabrication, thermal control, and in-situ adaptation	C-994–C-996	variation and thermal cost established; recoverable mismatch is scoped
labels, routing, uncertainty, and lifecycle ranking	C-997–C-1001	neuromorphic-label inference disputed; routing and uncertainty composition plausible; scoped lifecycle reversal established

The sources support the constraints and component mechanisms. They do not yet show that their full composition improves this project's quality–risk–latency– energy frontier. F-007 is therefore a hostile fixture, not an architecture promotion.

Speculative extensions

Learned operator compaction

Learn the smallest operator state that preserves a registered family of likelihoods, counterfactual acquisitions, and calibration decisions. Compare it with explicit metadata, sufficient-statistic storage, low-rank calibration, and recomputation from raw evidence. A compact state that cannot answer a new registered query is rejected.

Joint query, sensor, and substrate routing

Let one controller decide whether to answer from retained state, acquire a new physical observation, or move the transform to another substrate. The claim is interesting only if joint control beats three separately optimized controllers after coordination and monitoring cost.

Population-calibrated physical modules

Treat fabrication variation as measured device identity rather than nominal noise. Assign workloads by the calibrated envelope of each device, then test whether characterization, placement, spares, and migration work less than trimming every device to one specification.

Future-query-aware sensing

Choose acquisitions that serve both the immediate decision and declared future queries. This could favor a slightly more expensive measurement now if it prevents reacquisition or unsafe inference later. The future-query distribution must be registered before results are inspected.

Failure modes

SIGNATURE	INTERPRETATION AND REQUIRED RESPONSE
sharper reconstructions appear without improved held-out physical decisions or calibration	the prior changed appearance, not measured information; narrow the claim
null-space pairs receive confident different answers from the same observation	the decoder hides prior selection; expose alternatives or abstain
active sensing wins only with more photons, time, dose, or actuator work	extra opportunity explains the result; match the acquisition ledger
a multiplex advantage disappears when the dominant noise source changes	the result is regime-specific; retain the crossover, not a universal rule
drift monitoring reacts to scene shift or misses reference-channel failure	the validity detector is not identifiable; add controls or conservative fallback
fused confidence improves while shared calibration error remains unmodelled	covariance was double-counted; use robust fusion or keep sensors separate
an optical path wins on core propagation but loses sensor-to-decision energy	conversion, control, or movement dominates; retain the digital baseline
nominal-device accuracy hides die, package, temperature, and age spread	the hardware claim is not population-valid; stratify devices and lifetime
in-situ adaptation consumes unreported training measurements or human tuning	calibration work is omitted; charge it to deployment
compact storage answers current tasks but prevents a registered later query	compaction violated the preservation contract; retain raw or richer state
one scalar efficiency score hides worse risk, calibration, or maintenance	report the Pareto vector; do not average protected outcomes away

Measurable predictions

- 1. Null-space honesty.** On paired physical states that share an observation under the tested operator, an operator-aware system will retain ambiguity or abstain more accurately than a tensor-only decoder without reducing identifiable-task performance.
- 2. Prior-shift qualification.** Under held-out scene structure, operator-aware uncertainty will predict super-resolution and compressed-recovery failure better than confidence from the reconstruction model alone.
- 3. Costed active acquisition.** At equal photons, dose, latency, action count, risk, and joules, active selection will improve accepted task utility beyond fixed acquisition and one-step EVI—or the learned acquisition mechanism is retired.
- 4. Noise-regime crossover.** Multiplexed and focused acquisition will exchange rank at a reproducible noise boundary predicted before the confirmatory run.
- 5. Drift-aware recovery.** Versioned monitoring will reduce invalid confident outputs and recovery time under hidden alignment, gain, timing, and thermal changes beyond periodic calibration at equal reference and maintenance cost.
- 6. Heterogeneous crossover.** A physical path will improve end-to-end accepted outputs per joule only in preregistered regions of transform reuse, precision, input locality, utilization, and device validity; digital routing will win outside them.
- 7. Device-population validity.** Routing by measured device envelope will improve yield-adjusted lifetime service beyond nominal routing and uniform trimming after characterization, migration, spare, and control costs.
- 8. Query-preserving compaction.** A query-registered retained state will use fewer stored bytes and lifecycle joules than raw retention while meeting every registered reconstruction, uncertainty, provenance, and recalibration tolerance.

All predictions are evaluated through F-007. A positive result remains bounded to its measurement operator, physical regime, query set, hardware population, workload, and lifecycle boundary.

Active chemical sensing

Scope

A chemical sensor does not receive an odor, analyte identity, source, or hazard. It receives a time-dependent response produced jointly by source release, transport, reaction, surfaces, the receiver's path, the sampling action, the inlet and chamber, sensor chemistry, temperature, humidity, calibration, adaptation, ageing, contamination, and previous exposure. In a turbulent plume, even the material reaching the receiver arrives as intermittent whiffs and blanks rather than a smooth pointer to its source.

This chapter turns the olfaction, chemical sensing, and plume-tracking audit into readable architecture. The detailed definitions live in operator-qualified chemical-sensing mathematics, and Fixture F-011 tests the architecture across fourteen hostile tracks. The editable diagram is kept in operator-qualified-active-chemical-sensing.mmd.

The chapter connects existing project components rather than adding another principle or candidate:

1. **sensorimotor grounding**, because sniffing, pumping, orientation, locomotion, and receiver geometry change the evidence;
2. **operator-qualified sensing**, because every chemical result remains conditional on a physical forward operator;
3. **sparse predictive compute**, because temporal events and sparse representations earn efficiency credit only through total task work and measured energy;
4. **memory and consolidation**, because fast adaptation, learned associations, slow calibration, drift, and maintenance occupy different state and update timescales;
5. **reliability under mission profiles**, because humidity, contamination, ageing, poisoning, replacement, and out-of-support operation change the device rather than merely the data; and
6. **the energy model**, because motion, pumps, heaters, preconcentration, chromatography, vacuum/ionization, calibration gases, consumables, maintenance, human work, and embodied devices remain inside the service boundary.

The intended output is a calibrated decision, qualified retained observation, safe action, or abstention. Presence, molecular identity, perceptual odor identity, concentration, mixture composition, direction, source position, source attribution, intensity, valence, hazard, exposure, and absorbed dose are different outcomes. One may help predict another; none may silently replace it.

The evidence range for this chapter is C-1152–C-1203: 50 claims are **established** within their stated experiments or authoritative methods, C-1188 is **plausible**, and C-1192 is **disputed**. Those statuses qualify the claim boundaries; they are not votes for the architecture.

Biological observation

Receptor populations provide coverage, not self-describing identities

Mammalian olfaction begins with a large receptor family. Within the receptor panel and odorants studied, individual receptors responded to multiple odorants, individual odorants recruited multiple receptors, and the population pattern changed with concentration (C-1152–C-1154). This supports a distributed measurement basis. It does not supply a universal odor code, open-world chemical coverage, or concentration-independent token.

Chemosensation itself is not one architecture. Manipulations of mammalian sweet and umami pathways provide a scoped dedicated-cell counterexample (C-1155). A receptor-like array must therefore state which chemicals and concentrations it can distinguish, where responses overlap, and which outcomes require another channel. “Combinatorial,” “labelled-line,” or “olfactory” is a description of evidence organization, not an implementation credit.

Bulb and cortex transform concentration, gain, and timing

Divisive normalization in the studied fly antennal lobe scaled projection-neuron responses with pooled receptor activity. Mouse olfactory-bulb and piriform measurements show transformations that can make identity representations more tolerant to concentration, while still preserving useful intensity or concentration-change information in other activity (C-1156–C-1162).

Three constraints follow:

1. concentration tolerance is an outcome to measure, not an invariant built into an anatomical label;
2. suppressing absolute level can damage leak, dose, or safety tasks even when identity improves; and
3. pooled inhibition and recurrence must compete with robust scaling, explicit gain-state estimation, concentration-conditioned inference, and ordinary recurrent models at equal latency and energy.

Receptor adaptation includes causal calcium-dependent feedback, receptor current and spike output can span different concentration ranges, and habituation depends on duration, interval, and odor similarity (C-1163–C-1166). Receptor adaptation, behavioral habituation, short-term sensor recovery, calibration drift, and irreversible poisoning must remain separate states.

Sniffing is part of the observation operator

Rats can alter sniff rate rapidly during discrimination; changing sniffing changes the peripheral-to-bulb filter; early inhalation-locked activity can carry information quickly; and mice can use millisecond-scale sniff-phase differences under scoped protocols (C-1167–C-1171). That timing cannot recover fluctuations already removed by tubing, a chamber, or a slow sensor. End-to-end bandwidth, capture and receipt time, and the realized sampling waveform set the usable support.

Spatial acquisition is also regime-dependent. Serial sampling sufficed in one mouse gradient task, while bilateral temporal correlation supplied odor-motion information in a fly preparation (C-1172–C-1173). Neither result establishes universal stereo or universal serial sampling. Body size, receptor spacing, movement, wind, range, plume intermittency, and sensor bandwidth decide which comparison is informative.

Turbulent plumes turn localization into inference under intermittent evidence

Theory tested against simulation, laboratory, and field observations describes structured whiff and blank statistics rather than a smooth instantaneous gradient. Fine plume structure requires fast ground truth; field gaps change with environment; and several intuitive burst summaries converge too slowly or vary too weakly to guide short-horizon search reliably (C-1174–C-1177).

Animals demonstrate multiple bounded strategies. Moths can surge after odor contact and cast after loss. Walking flies use distinct odor-ON, odor-OFF, and wind transforms, and in irregular plumes their stochastic turns and walk/stop decisions depend on encounter timing rather than

continuous steering (C-1178–C-1180). Contact-correlated slowing in mice is an observation, not proof that slowing is optimal (C-1181).

Robotics already supplies strong nulls. Infotaxis, gas-plus-wind localization, transient processing for slow metal-oxide sensors, reactive off-zigzag search, geometry-aware modular policies, and particle-filter source belief have all been demonstrated in scoped settings (C-1182–C-1187). Compact temporal-memory reinforcement learning is only **plausible** for the studied simulated plumes until unchanged embodied transfer is shown (C-1188).

Sparse piriform activity is not a hardware or energy conclusion

Piriform odor responses can be sparse and distributed rather than neatly topographic; in one rat task, burst-count population information outperformed some precise-pattern accounts (C-1189–C-1191). The degree of sparsity is **disputed** as a universal characterization because it changes with concentration and protocol (C-1192). Longitudinal piriform ensembles can also drift while behavior remains stable (C-1193).

These observations motivate tests for sparse routing, event memory, and remappable readouts. They do not establish fixed semantic neuron addresses, low memory traffic, low organism energy, or a benefit over pruning, compression, low precision, sparse convolution, or dense execution on suitable hardware.

Association, valence, identity, and hazard are separable

Rapid reward-category coding appeared in olfactory tubercle within minutes in one task, while posterior piriform lacked the same explicit code even after overtraining. Arbitrary piriform ensembles could acquire opposite valence under different reinforcement, and innate aversion could be disrupted while learned detection or avoidance remained (C-1194–C-1196). Cortical-amygdala pathways causally contributed to scoped innate odor behavior (C-1197).

The architectural record must therefore preserve odor evidence, reinforcement, context, action, feedback, acquisition time, retention, transfer, reversal, and readout/module identity. Chemical identity, odor category, innate choice, learned choice, pleasantness, irritation, toxicity, external exposure, and hazard remain separately scored.

Mixtures and instruments expose the same ambiguity problem

Animals can learn a target in variable mixtures, but performance degrades with background count and overlap; chemically similar maskers can raise detection thresholds more in the tested regime (C-1198–C-1199). Cross-reactive artificial arrays are already an established sensing baseline, not a novel consequence of receptor analogy (C-1200).

Multi-year metal-oxide sensor data demonstrate drift. Humidity, stability, selectivity, and poisoning belong to the operator state, and validated analytical/safety workflows preserve sampling, calibration, recovery, identification support, exposure units, and uncertainty (C-1201–C-1203). A high closed-panel classifier score cannot turn an unsupported mixture into an identified chemical, a library match into source attribution, or odor detection into safety.

Proposed AI translation

Preserve the whole chemical episode

For episode e , preserve

$$\mathcal{C}_e = (S_e, X_e, A_e, R_e, O_e, K_e, H_e, T_e, U_e, B_e),$$

where:

- S_e records source identity, mixture, release in moles per second, temperature in kelvins, geometry in metres, phase, and motion in metres per second;

- X_e records domain, boundaries, surfaces, airflow in metres per second, pressure in pascals, relative humidity as a dimensionless fraction, temperature, turbulence, reaction, sorption, and chemical background;
- A_e records commanded and realized motion, orientation, sniff/pump flow in cubic metres per second, heater power in watts, valve, preconcentration, purge, query, confirmation, stopping, and abstention;
- R_e records receiver/body, bilateral or array geometry, pose, inlet, tubing, chamber, pump, heater, sensor, saturation, health, and feasible authority;
- O_e is the versioned observation operator: response/recovery, cross-sensitivity, nonlinearity, hysteresis, support, clock, latency, quantization, preprocessing, missingness, and selection;
- K_e records reference-gas composition and uncertainty, blanks, zero/span, flow, device and batch, compensation, age, drift, poisoning, maintenance, calibration validity, and traceability;
- H_e records prior exposure, adaptation, habituation, contamination, cleaning, training, reinforcement, feedback, previous actions, and readout remapping with timestamps;
- T_e declares the literal target, deadline in seconds, loss/utility, abstention policy, exposure rule, and safety constraint;
- U_e names the independent unit: sample, stock, injection, device, batch, day, source, plume realization, site, body, animal/subject, or model seed; and
- B_e is the componentwise ceiling in evidence, labels, standards, channels, actions, metres, seconds, bytes, searches, person-hours, joules, consumables, exposure, replacements, embodied devices, and opportunity.

This is the chemical instantiation of the versioned observation contract. The endogenous-observation candidate owns the coupling between acquisition action and future evidence; the latency-qualified authority envelope owns the action restriction when evidence is slow, stale, saturated, miscalibrated, or poisoned.

Model transport before interpreting the sensor

For analyte i , a minimum transport model is

$$\frac{\partial c_i}{\partial t} + \mathbf{u} \cdot \nabla c_i = \nabla \cdot (D_i \nabla c_i) + R_i(\mathbf{c}, T, P, H_r, \mathbf{x}, t) + q_i(\mathbf{x}, t),$$

where amount concentration c_i is in moles per cubic metre, position $\mathbf{x}$ is in metres, time t is in seconds, velocity $\mathbf{u}$ is in metres per second, diffusivity or declared effective dispersion D_i is in square metres per second, relative humidity H_r is dimensionless, and reaction/loss/phase-transfer R_i and volumetric source q_i are in moles per cubic metre per second. Every term has units of moles per cubic metre per second. Boundaries, buoyancy, droplets, thermal stratification, deposition, and unresolved turbulent fluxes remain explicit when they affect the task.

For channel m sampled at t_n , the measured trace is

$$y_{m,n} = g_{m,v} \left(\sum_i \int_0^\infty h_{m,i,v}(\tau; \mathbf{z}_n) c_i(\mathbf{x}_r(t_n - \tau), t_n - \tau) d\tau, \mathbf{z}_n \right) + \epsilon_{m,n},$$

where channel output $y_{m,n}$ and error $\epsilon_{m,n}$ use the calibrated sensor unit, causal response kernel $h_{m,i,v}$ is in reciprocal seconds, delay τ is in seconds, receiver path $\mathbf{x}_r$ is in metres, version v is dimensionless, and $\mathbf{z}_n$ contains flow, heater, temperature, humidity, interferents, adaptation, age, drift, saturation, and poisoning. A static feature vector is permitted only after this dynamic operator has been tested or shown irrelevant inside the declared support.

Represent non-identifiability instead of forcing a label

Let G_v be the calibrated mixture-to-sensor forward operator and $\mathcal{S}_c$ the supported set of nonnegative composition vectors in moles per cubic metre. For observation $\mathbf{y}$, retain

$$\mathcal{N}_v(\mathbf{y}) = \left\{ \mathbf{c} \in \mathcal{S}_c : \|\mathbf{y} - G_v(\mathbf{c}; \mathbf{z})\|_{\Sigma_y^{-1}} \leq \varepsilon_y \right\},$$

where error covariance Σ_y is in squared sensor-output units, the Mahalanobis norm and tolerance ε_y are dimensionless, and $\mathcal{N}_v$ is the observation-equivalent composition set. If materially different identities, concentrations, exposures, or hazard states remain in that set, the result is ambiguous. The system can acquire another measurement, request analytical confirmation, retain alternatives, or abstain; it cannot convert a prior-selected label into new chemical evidence.

This state links to [reset-coupled staged verification](#): an inexpensive array may screen, but escalation to GC-MS, PTR/SIFT-MS, IMS/FAIMS, or another qualified method must add conditionally useful evidence after sampling, standards, turnaround, analyst time, consumables, exposure, and energy are charged. The analytical method is itself an operator with blanks, recovery, retention, deconvolution, library support, calibration, and uncertainty—not an oracle.

Keep a literal outcome firewall

OUTPUT	NATIVE MEASUREMENT	MUST REMAIN SEPARATE FROM
presence	hits, false alarms, d' , criterion, matrix and concentration	identity or recognition accuracy
chemical identity	confusion/unknown set, standards, retention/spectral support and calibrated probability	odor name, valence or source
concentration	<code>mol/mol</code> , <code>mol/m^3</code> , or <code>kg/m^3</code> , temperature, pressure, bias/error and support	raw sensor output or perceived intensity
mixture	component identity/concentration, recovery, censoring and non-identifiable set	dominant label
direction/position	angular error; position error and coverage in metres; source-off false declarations	contact or instantaneous gradient
source attribution	competing emitters, transport evidence, association and posterior calibration	chemical identity or location alone
association/valence	learning curve, context, reinforcement, retention, transfer, reversal; innate and learned choices separately	identity, toxicity or hazard
exposure/dose	external concentration-time in <code>kg s/m^3</code> ; absorbed dose only with dosimetry	detection or sampling duty cycle
hazard/safety	chemical-, route-, population-, endpoint- and averaging-time-specific rule	odor threshold, intensity, preference or aversion
efficiency	protected outcomes plus evidence, time, bytes, actions, person-hours, consumables, exposure and lifecycle joules	event count, inference power or organism metabolism

Relative gradients are valid only for the field transformation tested

Multiplying an entire positive concentration field by a constant preserves its shape and every local ratio, but adding a background, changing transport, or clipping a receptor does not. The *E. coli* evidence links response rescaling to search only inside finite concentration regimes (C-1542); the social-amoeba evidence adds a density-qualified secrete-and-sense boundary rather than generic density-independence (C-1546).

An active-search policy therefore records which field transformation it is expected to ignore. It is challenged with same-ratio/different-difference and same-difference/different-ratio fields, additive backgrounds, unseen source strengths, saturation, transport change and values near zero. A relative channel may guide search only while its support gate remains valid; an absolute-critical exposure or load stays on a calibrated absolute channel. The full mathematical and protocol boundary is kept in [Interface-qualified scale symmetry](#) and Fixture F-026.

Use action to change observability, not to obtain a free second dataset

At decision time t , choose

$$a_t = \pi_q(\mathcal{H}_t, \hat{\mathbf{c}}_t, \hat{\mathbf{s}}_t, \hat{O}_t, \hat{U}_t, \mathcal{A}_t^{\text{safe}}, \mathbf{B}_t),$$

where $\mathcal{H}_t$ is causally received observation/action history, $\hat{\mathbf{c}}_t$ is concentration/mixture belief in moles per cubic metre, $\hat{\mathbf{s}}_t$ is source belief with position in metres and release in moles per second, $\hat{O}_t$ is operator/condition state, $\hat{U}_t$ is uncertainty, $\mathcal{A}_t^{\text{safe}}$ is the independently constrained action set, and $\mathbf{B}_t$ is remaining budget in its component units.

The action may move or orient the body, change bilateral spacing, sniff or pump, change heater or valve state, purge, resample, request confirmation, stop, or abstain. Its causal value must be tested against fixed, random, replayed, and dose-matched acquisition. Equal wall time is insufficient if one method inhales or pumps more material, experiences more whiffs, travels farther, uses more energy, or accepts more exposure.

For plume search, the system retains a joint belief

$$p(\mathbf{s}, \mathbf{c}_{0:t}, O_t \mid y_{1:t}, a_{1:t}, \mathcal{C}_e),$$

not one gradient arrow. Surge--cast/off-zigzag rules, wind-only anemotaxis, particle-filter belief control, infotaxis, finite-state search, POMDP/MPC/value of information, and matched-memory reinforcement learning remain mandatory nulls. Success, false source declarations, location error, posterior coverage, path in metres, time in seconds, collisions, exposure, and joules are reported separately.

Maintain fast response and slow condition as different states

Use at least two state transitions:

$$\mathbf{r}_{n+1} = f_r(\mathbf{r}_n, \mathbf{c}_n, a_n) + \xi_n, \quad \mathbf{d}_{e+1} = f_d(\mathbf{d}_e, \mathcal{E}_e, m_e) + \omega_e,$$

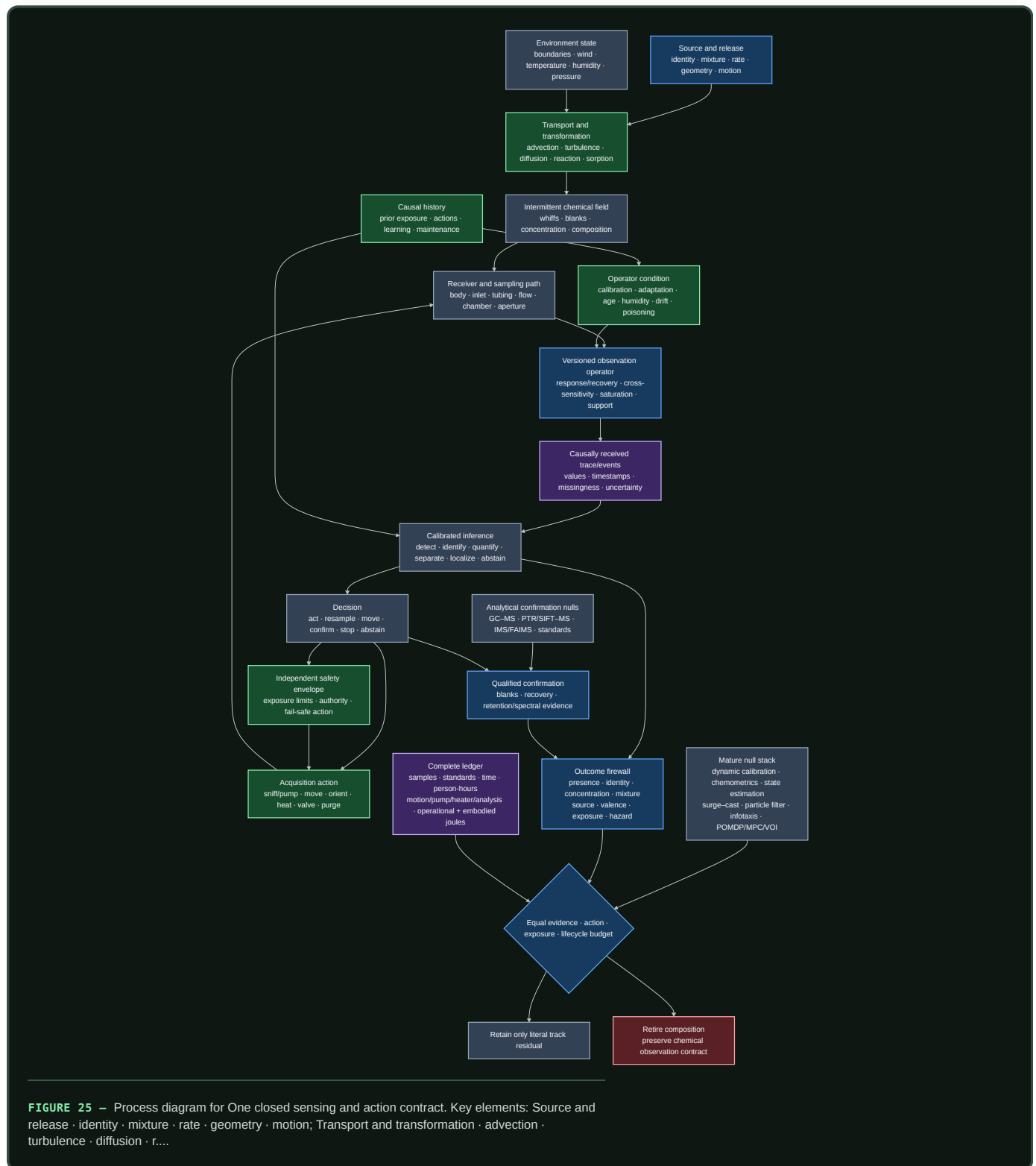
where within-episode state $\mathbf{r}_n$ includes response, adaptation, heater, and recovery; between-episode state $\mathbf{d}_e$ includes calibration, baseline/gain drift, contamination, ageing, and poisoning; concentration $\mathbf{c}_n$ is in moles per cubic metre; cumulative stress/exposure $\mathcal{E}_e$ retains its physical units; and maintenance action m_e records purge, cleaning, recalibration, repair, or replacement. A return to baseline does not prove restored selectivity or calibration. A task residual cannot by itself distinguish environmental change from device change.

The graded assurance envelope binds calibration and condition evidence to the exact operator version. The reversible physical-skill candidate receives credit for coatings, inlets, chambers, filters, heaters, or other physical transforms only after cross-sensitivity, reset, poisoning, replacement, fallback, and fabrication burden are measured.

Preserve evidence for recalibration and future interpretation

Raw traces, calibration/operator history, standards, sample lineage, analytical evidence, and retained physical samples have different reconstruction value. Contract-preserving compaction may replace them only for registered future queries; value- and reconstructability-aware tiering must survive hidden recalibration, changed-library, changed-exposure-rule, and poisoning-investigation queries without future-label leakage.

One closed sensing-and-action contract



Editable source: operator-qualified-active-chemical-sensing.mmd.

Efficiency mechanism

The architecture permits six distinct efficiency mechanisms. Each remains a hypothesis until it improves a protected outcome under F-011's matched budget.

MECHANISM	POSSIBLE SAVING	REQUIRED ACCOUNTING	IMMEDIATE RETIREMENT CONDITION
selective acquisition	avoid samples, motion, pumping, heating, or confirmation that cannot change the decision	sampled volume/mass, whiffs, actions, path, latency, exposure, wear and joules	fixed, random, replayed, or dose-matched acquisition reaches the same frontier
transient/event processing	act on causal onsets, offsets, whiffs and blanks without waiting for slow steady state	physical bandwidth, missed sustained signals, false events, bytes, memory traffic, decoder work and total energy	dynamic deconvolution, derivatives, matched filters, or a finite-state history matches it
cross-reactive population coverage	reuse partially selective channels across chemicals and mixtures	sensor chemistry/area, response support, calibration standards, interferents, unknowns, saturation and replacement	gain follows coverage, SNR, sampled material, or labels rather than the decoder
calibrated normalization and multiscale state	stabilize some identity information while retaining concentration, change, and device condition	absolute-signal error, rare targets, state updates, recurrence, calibration, latency and energy	robust scaling or explicit gain/state estimation matches it, or safety information is erased
staged analytical escalation	use a low-cost screen for easy cases and buy stronger separation/identification only when valuable	aliquots, standards, blanks, turnaround, analyst work, carrier gas, sorbents/columns, vacuum/ionization, exposure and joules	always-confirm, never-confirm, sequential tests, or an ordinary calibrated cascade matches it
qualified compaction and maintenance	retain only evidence needed for registered future queries; recalibrate, clean, remap, or replace only when justified	raw/sample retention, update writes, future-query loss, downtime, labels, maintenance, replacement, people and embodied burden	future recalibration, changed-library, exposure-rule, or poisoning queries cannot be reconstructed

Sparse or event-driven computation is not a seventh saving until its total physical cost is lower. A top- k or thresholded representation can reduce arithmetic while adding normalization, sorting, indices, irregular memory traffic, routing, remapping, idle hardware, missed-event risk, and maintenance. The relevant numerator is accepted task service, not active-unit count.

Lifecycle energy for method q over one accepted service interval is

$$E_q^{\text{life}} = E_q^{\text{data}} + E_q^{\text{train}} + E_q^{\text{move}} + E_q^{\text{pump}} + E_q^{\text{heat}} + E_q^{\text{sense}} + E_q^{\text{separate}} + E_q^{\text{ionize}} + E_q^{\text{infer}} + E_q^{\text{comm}} + E_q^{\text{store}} + E_q^{\text{cal}} + E_q^{\text{maint}} + E_q^{\text{facility}} + E_q^{\text{emb}},$$

where every term is energy in joules and covers data acquisition, training, receiver motion, pumping, heating, sensing, analytical separation, ionization/vacuum, inference, communication, storage, calibration, maintenance, facility overhead, and amortized embodied hardware. Carrier and calibration gases, sorbents, columns, dopants, filters, cleaning agents, samples, emissions, and disposal remain additionally reported in their native physical or lifecycle units.

Human work is

$$H_q^{\text{human}} = H_q^{\text{design}} + H_q^{\text{sample}} + H_q^{\text{label}} + H_q^{\text{cal}} + H_q^{\text{analyze}} + H_q^{\text{tune}} + H_q^{\text{safety}} + H_q^{\text{monitor}} + H_q^{\text{maint}},$$

where every term is in person-hours and roles are separated. An apparent energy gain is rejected if it moves work into sample preparation, calibration, chemical analysis, safety review, data curation, cleaning, or repair without counting it.

Evidence status

EVIDENCE BUNDLE	STABLE CLAIMS	STATUS	ARCHITECTURAL USE AND BOUNDARY
receptor family, combinatorial responses, concentration dependence, taste counterexample	C-1152–C-1155	4 established	justify population-coverage and dedicated-channel comparisons; no universal chemical code
normalization, bulb/piriform concentration transforms, intensity/change and sniff-phase state	C-1156–C-1162	7 established	test joint concentration–identity and explicit gain-state mechanisms; never erase safety-relevant level by default
receptor adaptation, transduction range and habituation	C-1163–C-1166	4 established	require separate response, adaptation, habituation, recovery and slow-condition state
active sniffing, response timing, serial and bilateral acquisition	C-1167–C-1173	7 established	justify causal sampling/body-action tests inside measured end-to-end bandwidth; no universal stereo/serial rule
plume intermittency, measurement bandwidth, environment and weak directional summaries	C-1174–C-1177	4 established	require measured transport, whiff/blank statistics and temporal controls rather than smooth-gradient assumptions
animal and robotic plume navigation, reactive/belief/search nulls	C-1178–C-1187	10 established	establish a regime-dependent policy library and strong robotics null stack; behavior is not optimality proof
compact temporal-memory RL in simulated plumes	C-1188	1 plausible	eligible only as a frozen simulation-to-embodiment hypothesis

EVIDENCE BUNDLE	STABLE CLAIMS	STATUS	ARCHITECTURAL USE AND BOUNDARY
sparse/distributed piriform codes and concentration-dependent sparsity	C-1189–C-1192	3 established; C-1192 disputed	motivate causal sparse/readout tests; no fixed sparseness, hardware, or energy conclusion
representational drift, rapid value learning, flexible and innate valence pathways	C-1193–C-1197	5 established	require remapping cost and separate association, region/readout, innate, learned and hazard outcomes
mixture foreground, masking, arrays, drift, condition and analytical/safety boundaries	C-1198–C-1203	6 established	require unknown/mixture tests, future-device splits, calibration/poisoning state, qualified analytical confirmation and exposure rules

The totals are exactly 50 established, one plausible, and one disputed claim. The established status applies only to the cited biological preparation, behavior, instrument, dataset, method, or authoritative standard. It does not establish that the project composition improves an engineering frontier.

The complete mature null is a composition, not a token baseline:

1. traceable sampling, standards, blanks, duplicates, recovery, flow and calibration;
2. GC--MS/GC--FID/PID and GC×GC where justified, retention indices, authentic standards, PTR/SIFT--MS, IMS/FAIMS, electrochemical/PID, and targeted spectroscopy under their support;
3. dynamic system identification, response/recovery modelling, deconvolution, filtering, robust scaling and temperature/humidity compensation;
4. PCA/PLS, LDA/QDA, calibrated regression/classification, SVMs, trees, ensembles, neural models, open-set detection, conformal/selective prediction, mixture models, domain adaptation and abstention;
5. measured/validated flow models, Kalman/particle filtering, Gaussian-process plume inference, observability and posterior calibration;
6. correlated random walk, gradient and wind baselines, surge--cast, off-zigzag, infotaxis, particle-belief control, finite-state search, POMDP/dual control/MPC/value of information and matched-memory RL; and
7. detector health, poisoning/out-of-support alarms, staged verification, exposure constraints, independent authority, fallback, maintenance, sample lineage, human work, consumables, and lifecycle accounting.

F-011 compares against that full stack. A weak static classifier, uncalibrated e-nose, single gradient controller, or instrument name is not the baseline.

Speculative extensions

The following are experiment-generating compositions. None is promoted by this chapter.

Action-conditioned identifiability

Use the current observation-equivalent set $\mathcal{N}_v(\mathbf{y})$ to select the cheapest safe action expected to separate decision-relevant alternatives. The action might change path, wind-relative orientation, bilateral geometry, flow, heater state, temporal support, or analytical method. This joins Candidate 007 with Candidate 014. It survives only if explicit Bayesian design, value of information, POMDP/dual control, and ordinary staged testing cannot reach the same calibrated decision frontier.

Dual identity--concentration state

Maintain shared evidence with separately protected readouts for chemical/odor identity, absolute concentration, concentration change, and operator condition. Normalization or recurrence may stabilize the identity readout while the other paths preserve dose and condition. A useful implementation must beat concentration-conditioned generative models and explicit gain-state estimators; it is rejected when one task improves by destroying another.

Remappable sparse population memory

Treat sensor/receptor channels and sparse learned units as replaceable evidence contributors rather than permanent semantic addresses. A readout-maintenance layer would detect drift, remap channels, preserve uncertainty, and request labels or calibration selectively. It is worth retaining only if future-time performance improves after update writes, labels, monitoring, downtime, memory traffic, replacement, and energy are charged. Standard recalibration, domain adaptation, ensemble remapping, pruning, compression, and dense low-precision execution remain the nulls.

Operator-matched event front end

Co-design inlet, chamber, sensor physics, heater/pump action, deconvolution, and event thresholds so the retained trace preserves task-bearing whiff/blank and transient information at lower traffic. This is a scoped extension of [Candidate 006](#), not a claim that physical or event-driven sensing is intrinsically efficient. It must transfer across hardware, humidity, drift and plume timescale and must beat a calibrated dynamic model on the same device.

Qualified screen--confirm--retain loop

Compose an inexpensive cross-reactive screen, calibrated abstention, conditional analytical confirmation, and query-aware retention. The screen can provisionally act only inside the [latency-qualified authority envelope](#); [Candidate 010](#) owns escalation; [Candidates 017 and 018](#) own retained evidence. The composition is rejected if an ordinary calibrated cascade or always-confirm policy matches protected risk, latency, and total cost.

Regime-switching plume search

Use calibrated transport, sensor-condition and source beliefs to switch among reactive ON/OFF behavior, wind-relative movement, local search, belief-driven exploration, confirmation, and safe withdrawal. The controller must expose the regime evidence that authorized the switch and must abstain when operator or wind evidence is invalid. It is rejected if one finite-state controller, particle-belief policy, infotaxis, POMDP/MPC, or matched-memory learner reaches the same held-out source-search frontier.

Failure modes

Physics and operator failures

1. **Smooth-gradient fiction:** instantaneous concentration is treated as a stable source direction despite intermittent transport.
2. **Static-vector fiction:** inlet, chamber, response, recovery, hysteresis, saturation, humidity, and prior exposure are discarded before inference.
3. **Bandwidth invention:** millisecond or event information is claimed after the physical transport or sensor has filtered it away.
4. **Mixture over-identification:** a single label is emitted while materially different compositions, concentrations, exposures, or hazards remain observation-equivalent.
5. **Conversion error:** parts per million are converted to mass concentration without molar mass, temperature, pressure, and fraction definition.
6. **Transport/operator confounding:** policy performance is credited to the learner although one arm received better wind, likelihood, sensor dynamics, field truth, calibration, or source prior.

Representation and learning failures

7. **Label-as-mechanism:** “receptor-like,” “bulb,” “piriform,” “sparse,” “temporal,” or “neuromorphic” replaces a causal ablation and literal endpoint.
8. **Coverage-as-decoder gain:** more sensor chemistry, area, concentration, standards, or SNR is attributed to architecture.

9. **Concentration erasure:** identity appears invariant because the system discarded information required for leak, exposure, or safety decisions.
10. **Stable-address assumption:** drifting or replaced sensor/representation units retain fixed semantic addresses without remapping cost.
11. **Event-count efficiency:** fewer active events are reported without bytes, memory traffic, routing, decoding, idle hardware, missed hazards, and lifecycle joules.
12. **Association collapse:** endpoint accuracy substitutes for acquisition curve, reinforcement/context, retention, transfer, reversal, and selective representation/readout intervention.
13. **Valence collapse:** innate choice, learned choice, pleasantness, irritation, toxicity, exposure, and hazard are merged into one score.

Evaluation, reliability, and safety failures

14. **Temporal/batch leakage:** random rows or adjacent windows share stock, dilution, sample, plume seed, sensor, device, batch, calibration, day, site, subject, or analytical run across splits.
15. **Simulation inverse crime:** training and evaluation share CFD mesh, response kernel, source schedule, random seed, or post-test retuning.
16. **No-source omission:** every episode contains a source, so unconditional declarations look successful and false reassurance stays invisible.
17. **Compensation without condition detection:** expected drift is corrected while humidity, contamination, poisoning, replacement, or out-of-support state remains undetected.
18. **Instrument-as-oracle:** GC--MS or another method is credited without sample lineage, blanks, recovery, breakthrough/carryover, separation, retention/spectral support, standards, library scope, and uncertainty.
19. **Odor-as-safety:** detection threshold, intensity, preference, or aversion substitutes for a chemical-, route-, population-, endpoint-, and averaging- time-specific exposure rule.
20. **Self-certified authority:** the same uncertain sensor/inference path defines its own safety envelope and fallback.
21. **Free active sensing:** an adaptive method samples more material, sees more whiffs, moves farther, waits longer, consumes more pump/heater energy, or accepts more exposure than its baseline.
22. **Incomplete lifecycle boundary:** analytical preparation, standards, calibration gases, consumables, motion, pumps, heaters, chromatography, vacuum/ionization, facility power, cleaning, replacement, human work, embodied devices, emissions, and disposal disappear from the ledger.

Any failure that creates the reported advantage retires the architectural claim for that track. The operator record can remain useful even when the proposed mechanism does not.

Measurable predictions

Fixture F-011 implements these predictions with frozen splits, common operator/action budgets, causal ablations, source-off trials, prospective device/time tests, and full resource accounting.

TRACK	TESTABLE PREDICTION	STRONGEST DECISIVE COMPARISON	RETIRE WHEN
T1 concentration--identity	a shared qualified state improves identity across held-out concentration while retaining calibrated absolute concentration	raw/dynamic calibrated models, concentration-conditioned generative inference, divisive normalization and recurrence	identity gain disappears when concentration is protected or relies on seen concentration/matrix
T2 coverage versus architecture	the proposed decoder extracts more task value from an equal channel basis, area, bandwidth and SNR	dense/sparse linear, kernel, tree, Bayesian and neural decoders on matched arrays	gain follows broader chemistry, more sampled material, labels or SNR
T3 normalization	state-qualified normalization improves the joint identity--concentration--rare-target--calibration frontier under saturation and interferences	no normalization, robust scaling, explicit gain-state estimation, divisive and recurrent alternatives	a conventional method matches, or absolute/safety information worsens
T4 temporal code	causal temporal order adds held-out chemical/plume information beyond the measured response operator	instantaneous/derivative features, matched filters, state-space deconvolution and event models	advantage vanishes under held-out inlet/sensor operators or marginal-preserving shuffle

TRACK	TESTABLE PREDICTION	STRONGEST DECISIVE COMPARISON	RETIRE WHEN
T5 adaptive sniff/pump	closed-loop acquisition improves literal decision value per sampled material, exposure, time and joule	fixed-rate, random, replayed and dose-matched schedules with VOI/POMDP control	it receives more dose/opportunity or fixed acquisition reaches the frontier
T6 bilateral/serial/wind	cue value changes predictably with range, plume regularity, body spacing and bandwidth	instantaneous bilateral gradient, lag correlation, unilateral history, wind-only and calibrated fusion	one cue is claimed universally or gain fails held-out bodies/regimes
T7 plume statistics	whiff/blank history contains source-bearing information not captured by simpler causal summaries	mean, peak, slope, duration, frequency, time-since-hit, bilateral lag and full history at equal window	a simpler statistic matches, or prediction uses downstream distance/leakage rather than source evidence
T8 source search	the composition improves success, false declaration, calibrated position, path, time, risk, exposure and energy jointly	random walk, gradient, wind, surge--cast/off-zigzag, infotaxis, particle belief, finite-state, POMDP/MPC and matched-memory RL	B10 or any simpler policy matches on held-out sources/plumes with failures included
T9 embodiment transfer	a frozen simulated policy retains value under measured tubing, sensor, humidity, drift, saturation and action latency	finite-state, system-identified belief control, domain-randomized and recurrent/RL baselines	gain requires post-test retuning or disappears under the physical operator
T10 sparse total cost	sparse/event representation lowers complete accepted-service cost without losing rare, sustained, calibration or hazard signals	dense low precision, pruning, compression, top- k , threshold events, sparse convolution and indexed retrieval	only active count/FLOPs fall, or total bytes/joules and protected task do not improve
T11 drift/readout maintenance	condition-aware remapping improves prospective future-device performance at lower full maintenance cost	frozen readout, scheduled recalibration, state estimation, orthogonal correction, domain adaptation, ensembles and replacement	future labels leak, or conventional maintenance reaches the frontier
T12 mixture/masking	the system recovers or correctly abstains on unseen compositions while preserving target detection and component concentration	calibrated multivariate, nonnegative/generative mixture, open-set and selective-prediction baselines	closed schedules or dominant labels create the score, or non-identifiability is hidden
T13 humidity/poisoning	the system distinguishes reversible condition, drift, contamination, poisoning, replacement and unknown input early enough for safe degradation	no correction, dynamic calibration, supervised/unsupervised adaptation, condition diagnostics, redundancy and fallback	correction works only on expected drift or cannot constrain unsafe action prospectively
T14 tiered analysis	screen--abstain--confirm reduces protected risk/latency/cost across knowns, unknowns, mixtures, blanks and exposure boundaries	always-confirm, never-confirm, sequential probability tests, calibrated cascades and VOI escalation	false reassurance exceeds its ceiling or analytical/human/consumable/lifecycle cost removes the gain

For every track, report literal outcomes, calibrated uncertainty, independent unit, failures, abstentions, unused budget, and the complete cost vector. A residual must replicate across at least two target chemical families, interferent/matrix families, concentration and source/plume regimes, sensor chemistries, manufacture batches, operator/calibration versions, future times, sites, model families, and hardware classes. Active tracks additionally require unseen source positions, plume seeds, bodies, action limits, and paired counterfactual seeds.

If the complete mature stack matches the composition, if no selective ablation isolates value, or if the gain disappears after calibration, analytical work, exposure, maintenance, human effort, consumables, facility and lifecycle energy are charged, retire the architectural residual. Keep the chemical observation contract and the negative result; create no new principle or candidate.

Reliability under mission profiles

Scope

An AI system is not deployed on a nominal device. It is deployed on a physical population: lots, wafers, dies, blocks, packages, boards, power paths, cooling paths, memories, interconnects, converters, sensors, and controllers that begin different and continue to change. Reliability is therefore a relation among a required function, an acceptance criterion, a physical unit, its operating history, its environment, and a time horizon—not a scalar property of a hardware name (C-1002).

This chapter defines the architecture required to preserve accepted AI service under that reality. It separates:

1. time-zero variation;
2. reversible drift;
3. cumulative irreversible degradation;
4. abrupt permanent failure;
5. transient upset;
6. correction or compensation applied by the system; and
7. measurement, calibration, model-support, and classification error.

Those states are not interchangeable. Restored task score after recalibration does not prove physical recovery. A transient error does not establish wear. A monitor alarm does not identify a mechanism. A passed qualification does not cover every future workload or environment (C-1003, C-1022, C-1030).

The chapter joins four maintained artifacts:

1. the semiconductor device and circuit reliability audit, which establishes the evidence boundary;
2. the mission-profile-qualified reliability mathematics, which defines the observation, damage, service, and lifecycle ledgers;
3. Fixture F-008, which tests ten hostile experiment tracks against the complete conventional reliability stack; and
4. the editable degradation-and-recovery diagram.

The result composes existing candidates. It does not allocate a new principle or candidate (C-1053). Its purpose is narrower and more useful: make physical reliability a versioned runtime and lifecycle contract rather than an assumption hidden below the model.

Biological observation

Living systems maintain useful function while their physical components differ, accumulate stress, repair imperfectly, and turn over at different rates. They do not depend on every component remaining identical to an original specification. They combine local maintenance, selective

replacement, redundancy, reserve, containment, feedback, and changes in activity. At larger scales, function can persist even while particular molecules, organelles, cells, or tissue regions are repaired, removed, replaced, or assigned less demanding work.

The useful abstraction is not a particular organ or biochemical mechanism. It is a substrate-independent maintenance contract:

- components have individual histories and unequal remaining capability;
- current output does not reveal all accumulated damage;
- stress and repair occur at several timescales;
- local defects are tolerated only while they remain contained;
- surveillance consumes resources and can itself fail;
- recovery may restore function without restoring the original substrate;
- reserve is finite; and
- replacement is part of operation, not evidence that degradation never occurred.

The translation must stop at that contract. Semiconductor bias-temperature instability, electromigration, radiation upset, solder fatigue, and memory-cell endurance do not become biological processes because they share words such as stress, repair, or recovery. Their causal models, units, evidence, and failure criteria remain physical-domain specific. The analogy contributes questions—what is damaged, what is observed, what can recover, what must be replaced, and what function remains acceptable—not answers.

This boundary is especially important because equal present performance can hide unequal remaining life. Some electrical effects partly recover after rest, while interconnect voids, consumed endurance, package fatigue, and other damage persist (C-1014, C-1017, C-1019, C-1022). The architectural lesson is to carry native margin, reversible state, irreversible damage, compensation, and reserve separately instead of naming every return of output quality “recovery.”

Proposed AI translation

Make the physical population part of system identity

Every deployed unit receives an immutable identity chain for fabrication lot, wafer, wafer coordinates, die, block or array, package, board, power and clock path, cooling path, tester, calibration, hardware, firmware, compiler, model, repair map, and site. Replacement, remapping, recalibration, repair, or firmware change adds a versioned transition; it does not overwrite history.

The population is the denominator. Production yield, accepted yield, early-life failure, field reliability, and wear-out answer different questions (C-1007). Critical area, clustered defects, layout, repair, and edge structure affect yield (C-1008). Spare rows, blocks, experts, or tiles can raise accepted yield only when diagnosis and defect geometry permit repair, and they consume area, routing, test, latency, and residual common-mode reserve (C-1011).

For fabricated unit i and jointly required criterion q , let $A_{i,q} = 1$ when the unit passes that criterion and $A_{i,q} = 0$ otherwise. Joint accepted yield is

$$\bar{Y}_{\text{joint}} = \frac{1}{N_{\text{fab}}} \sum_{i=1}^{N_{\text{fab}}} \prod_{q=1}^Q A_{i,q},$$

where N_{fab} is the number of fabricated units [unit], Q is the number of required criteria [criterion], and $\bar{Y}_{\text{joint}}$ is dimensionless. Dead-on-arrival, untestable, unpackageable, screened-out, and discarded devices remain in N_{fab} . High marginal pass rates do not imply a high joint pass rate when timing, power, noise, memory, and analog constraints are correlated (C-1010).

Population evaluation groups data by lot, wafer, die, region, device, site, and future time. Randomly splitting transactions can leak shared fabrication and aging history (C-1013). Per-device control is credible only after it transfers across those grouped holdouts.

Carry typed physical state

For physical unit i at time t , maintain

$$z_i(t) = (\theta_i, D_i^{\text{perm}}(t), R_i^{\text{rev}}(t), W_i(t), \mathcal{F}_i(t), Q_i^{\text{res}}(t)),$$

where:

- θ_i is time-zero physical state in declared native units;
- D_i^{perm} is cumulative irreversible damage [declared damage unit];
- R_i^{rev} is reversible degradation [same declared damage unit];
- W_i is consumed endurance [cycle, write, or normalized wear unit];
- $\mathcal{F}_i$ is abrupt or latent fault state [state]; and
- Q_i^{res} is remaining correction, timing, thermal, repair, and spare reserve [declared reserve unit].

Keep the evidence record separate:

$$O_i(t) = (o_i, v_i, \Sigma_i^{\text{cal}}, m_i, c_i^L, c_i^U, a_i^{\text{ev}}, \mathcal{V}_i),$$

where o_i is observed telemetry in sensor-native units, v_i is calibration and instrument version [version], Σ_i^{cal} is calibration covariance in squared native units, m_i is observation-availability mask [dimensionless], c_i^L and c_i^U are censoring bounds in native units, a_i^{ev} is evidence age [s], and $\mathcal{V}_i$ is the calibrated validity envelope.

This is the physical-device specialization of the [versioned observation contract](#): telemetry may be compacted, but calibration, missingness, censoring, mission history, and version validity must remain available to the decisions they support.

This separation prevents five common category errors:

OBSERVATION	WHAT IT DOES NOT PROVE
a device began in a slow or leaky tail	that it has aged
a parameter moved and later returned	that irreversible damage was absent
a task output recovered after calibration	that native physical margin recovered
a fault disappeared after retry	that it was harmless or non-recurring
a monitor stayed quiet	that the protected path stayed inside margin

Time-zero mismatch and systematic gradients do not reduce to one global offset (C-1009). Device-specific trap activity can also produce stochastic tails that a deterministic shift misses; that stronger claim remains plausible rather than established (C-1012). Tester accuracy, calibration, sampling, guard bands, and the conformity rule determine false acceptance and rejection near limits (C-1006).

Drive degradation models from the actual mission profile

For episode e , record the measured mission profile

$$M_e = \left\{ u_n, V_n, f_n, T_n, J_n, a_n, \phi_n, \dot{D}_n^{\text{ion}}, c_n, r_n, \Delta t_n \right\}_{n=0}^{N_e-1},$$

where u_n is workload class [class], V_n measured voltage [V], f_n operation or clock rate [Hz], T_n absolute temperature [K], J_n current density [A m^{-2}], a_n activity [dimensionless], ϕ_n particle flux [particle $\text{m}^{-2} \text{s}^{-1}$], $\dot{D}_n^{\text{ion}}$ ionizing-dose rate [Gy s^{-1}], c_n cooling state [state], r_n route and protection state [state], Δt_n interval duration [s], and N_e interval count [interval]. Commanded voltage, nominal temperature, calendar age, and mean utilization remain useful metadata but cannot replace the measured history.

The history matters because electric field and switching affect interface and oxide damage (C-1015); current density, waveform, geometry, material, and temperature affect electromigration (C-1017); local power changes temperature, which feeds back into delay, leakage, material transport, and degradation (C-1018); and thermal cycles can damage packages even when average temperature is similar (C-1019).

For physical mechanism k , a mission-qualified damage proxy is

$$D_k(t) = \int_0^t r_k(V(\tau), T(\tau), J(\tau), a(\tau), \phi(\tau), u(\tau)) \, d\tau,$$

where r_k is damage rate [damage unit s^{-1}], τ is time [s], and $D_k(t)$ is accumulated mechanism-specific damage [damage unit]. A scalar “hardware age” can summarize this vector for a particular decision, but it cannot replace the underlying mechanism and support record.

The analytical illustration below applies one hypothetical Arrhenius-rate model to two constructed temperature histories. It demonstrates the accounting consequence of nonlinearity; it is not calibrated device damage.

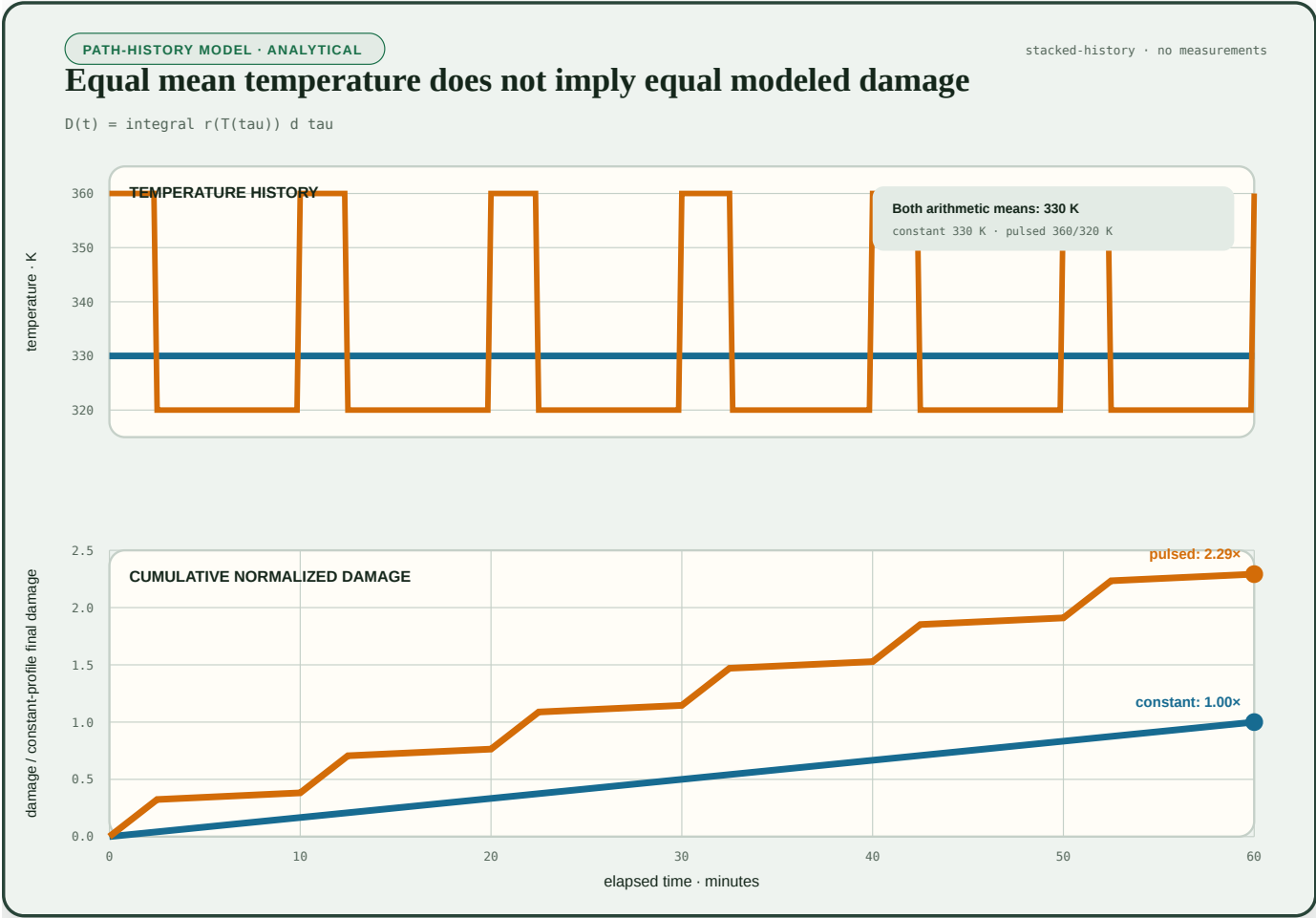


FIGURE 26 – Two devices can have the same elapsed time and arithmetic-mean temperature yet accumulate different modeled damage because the rate is nonlinear in temperature; a nominal or averaged operating point cannot replace mission history.

Treat accelerated tests as supported models, not timeless constants

Accelerated tests are useful only while stress and use conditions share the relevant mechanism and the fitted model remains supported (C-1004). Excessive temperature, field, current, humidity, cycling, or radiation can change the failure mechanism. Dielectric breakdown distributions, for example, require stable mechanism and spatial assumptions for area or Weibull extrapolation (C-1016).

Each extrapolation therefore travels with:

1. material stack, geometry, fabrication population, and failure criterion;
2. stress variables, waveforms, duty cycles, recovery intervals, and sampled range;
3. fitted mechanism, parameter intervals, residuals, and model version;
4. censored and failed units, missingness model, and failure-analysis agreement;
5. distance from training support and any detected mechanism transition; and
6. held-out use-like coverage and false-safe rate.

Zero observed failures is not zero hazard. Exposure, censoring, distribution, sampling, and interval coverage determine the supported upper bound (C-1005). Temperature, voltage, current, fabrication state, and packaging may also interact; simple addition of constant failure rates can then misstate system risk. That cross-mechanism claim is plausible and must be tested rather than assumed (C-1020).

Couple sparse execution to temperature and wear

Conditional routing can reduce switching and data movement while repeatedly selecting the same experts, memory banks, converters, links, or power regions. The same logical sparsity may therefore lower total work but raise local duty cycle, temperature, current density, droop, and accumulated wear. This concentration effect is plausible, not yet an established universal outcome (C-1021). Nominal area, aggregate load, or global utilization likewise cannot identify the tail of local contact, thermal, write, traffic, or damage exposure (C-1498).

The **sparse and predictive compute** chapter therefore gains a physical routing state. Every routing decision observes or estimates:

- spatial power and temperature;
- current density and voltage droop;
- thermal-cycle and recovery history;
- timing, SRAM, analog, and interconnect margin;
- write, program, erase, and remap counts;
- native degradation and compensation;
- remaining spares and repair paths; and
- migration, calibration, replay, cooling, and replacement cost.

A route is efficient only if it improves the lifetime-adjusted accepted-service frontier. “Fewer active parameters” and “lower average chip power” are not physical reliability results. Local temperature and thermal time constants also constrain how quickly a controller may act (C-1036).

Put faults through a typed containment ladder

Radiation and electrical faults differ by environment, particle, energy, cross-section, operating state, persistence, geometry, and consequence (C-1023). A stored-state upset, transient pulse, destructive latch-up, burnout, and cumulative dose degradation require different responses (C-1024). Field populations can also violate convenient proxies: production DRAM and flash studies found error relationships that did not reduce to simple transient or wear assumptions (C-1025, C-1047).

Every fault event therefore carries:

- physical or injected provenance;
- spatial geometry: bit, word, row, bank, chip, link, route, or domain;
- occurrence time and duration [s];
- persistence: transient, intermittent, or permanent;
- common-cause identity;
- activation and masking path; and
- external-side-effect state.

The containment ladder is explicit:

1. **detect:** syndrome, residual, shadow sample, watchdog, or reference channel;
2. **correct locally:** ECC, retry, refresh, verify-and-program, or bounded compensation;
3. **replay before commitment:** discard provisional work and repeat from a protected checkpoint;
4. **scrub or revalidate:** read, correct, rewrite, calibrate, or retest before errors accumulate;
5. **contain the domain:** isolate a codeword, block, expert, chiplet, clock, supply, or route;
6. **repair or remap:** consume a spare, migrate state, change mapping, or replace a failed component;
7. **derate or repurpose:** allow only work inside the reduced qualified envelope; and
8. **retire:** stop accepting protected work when containment or reserve fails.

A SEC–DED code has a specific one-error-correction and two-error-detection contract; burst, chip, address, decoder, correlated, and higher-multiplicity faults need different geometry or codes (C-1027). Scrubbing shortens accumulation time but costs bandwidth, energy, controller activity, and sometimes endurance (C-1028). Replication works only outside shared supply, clock, thermal, radiation, design, voter, software, and update failure domains (C-1029). Logical and application masking are measured because they may suppress or amplify a device event before an accepted output (C-1026).

This ladder joins the staged verification candidate with severity-ordered containment. Detection alone never proves diagnosis, containment, repair, side-effect absence, or future reliability (C-1030). Fault injection is accepted only to the extent that its location, timing, duration, activation, workload, and observation represent the target physical population (C-1031).

Bind margin authority to fresh independent evidence

Lower voltage can reduce switching energy approximately with V^2 , but also increases delay and can trigger timing, memory, analog, or retention failure; the useful point depends on device and workload (C-1032). Shadow sampling, canaries, replica paths, ring oscillators, and thermal sensors help only while they track the protected circuits across space, data, time, aging, and regime (C-1033, C-1034). Body bias changes speed and leakage but also junction and technology-specific limits (C-1035).

Let $m_i^{\text{lb}}(t)$ be a conservative lower bound on physical margin in a declared unit, $c_i^{\text{op}}(t)$ the margin consumed by the proposed operating point in the same unit, r_i^{min} the required reserve, and a_i^{ev} evidence age [s]. Adaptive authority is admissible only when

$$m_i^{\text{lb}}(t) - c_i^{\text{op}}(t) \geq r_i^{\text{min}}, \quad a_i^{\text{ev}} \leq a_{\text{max}}, \quad x_i(t) \in \mathcal{V}_i,$$

where a_{max} is the maximum allowed evidence age [s], $x_i(t)$ is the current operating covariate vector, and $\mathcal{V}_i$ is the validated envelope. If any condition fails, the controller moves to a preregistered safe point, fallback route, or non-accepting state.

The optimizing controller is not its own final assurance authority. The claim that a learned controller needs an independent protection layer is plausible and is tested by corrupting controller, monitor, regulator, clock, firmware, calibration, and fallback paths (C-1037). Approximate state

may cross into model data or tolerant arithmetic only when exact addressing, control, accounting, safety, and acceptance remain protected (C-1038). This extends the latency-qualified authority candidate and the graded assurance candidate.

Treat analog and in-memory computation as changing physical state

Resistive arrays can perform vector–matrix products through stored conductance and circuit laws, reducing some weight movement while making device, wire, and peripheral state part of the computation (C-1039). The programmed matrix is therefore not just a tensor checkpoint. For conductance matrix G^0 [S], write

$$G(t) = G^0 + \Delta G^{\text{prog}} + \Delta G^{\text{drift}}(t, T) + \Delta G^{\text{cycle}} + \Delta G^{\text{stuck}},$$

where each ΔG term is in siemens [S] and denotes programming error, time- and temperature-dependent drift, cycling variation, or stuck-cell error. For input voltage vector v [V], ideal current is $i = Gv$ [A], but the accepted system result must include source and access resistance, line drop, sneak paths, converter limits, peripheral noise, tiling, communication, digital completion, and calibration (C-1042).

Programming is state dependent and variable, and verify-and-program consumes pulses, time, and endurance (C-1040). PCM drift depends on device, state, elapsed time, temperature, and reference (C-1041). Endurance is finite, variable, and acceptance-threshold dependent (C-1045). Hardware-aware training can recover performance for represented nonidealities but does not establish transfer to unseen devices, lots, drift laws, temperatures, faults, or correlations (C-1043). A high-precision residual can correct a physical proposal when conditioning and error bounds permit, but converters, digital work, and iterations remain in the ledger (C-1044).

The architecture therefore treats a physical array as a versioned, calibrated, wearing service provider. The reversible physical skill candidate must expose device population, program distribution, drift, endurance, peripheral work, residual correction, revalidation, and digital fallback. The adaptive topology candidate may rotate or migrate operators only after movement, recalibration, state reconstruction, and wear are charged.

Manage repair, placement, retention, and replacement together

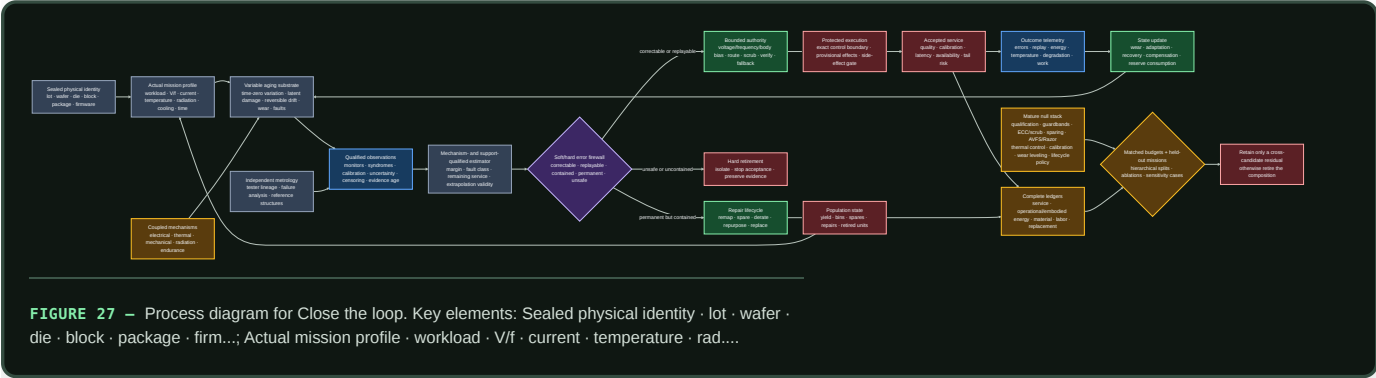
Changing logical-to-physical mapping can spread wear, but it costs metadata, movement, latency, energy, recovery logic, and resilience to adversarial writes (C-1046). The value-and-reconstructability candidate may improve placement only if independently measured value or reconstruction cost adds benefit beyond ordinary endurance-aware wear leveling. The semantic-compaction candidate must preserve fault, calibration, remap, firmware, and incident queries across medium aging.

Each physical unit follows a lifecycle state machine:

STATE	ALLOWED ACTION	EVIDENCE REQUIRED
qualified	accept work inside the current envelope	fresh calibration, margin, correction, and version evidence
degraded	derate voltage, frequency, precision, load, or duty cycle	bounded service and risk under the reduced envelope
repairable	scrub, reprogram, remap, replace a block, or consume a spare	diagnosed containment and successful post-repair qualification
repurposable	assign a less demanding service class	independent assurance that the new class fits remaining capability
replaceable	transfer state and service to another unit	replacement inventory, migration, embodied cost, and recovery evidence
retired	isolate and stop protected acceptance	any hard limit, uncontained fault, invalid evidence, or exhausted reserve

No transition resets the ledger. Repair does not erase the original fabrication burden; replacement adds another one. Whether guardbands, cooling, calibration, refresh, repair, and longer life beat replacement across the whole lifecycle remains plausible and case dependent (C-1051).

Close the loop



Editable source: [mission-profile-qualified-degradation-recovery.mmd](#). The working architecture supplies the runtime control plane; system synthesis supplies execution and maintenance timescales; the energy model supplies the accepted-service denominator; and operator-qualified sensing supplies the rule that physical observations remain versioned and calibration-bound.

Efficiency mechanism

Reliability work is not pure overhead. It can create efficiency by recovering margin that would otherwise be reserved uniformly, preventing expensive recomputation or replacement, and matching physical resources to the service they can still provide. Each saving has a corresponding debit.

MECHANISM	POTENTIAL SAVING	REQUIRED DEBIT AND DECISIVE NULL
population-qualified operating points	less worst-case voltage, timing, cooling, and test margin	sensors, characterization, controller area, calibration, tail risk; compare corners, binning, AVS, Razor, and body bias
conditional routing	fewer switches and bytes moved	local heat, droop, migration, concentrated aging, monitoring, reserve; compare uniform and electrothermal-wear-aware routing
staged correction	avoid full replication or discard by correcting, scrubbing, or replaying locally	ECC bits, bandwidth, verification, latency, write endurance, common causes; compare the strongest fixed correction stack
repair and remapping	retain usable capacity after local defects or wear	spares, area, movement, metadata, recovery, requalification; compare mature redundancy and wear leveling
physical compilation	reduce repeated programmable movement or arithmetic	design, fabrication, yield, converters, calibration, drift, utilization, lifetime; compare matched digital hardware
derating and repurposing	extract safe service from reduced capability	lower throughput or quality, routing complexity, assurance work, longer operating energy; compare replacement and ordinary asset management
timely replacement	reduce use-stage energy or risk	new fabrication, packaging, migration, material, downtime, and stranded reserve; compare repair and life extension under sensitivity cases

For transaction j , define $A_j = 1$ only when quality, calibration, latency, constraint, and fault-containment requirements are all satisfied; otherwise $A_j = 0$. Accepted service is

$$S_{acc} = \sum_{j=1}^{N_{tx}} A_j \omega_j,$$

where N_{tx} is transaction count [transaction] and ω_j is registered service value [service unit/transaction]. Silent corruption, miscorrection, or an escaped unsafe effect makes $A_j = 0$ regardless of average task score.

Complete lifecycle energy is

$$E_{life} = E_{fab} + E_{package} + E_{test} + E_{op} + E_{repair} + E_{replace} + E_{eol},$$

where every term is energy [J] allocated by a published rule, and operational energy includes compute, memory, movement, conversion, monitoring, correction, calibration, cooling, idle, and recovery. Energy intensity is

$$\eta E = \frac{E_{\text{life}}}{S_{\text{acc}}} \text{ [J/accepted service]},$$

and is undefined when $S_{\text{acc}} = 0$. Material [kg by category], carbon [kg CO₂e under a versioned inventory], person-hours [person-hour by role], availability [dimensionless], tail latency [s], data loss [event], repair [repair], and replacement [replacement] remain separate protected outcomes.

Fabrication is inside the boundary. Primary inventories show substantial electricity, fuels, ultrapure materials, gases, water, and infrastructure burden, with strong process and allocation dependence (C-1048). Facility idle and support energy make energy per good die depend on yield, throughput, rework, and utilization (C-1049). Process, idle, rest, and sleep modes and non-electric utilities therefore require explicit rates, durations, conversions, and boundaries (C-1050).

Moving recurring work into an ASIC or physical array can reduce repeated programmable work only when reuse, utilization, yield, lifetime, calibration, and future stability amortize the commitment (C-1052). That crossover—not nominal operations per joule—is the efficiency claim tested by F-008.

Evidence status

The stable ledger contains 52 claims for this chapter: 47 established and 5 plausible. It contains no speculative or disputed claim in this range.

CLAIMS	LEDGER STATUS	WHAT THE CHAPTER MAY USE THEM FOR
C-1002–C-1011	10 established	functional reliability scope, qualification limits, acceleration, censoring, metrology, yield, hierarchical population, joint constraints, and repair
C-1012	1 plausible	device-specific stochastic aging tails; must remain a tested model rather than a default
C-1013–C-1019	7 established	hierarchical leakage control, reversible BTI observation, HCI, dielectric breakdown, electromigration, electrothermal feedback, and package fatigue
C-1020–C-1021	2 plausible	cross-mechanism coupling and sparse-route aging concentration; require combined-stress and lifetime-frontier tests
C-1022–C-1036	15 established	native recovery distinction, radiation, field evidence, masking, ECC, scrub, replication, containment, injection limits, voltage scaling, monitors, body bias, and thermal control
C-1037	1 plausible	independent protection for learned or optimizing hardware authority; test controller and fallback faults
C-1038–C-1050	13 established	exact/approximate firewall, physical arrays, programming, drift, peripherals, transfer, residual correction, endurance, wear placement, field proxies, and lifecycle inventories
C-1051	1 plausible	lifecycle ranking of guardband, maintenance, repair, life extension, and replacement
C-1052–C-1053	2 established	physical compilation as amortization and the no-new-invariant disposition

The 47 established claims support the measurement constraints, physical mechanisms, mature correction methods, and lifecycle boundary. They do not show that this project’s full composition beats the complete conventional stack. The five plausible claims identify exactly where confirmatory work is needed: stochastic device tails, coupled mechanisms, sparse-wear concentration, independent authority protection, and lifecycle policy ranking.

F-008 therefore uses a complete mature null: fabrication statistics and design margin; mechanism-based qualification; screening, binning, repair, and redundancy; ECC, interleaving, scrub, replay, and replication; AVS, Razor, body bias, droop and thermal control; degradation and wear management; approximate and mixed-precision safeguards; measured analog/in-memory paths; endurance-aware placement; and condition-based lifecycle management. A composition that beats a weakened subset but not that stack has no residual.

Speculative extensions

Mechanism-qualified physical state estimator

Learn a compact state that predicts native margin, reversible drift, irreversible damage, fault class, remaining reserve, and model-support distance from sparse fleet telemetry. The estimator must expose which mechanism and population support each prediction. Compare it with mechanism-specific engineering models, hierarchical mixed effects, conventional prognostics, and a generic learned health score. Retire it if the compact state gains apparent accuracy by pooling incompatible mechanisms or leaking future failure evidence.

Reliability-aware expert routing

Extend conditional routing so expert utility is divided by the marginal electrothermal and wear consequence of choosing its physical location now. The router may rotate, migrate, or derate experts, but must preserve quality, calibration, latency, fault domains, and reconstruction obligations. Compare it with energy-minimal, thermal-aware, wear-rotating, and cumulative-damage-aware policies. The extension is useful only if a held-out lifetime frontier improves after monitoring and migration are charged.

Self-testing physical operators

Interleave task execution with low-cost reference operations that identify array drift, converter change, stuck cells, timing loss, and monitor failure. Reference scheduling becomes a value-of-information problem under endurance and availability constraints. The self-test never becomes its own assurance root; an independent reference or conservative fallback remains necessary for protected decisions.

Query-preserving reliability traces

Compress physical telemetry while preserving registered future queries: incident reconstruction, acceleration-model refit, fault-geometry audit, calibration lineage, repair qualification, lifecycle allocation, and replacement decision. Compare the retained state with raw traces, conventional logs, sufficient- statistic storage, and recomputation. If a later registered query cannot be answered with qualified uncertainty, the compaction is rejected.

Service-aware turnover market

Treat qualified hardware capability as a changing portfolio. A scheduler can move precise, high-consequence, or high-write work toward units with appropriate margin and move tolerant work toward derated units. Repair, repurpose, and replacement decisions then optimize accepted service under separate risk, energy, material, work, and inventory constraints. This remains speculative until prospective cohorts outperform ordinary condition-based maintenance and replacement optimization across registered sensitivity cases.

Failure modes

SIGNATURE	INTERPRETATION	REQUIRED RESPONSE
nominal accuracy is reported from selected good devices	fabrication population and yield loss are hidden	restore all-unit denominator; group by lot, wafer, die, site, and future time
one deterministic aging shift fits the mean but misses per-device tails	time-zero state and stochastic variation were collapsed	retain hierarchical state and calibrated tail intervals; test C-1012 rather than assuming it
task score returns after rest, voltage increase, calibration, or remapping	compensation or reversible drift is being called physical recovery	report native margin, reversible state, irreversible damage, and compensation separately
accelerated data fit well but fail use-like low stress	model crossed support or mechanism	withdraw extrapolation and authority; repeat with mechanism-qualified design and censoring
zero failures are treated as zero risk	exposure and censoring were ignored	publish upper bounds, intervals, missingness, and all censored units
sparse routing saves dynamic energy while a few blocks heat and wear rapidly	logical activity was detached from physical placement	compare lifetime-adjusted accepted service with electrothermal and endurance ledgers
random bit flips show resilience	injection does not represent radiation, timing, burst, decoder, permanent, or common-cause geometry	calibrate the fault population and withhold physical classes

SIGNATURE	INTERPRETATION	REQUIRED RESPONSE
ECC corrects words but external side effects or controller state are corrupted	correction was not containment	add provisional execution, side-effect gates, replay, domain isolation, and incident tracing
replication fails under shared supply, clock, thermal, software, or voter state	copies share a failure domain	redraw physical domains or keep the failure unmitigated
voltage controller trusts a stale or co-degrading monitor	authority outlived its evidence	enforce evidence-age, tracking, independent lower bound, and fallback
learned controller suppresses alarms that would reduce its performance	assurance is inside the optimizer's objective loop	remove final authority from the learner; use a protected independent limit
analog core wins on array energy but loses after DAC/ADC, host, calibration, cooling, and writes	component work was substituted for service work	report complete sensor- or memory-to-accepted-output crossover
hardware-aware training fails on a new lot or combined nonideality	the device-error simulator was overfit	expose support distance, abstain, recalibrate, or route to digital fallback
wear leveling moves data more than it saves or exposes high-value placement	metadata, movement, or attack surface dominates	revert to the strongest endurance-aware null or narrow the value policy
a repaired unit returns to service without requalification	repair was treated as erasure of history	create a new version, preserve consumed reserve, and rerun the applicable envelope
old hardware is retained because manufacture is sunk, or replaced because the new device uses less power	lifecycle ranking uses one boundary	compare keep, derate, repair, repurpose, and replace with common inventories and sensitivity cases
one score combines energy, safety, reliability, material, and work	protected outcomes can compensate for each other invisibly	restore the outcome firewall and apply hard constraints before Pareto ranking

Hard retirement applies when a protected effect escapes containment; the upper risk bound exceeds its limit; native margin or reserve crosses a hard lower bound; evidence, calibration, or version leaves the qualified envelope without a safe fallback; a controller or repair path bypasses the transaction boundary; a mechanism changes outside model support; or incident lineage is no longer auditable. Retirement stops protected acceptance and preserves evidence. A unit may still be isolated for analysis or later qualified for a different service.

Measurable predictions

The predictions map one-to-one to the ten hostile tracks in F-008. All comparisons use the same physical cohort or preregistered blocked allocation and matched fabrication, sensing, protection, compute, energy, material, labor, risk, reserve, and wall-time budgets.

- Hierarchical yield transfer.** A population-qualified mapper will improve joint accepted and post-aging yield beyond corners, screening, binning, repair, and regional calibration on held-out dies, wafers, lots, sites, and future time—or per-device adaptation is retired.
- Use-condition extrapolation.** A mechanism- and support-qualified survival model will achieve registered interval coverage and fewer false-safe predictions on withheld use-like low stress than a single apparent acceleration model, while retaining censored units and failure-analysis disagreement.
- Sparse lifetime frontier.** Electrothermal-wear-aware routing will improve lifetime-adjusted accepted service beyond uniform, energy-minimal, thermal-aware, and ordinary wear-rotation policies; short-run switching energy alone will not predict the winner.
- Fault-geometry match.** A typed correction, scrub, replay, sparing, and containment policy will reduce SDC and escaped effects on withheld fault geometries beyond the strongest fixed compatible stack at equal area, bandwidth, latency, energy, wear, reserve, and availability—or the adaptive composition is removed.
- Evidence-age authority.** A voltage controller with a fresh independent margin bound and hard fallback will deliver more accepted service per joule than datasheet voltage, characterized DVFS, canary AVS, and Razor/replay without exceeding timing, SRAM, analog, SDC, escape, or damage limits under monitor, regulator, clock, calibration, and policy faults.
- Approximation containment.** Typed approximation with exact control, verification, and fallback will preserve rare protected outcomes under error, distribution, and objective shift better than untyped approximation after verifier energy and latency are charged.
- Physical-compute crossover.** A measured analog or in-memory path will beat matched quantized digital hardware only in preregistered regions of operator shape, precision, reuse, device state, temperature, yield, endurance, and peripheral work; the digital route will win outside those regions.

8. **Nonideality support.** Hardware-aware training with support detection, residual monitoring, calibration, and fallback will produce fewer confident silent failures than ordinary hardware-aware training on held-out lots, drift ages, correlations, converter laws, stuck cells, wires, and compound nonidealities.
9. **Wear and value placement.** Independently qualified value and reconstructability will improve accepted service before failure beyond Start-Gap-class and endurance-aware placement under held-out skewed, shifting, burst, and adversarial writes—or value-aware placement is retired.
10. **Lifecycle crossover.** A telemetry-qualified keep, derate, repair, repurpose, or replace policy will remain Pareto-competitive across registered fabrication, electricity, utilization, repair-yield, workload-growth, material, and replacement cases. No universal maximum-life or replace-early rule is predicted.

The cross-candidate architecture passes only if at least one preregistered accepted-service or lifecycle outcome improves beyond the complete mature null, no hard constraint fails, the effect survives population and mission-profile holdouts, and the gain remains after calibration, correction, recovery, failed units, spare consumption, repair, replacement, embodied energy, material, and human work are charged. Otherwise the relevant component—or the composition—is retired without creating a new principle or candidate.

Physical computation requires six boundaries

Scope

“Energy per operation” is meaningful only after both *operation* and *energy boundary* are fixed. A logical erasure bound, terminal energy of one device, energy of a clocked circuit, wall-plug energy of a workload, facility energy, and fabrication-to-retirement burden are related measurements, but they are not substitutes. Crossing those boundaries without changing the claim is the main category error this chapter prevents (C-1100, C-1151).

The working contract has six layers:

BOUNDARY	QUESTION ANSWERED	EVIDENCE REQUIRED	INVALID SHORTCUT
1. fundamental generalized erasure	what is the minimum expected work for this declared physical-state transformation?	initial distribution, Hamiltonian, bath temperature, side information, correlations, final error, duration, controls, cycle closure	$k_B T \ln 2$ per gate, FLOP, token, parameter, or model
2. device transition	what energy or heat crossed this device boundary during the transition?	waveform, terminals, parasitics, duration, temperature, state preparation, transition-error distribution, calibrated instruments	theoretical minimum or simulated internal energy as measured device energy
3. circuit and control	what did the complete physical circuit spend?	clock or power clock, control, wires, converters, leakage, sensing, ancillae, history, correction, I/O, reset	active element or reversible truth table alone
4. workload and data movement	what did the implemented system spend per accepted useful outcome?	software, precision, hierarchy, bytes moved, routing, utilization, idle, retries, quality, latency, throughput	peak TOPS/W, TDP, arithmetic count, or one kernel
5. facility and cooling	what facility energy is attributable under a declared interval and allocation?	IT and facility meters, cooling, power conversion, network/storage share, site, weather, interval, PUE category	generic PUE multiplier or PUE as carbon intensity
6. embodied lifecycle	did operational savings repay fabrication and ownership burden over delivered service?	yield, packaging, transport, deployment, maintenance, utilization, support life, replacement, end of life, geography, uncertainty	operational electricity alone

The evidence base is the information thermodynamics and physical computation audit, normalized as C-1100 through C-1151. The maintained quantitative definitions are in the boundary-qualified mathematical contract, and the twelve hostile comparisons are in Fixture F-010. The result composes current candidates; it adds no principle or candidate.

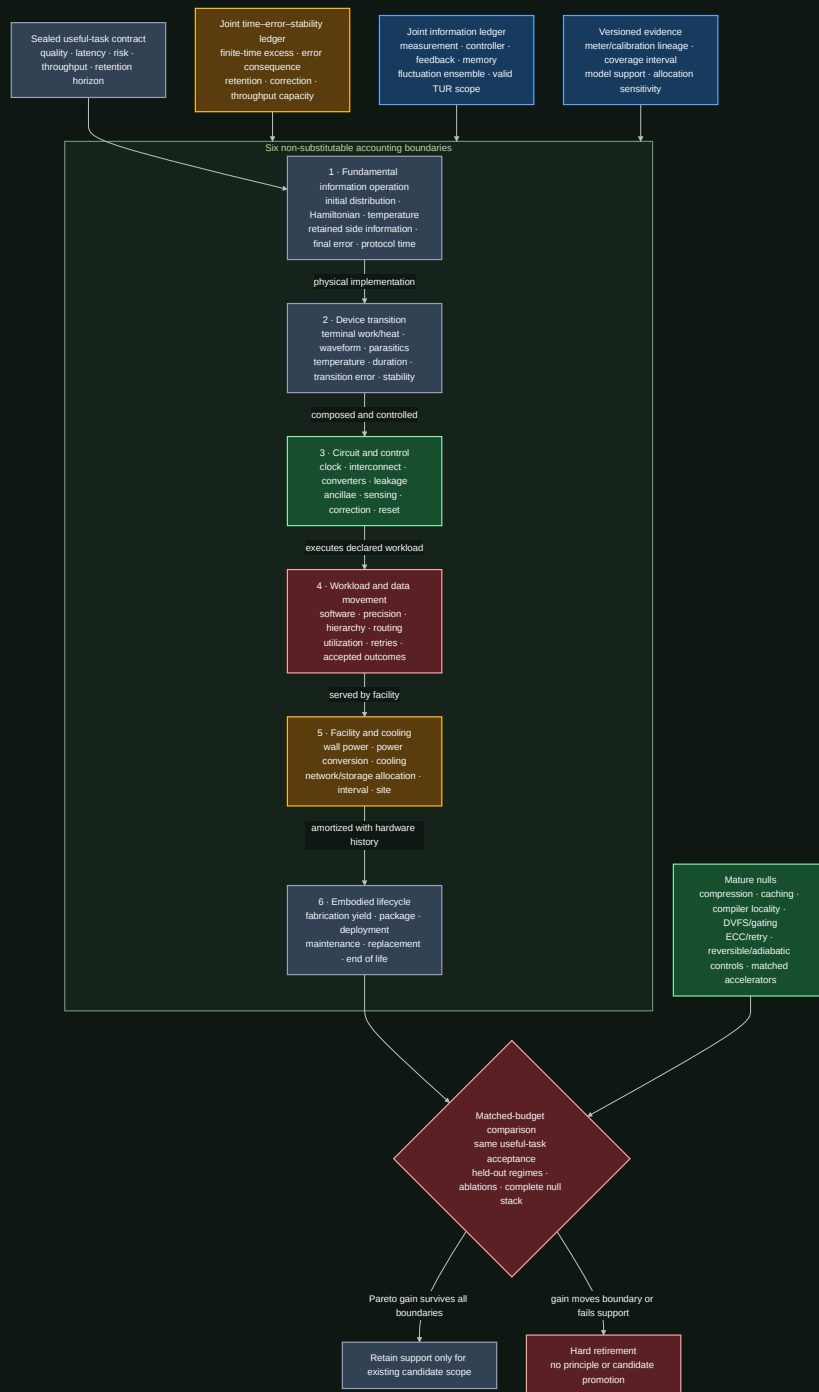


FIGURE 28 – Process diagram for Scope. Key elements: Sealed useful task contract · quality · latency · risk · throughput · ...; Six non substitutable accounting boundaries.

Editable source: boundary-qualified-physical-computation.mmd.

The useful outcome is the firewall. Every lower-level saving must survive the next boundary without lowering quality, raising unacceptable risk, missing the latency/throughput contract, or exporting work.

Biological observation

Biochemical sensing and adaptation make the boundary problem concrete. In the audited models, copy number, integration time, receptor statistics, energy supply, precision, and response speed remain distinct resources (C-1135, C-1136). A result for one biochemical network does not

become a universal information price, and it does not set accelerator energy. Its useful contribution is a measurement discipline: name the physical states, dynamics, observation interval, error variable, and supplied work.

Feedback experiments add a second lesson. A controlled subsystem can extract work or cool while the sensor, memory, controller, actuator, or coupled demon dissipates energy. The joint boundary restores the balance (C-1130, C-1131, C-1132, C-1133). Continuous information flow can be assigned to parts only when the joint transition structure supports that decomposition (C-1134).

Three observations transfer:

1. sensing, state retention, response, and reset are physical parts of the same loop;
2. precision, speed, stability, and energy form a frontier, not one scalar; and
3. a subsystem benefit is provisional until the coupled system closes.

The transfer stops there. The audited biological models do not establish a digital-training lower bound. Predictive-information and stochastic-learning connections remain plausible rather than established for deployed AI (C-1137, C-1138).

Proposed AI translation

Begin with an accepted useful outcome

For requested outcome j , define a preregistered acceptance indicator

$$A_j = \mathbf{1}[Q_j \succeq q_j \wedge L_j \leq L_j^{\max} \wedge \rho_j \preceq \rho_j^{\max} \wedge \Theta \geq \Theta_j],$$

where $A_j \in \{0, 1\}$ is acceptance [dimensionless], Q_j and q_j are measured and required task-quality vectors in the same task-native units, L_j and $L_j^{\max}$ are measured and maximum latency [s], ρ_j and $\rho_j^{\max}$ are measured and maximum risk vectors [failure/request], and Θ and Θ_j are delivered and required throughput [accepted outcome/s]. Rejection, abstention, timeout, retry, silent corruption, and blocked side effects remain in the request and resource ledgers. $\mathbf{1}[\cdot]$ is the dimensionless indicator; $\succeq$ and $\preceq$ mean that every registered vector component passes in its declared direction.

Let

$$N_{\text{acc}} = \sum_{j=1}^{N_{\text{req}}} A_j,$$

where N_{req} is requested outcomes [request] and N_{acc} is accepted outcomes [accepted outcome]. Every energy intensity in this chapter uses that denominator. If $N_{\text{acc}} = 0$, the intensity is undefined and the arm fails; it is not zero.

Carry six typed records

Each run produces six linked records rather than one “energy” field:

RECORD	MINIMUM FIELDS	PRIMARY ARCHITECTURAL OWNERS
fundamental_operation	logical map, physical encoding, p_0 , p_1 , Hamiltonians, T , correlations, error, duration, controls, theorem support	Candidate 009, Candidate 014
device_transition	device identity, terminals, waveform, energy/heat sign, temperature, duration, error, stability, meter/calibration	Candidate 006, Candidate 014
circuit_control	clock, power clock, wires, converters, leakage, sensing, controller, history, correction, reset, I/O	Candidate 010, Candidate 012

RECORD	MINIMUM FIELDS	PRIMARY ARCHITECTURAL OWNERS
workload_hierarchy	software/model version, precision, bytes by level, routes, utilization, idle, retries, quality, latency, throughput	Candidate 001, Candidate 017, Candidate 018
facility_interval	synchronized IT/facility meters, cooling, storage/network share, site, weather, PUE category, allocation sensitivity	Candidate 014
lifecycle_cohort	started and accepted devices, yield, package, deployment, maintenance, utilization, lifetime, replacement, end of life	Candidate 005, Candidate 006, Candidate 018

Every record carries immutable hardware/software/calibration versions, timestamps, uncertainty, validity support, missingness, and the parent/child identity needed to trace replacement or recalibration. This is the physical- energy specialization of the project's **versioned observation contract**.

Make boundary escalation explicit

A result can support only its measured level:

1. theorem or ideal protocol result;
2. isolated device result;
3. closed circuit result;
4. accepted workload result;
5. allocated facility result; or
6. amortized lifecycle result.

Promotion from one level to the next requires a new measurement, not a larger claim. This keeps the chapter aligned with **reliability under mission profiles**, where device population and history are part of the evidence, and with **operator-qualified sensing**, where physical observations remain tied to their operator, calibration, and support.

Efficiency mechanism

One vector, not one number

For a sealed run, report

$$\mathbf{E} = (E^{\text{fund}}, E^{\text{dev}}, E^{\text{circ}}, E^{\text{IT}}, E^{\text{fac}}, E^{\text{emb}}) \quad [\text{J}],$$

where the components are respectively theorem-qualified fundamental lower bound, device-terminal energy, complete circuit/control energy, workload IT energy, allocated facility energy, and allocated embodied energy [J]. The vector is not a sum: in many measurements $E^{\text{dev}} \subset E^{\text{circ}} \subset E^{\text{IT}} \subset E^{\text{fac}}$. For boundary b , useful intensity is

$$e^b = \frac{E^b}{N_{\text{acc}}} \quad [\text{J/accepted outcome}],$$

where $b \in \{\text{dev}, \text{circ}, \text{IT}, \text{fac}, \text{emb}, \text{life}\}$ and N_{acc} is accepted outcomes [accepted outcome]. Distance between E^{fund} and any measured component is descriptive only after their operations and denominators match; it is not a system ranking (C-1151).

Boundary 1 — generalized erasure

For physical microstate $z \in \mathcal{Z}$, probability $p(z)$ [dimensionless], Hamiltonian $\mathcal{H}(z)$ [J], bath temperature T [K], and Boltzmann constant $k_B = 1.380649 \times 10^{-23}$ J/K, define nonequilibrium free energy

$$\mathcal{F}[p, \mathcal{H}] = \sum_{z \in \mathcal{Z}} p(z) \mathcal{H}(z) + k_B T \sum_{z \in \mathcal{Z}} p(z) \ln p(z) \quad [\text{J}].$$

For an isothermal transformation under the selected theorem's assumptions, expected work on the system obeys

$$\langle W_{\text{on}} \rangle \geq \Delta \mathcal{F} = \mathcal{F}[p_1, \mathcal{H}_1] - \mathcal{F}[p_0, \mathcal{H}_0] \quad [\text{J}],$$

where p_0, p_1 are initial and final microstate distributions, $\mathcal{H}_0, \mathcal{H}_1$ are initial and final Hamiltonians [J], and W_{on} is work on the system [J]. Logical entropy alone is insufficient for nondegenerate or nonequilibrium memories (C-1103). Side information, correlation, and a finite reservoir change the accounting (C-1105, C-1107).

The familiar binary special case is

$$E_{\text{reset}}^{\text{fund}}(T, \epsilon) = k_B T [\ln 2 - h(\epsilon)] \quad [\text{J}], \quad h(\epsilon) = -\epsilon \ln \epsilon - (1 - \epsilon) \ln(1 - \epsilon),$$

where $\epsilon \in [0, 1/2]$ is symmetric reset-error probability [error/transition] and $h(\epsilon)$ is binary entropy [nat]. At $\epsilon = 0$, a uniformly distributed degenerate bit yields $k_B T \ln 2$ (C-1101). A biased state instead follows its entropy (C-1102). The system must still charge the consequence of allowed errors (C-1104).

Finite duration adds a separate coordinate. For protocol π of duration τ_π [s], define excess work

$$W_\pi^{\text{ex}} = \langle W_{\text{on}, \pi} \rangle - \Delta \mathcal{F} \quad [\text{J}].$$

Compare it only for matched initial/final state, error, bath, and allowed controls. Finite-time excess, finite error, and stability are distinct (C-1118, C-1119, C-1122). Finite-time quantum erasure has additional model-specific cost and fluctuation structure; classical quasistatic expressions cannot simply be relabeled (C-1120).

Information, fluctuations, and theorem scope

Individual trajectories below a mean bound are compatible with fluctuation relations (C-1106, C-1127, C-1128, C-1129). The implementation therefore stores the full work distribution, sample-selection rule, reverse protocol, rare-event coverage, and estimator uncertainty rather than only a mean or minimum.

For feedback, close the physical loop:

$$E_{\text{feedback}}^{\text{joint}} = E^{\text{plant}} + E^{\text{sense}} + E^{\text{record}} + E^{\text{control}} + E^{\text{actuate}} + E^{\text{reset}} \quad [\text{J}],$$

where the six terms are energy crossing the plant, sensor, record memory, controller, actuator, and reset boundaries [J]. Extracted plant work is signed; it cannot cancel an unmeasured controller.

For a stationary continuous-time Markov jump process and registered integrated current J_t over time t [s], the original steady-state thermodynamic uncertainty relation has the scoped form

$$\frac{\text{Var}(J_t)}{\langle J_t \rangle^2} \Sigma_t \geq 2,$$

where Σ_t is expected entropy production in units of k_B [dimensionless] and J_t is measured in its registered integrated-current unit. The process, current, stationarity, Markov property, time-reversal convention, observation support, and entropy-production estimator must be established first (C-1139). Other finite-time, initial-state, non-Markovian, deterministic, or quantum settings do not inherit this formula unchanged (C-1140, C-1141).

Boundaries 2 and 3 — device, circuit, and real crossover

For device transition k over $[t_k^0, t_k^1]$, terminal energy is

$$E_k^{\text{dev}} = \sum_{c=1}^{C_k} \int_{t_k^0}^{t_k^1} V_{k,c}(t) i_{k,c}(t) dt \quad [\text{J}],$$

where C_k is supplied channels [channel], $V_{k,c}$ is calibrated voltage [V], $i_{k,c}$ is signed current [A], and t is time [s]. Heat requires a calibrated thermal balance; terminal electrical energy is not relabeled as heat.

Logical reversibility can avoid compulsory erasure at intermediate steps (C-1112), but history, ancillae, output copy, communication, uncomputation, retention, and final reset remain physical (C-1113, C-1114). Its strongest comparison is therefore against reversible pebbling, checkpoint/recompute, compiler elimination, and an optimized irreversible circuit at equal service.

For an idealized adiabatic RC path, the crossover model is

$$E^{\text{adiabatic}}(\tau) = \gamma \frac{RC}{\tau} CV^2 + P_{\text{leak}}\tau + E^{\text{clock}}(\tau) + E^{\text{control}} + E^{\text{I/O}} + E^{\text{reset}} \quad [\text{J}],$$

where R is resistance [ohm], C is capacitance [F], τ is transition duration [s], V is voltage [V], γ is a waveform coefficient [dimensionless], P_{leak} is leakage power [W], and the remaining terms are clock, control, I/O, and reset energy [J]. Resistive loss may decrease approximately with RC/τ in its slow-ramp regime (C-1115); leakage and power-clock costs can create a finite optimum (C-1116). A fabricated energy-recovery processor establishes feasibility in its measured range, not zero energy or universal superiority (C-1117).

The crossover is real only when $E^{\text{adiabatic}} < E^{\text{ordinary}}$ at matched quality, error, throughput, capacity, layout/process, temperature, and cyclic closure. Here, E^{ordinary} is complete energy of the conventional comparison [J]. The ordinary CV^2 switching-loss expression is a circuit model, not Landauer erasure (C-1147).

Retention and correction

For memory tier m , charge

$$E_m^{\text{memory}} = N_m^{\text{w}} e_m^{\text{w}} + N_m^{\text{r}} e_m^{\text{r}} + N_m^{\text{ref}} e_m^{\text{ref}} + E_m^{\text{ECC}} + E_m^{\text{scrub}} + E_m^{\text{move}} + E_m^{\text{idle}} \quad [\text{J}],$$

where N_m^{w} , N_m^{r} , and N_m^{ref} are write, read, and refresh counts [operation]; the corresponding e terms are measured energy [J/operation]; and the remaining terms are correction, scrub, movement, and idle energy [J]. Report retention distribution, raw and post-correction errors, miscorrection, silent loss, endurance, and accepted retrievals.

Retention depends on barrier, temperature, time, and loss probability in the activated bistable null (C-1121). Analog storage adds thermal/device noise, finite usable precision, drift, calibration, and conversion (C-1123, C-1124). Noise-assisted or stochastic advantage is plausible only for selected workloads against matched deterministic and pseudorandom nulls (C-1125).

Boundary 4 — workload and locality

For hierarchy link $\ell \in \mathcal{L}$, let B_ℓ be transferred bytes [byte] and $\hat{e}_\ell$ be calibrated energy [J/byte] at the registered process, voltage, precision, distance, rate, and utilization. Movement energy is

$$E^{\text{move}} = \sum_{\ell \in \mathcal{L}} B_\ell \hat{e}_\ell \quad [\text{J}].$$

Measure register, local memory, cache, on-chip network, off-package memory, host, storage, and network separately. Sparse or modular execution also pays indices, routing, load imbalance, synchronization, conversion, cache misses, and idle capacity. Data movement can dominate arithmetic on measured accelerators (C-1145); a hierarchy-aware model's transfer to another system remains plausible until wall-plug validation (C-1146).

Modularity can save locality while losing accessible correlation or adding reset and communication cost (C-1142). A physical process optimized for one input prior can add mismatch dissipation under drift (C-1143), and circuit topology can change thermodynamic cost for the same logical function (C-1144). These are direct constraints on sparse predictive computation and the working architecture.

Boundaries 5 and 6 — facility and lifecycle

For a synchronized facility interval r ,

$$\text{PUE}_r = \frac{E_r^{\text{fac}}}{E_r^{\text{IT}}} \quad [\text{dimensionless}],$$

where E_r^{fac} is total facility energy [J] and E_r^{IT} is IT-equipment energy [J] under the declared ISO/IEC 30134-2 boundary and measurement category. PUE is neither task energy nor carbon intensity (C-1148). Workload attribution still needs synchronized meters, accepted outcomes, network/storage shares, weather, and sensitivity to cooling-overhead allocation.

For hardware cohort h , lifecycle energy is

$$E_h^{\text{life}} = E_h^{\text{fab}} + E_h^{\text{pack}} + E_h^{\text{transport}} + E_h^{\text{deploy}} + E_h^{\text{op}} + E_h^{\text{maint}} + E_h^{\text{replace}} + E_h^{\text{EOL}} \quad [\text{J}],$$

where the terms are fabrication, packaging, transport, deployment, operation including facility share, maintenance, replacement, and end-of-life primary energy [J]. Fabrication and packaging can change a use-phase ranking (C-1149). Specialized hardware is superior only when saved accepted-service energy outweighs new fabrication, low utilization, support life, maintenance, and replacement across uncertainty (C-1150).

Evidence status

The ledger contains exactly **52 claims**:

- **46 established** within their stated theorem, experiment, device, circuit, workload, facility, or lifecycle boundary;
- **5 plausible** transfers that still require target-system evidence; and
- **1 disputed** system-level inference.

The status distribution is not a confidence score for one architecture. It describes separate claims with separate support:

CLAIM BLOCK	STATUS	WHAT IS SUPPORTED
C-1100–C-1124	25 established	encoding dependence, generalized and finite-error erasure, finite reservoirs, four experimental platforms, reversible/adiabatic computation, finite-time cost, retention, switching error, and analog noise/precision
C-1125	1 plausible	selected stochastic/noise-assisted workloads may save energy against matched nulls
C-1126–C-1136	11 established	AWGN communication scope, fluctuation relations, feedback/information engines, continuous information flow, and scoped sensing tradeoffs
C-1137–C-1138	2 plausible	predictive-information and stochastic-learning transfer to deployed AI
C-1139–C-1145	7 established	TUR scope, modularity/mismatch cost, circuit topology, and measured importance of data movement

CLAIM BLOCK	STATUS	WHAT IS SUPPORTED
C-1146	1 plausible	hierarchy-aware energy prediction across routed workloads and systems
C-1147–C-1149	3 established	circuit charging differs from erasure, PUE scope, and fabrication/packaging burden
C-1150	1 plausible	lifecycle superiority of specialized low-operational-energy hardware
C-1151	1 disputed	using distance above $k_B T \ln 2$ as an actionable AI-system ranking

Four platforms experimentally approach or test Landauer-scale erasure: a colloidal memory (C-1108), a feedback trap (C-1109), a nanomagnetic bit (C-1110), and a cryogenic molecular nanomagnet (C-1111). They establish the physical principle in their declared protocols. They do not measure complete computers. The same evidence discipline applies upward:

1. a theorem establishes a bound only for its model;
2. a device experiment establishes its controlled physical transition;
3. a processor or accelerator establishes its measured circuit/workload range;
4. facility metering establishes its interval and allocation; and
5. lifecycle assessment establishes its functional unit and inventory cases.

The unresolved scientific object is not another universal constant. It is the measured crossover between implementations at equal accepted service. Fixture F-010 preserves that question without promoting its evaluation contract into a new P- bundle.

Speculative extensions

Only the five plausible claims license active extension work.

Physical stochasticity for matched workloads

Physical noise could be useful when the task already requires sampling, probabilistic search, or exploration. The comparison must match stationary distribution or target posterior, bias, mixing, tail coverage, latency, task quality, device/circuit energy, and lifecycle cost against high-quality digital pseudorandom sampling (C-1125). A noisy device does not earn credit merely for producing variation.

Predictive retention rather than historical retention

If a continual system stores only state that improves prediction, it may avoid updates whose information is nonpredictive. The surviving question is physical: does the predictive objective reduce writes, movement, correction, and retained capacity after ordinary predictive compression, caching, event-triggered updates, and recomputation are matched (C-1137)? This joins memory and consolidation with Candidate 017 and Candidate 018.

Thermodynamic learning efficiency on actual hardware

Toy stochastic-learning results can generate hypotheses about which updates carry useful information, but they do not bind digital gradient training (C-1138). A valid transfer would jointly measure optimizer, arithmetic, activation/gradient memory, communication, data loading, checkpointing, accepted validation outcomes, and hardware/facility energy under the same learning contract.

Hierarchy-aware conditional execution

A route-energy model could decide whether skipping arithmetic saves more than its indices, movement, arbitration, synchronization, imbalance, and idle capacity cost (C-1146). It must transfer across held-out model, sequence/graph length, sparsity pattern, cache-fit, precision, topology, and utilization regimes. This is the energy test for Candidate 001, not a new routing candidate.

Lifecycle-qualified physical compilation

Specialized hardware may move recurring computation into a lower-energy substrate, but the gain becomes real only after yield, package, converters, calibration, utilization, service life, software support, maintenance, replacement, and displaced-hardware assumptions are propagated (C-1150). Candidate 006 already owns that experiment. The governing quantity is accepted lifetime service, not peak component efficiency.

Failure modes

FAILURE	WHY THE CLAIM FAILS	REQUIRED REPAIR
assign $k_B T \ln 2$ to every operation	Landauer attaches to a declared physical information reduction, not an operation label	state the physical encoding, distribution, Hamiltonian, bath, error, duration, and cycle
treat a biased or known bit as uniform	information erased depends on the prior and usable side information	measure p_0 , correlations, and preparation/reset of helpers
call Landauer a power bound	work [J] lacks protocol time and throughput	report duration [s], useful rate [outcome/s], and capacity [device s]
claim a violation from one low-work trajectory	fluctuation relations constrain ensembles	preserve the full distribution, reverse protocol, rare-event support, and estimator uncertainty
reduce work by allowing errors	the transformation and accepted service changed	charge detection, correction, retry, fallback, silent loss, and harm
equate logical with physical reversibility	an invertible map says nothing about dissipative dynamics	close ancillae, history, output copy, uncomputation, clock, leakage, I/O, and reset
call adiabatic switching lossless	slower resistive loss can reveal leakage and power-clock cost	measure a real throughput-matched crossover across frequency, load, temperature, and utilization
omit memory stability	low write energy is useless if state expires or refresh dominates	measure retention distribution, refresh, correction, endurance, and accepted retrievals
treat analog state as an exact real number	useful precision depends on signal, noise, bandwidth, drift, calibration, and conversion	match end-to-end precision and tail quality with converters and host included
reuse $E_b/N_0 \geq \ln 2$ as a gate bound	it is an AWGN reliable-communication asymptote under a rate/coding regime (C-1126)	include transmitter, receiver, bandwidth, code, latency, and error in the communication service
draw an information engine around the plant	sensing, record memory, controller, actuation, and reset were exported	use the joint feedback ledger
apply a TUR to arbitrary AI metrics	training loss or accuracy is not automatically a physical Markov current	prove process, observable, stationarity, reversal, and entropy-production support first
assume modules always save energy	boundaries can discard correlations and add communication/reset	compare joint, modular, and shared-sufficient-statistic implementations under shifted priors
quote component picojoules as constants	process, voltage, precision, hierarchy, distance, rate, and utilization differ	calibrate the energy model to the measured implementation and top-level meter
report skipped arithmetic as task savings	routing, metadata, movement, imbalance, synchronization, and idle capacity may dominate	report bytes and joules at every hierarchy level per accepted outcome
multiply by generic PUE	PUE is interval- and facility-bound and lacks a task denominator	synchronize IT/facility meters and test registered overhead allocations
report operational energy as lifecycle efficiency	fabrication, yield, utilization, lifetime, maintenance, and replacement may reverse the ranking	use one cradle-to-retirement functional unit and uncertainty cases
compare unlike useful tasks	lower quality, longer latency, narrower support, or more failures created the saving	apply the common outcome firewall before comparing energy

The broader energy model should consume these typed records, while reliability under mission profiles supplies the device population, temperature, wear, correction, repair, and retirement state. Neither chapter can replace the other’s denominator.

Measurable predictions

Fixture F-010 converts the chapter into twelve equal-budget experiments:

TRACK	PREDICTION THAT MAY SURVIVE	STRONG NULL	HARD RETIREMENT CONDITION
T1 generalized erasure	a proposed protocol lowers the matched work distribution at fixed initial/final state, error, duration, bath, and controller boundary	best full/restricted-control protocol plus slow reference on the same memory	advantage vanishes when error, duration, correlation, finite reservoir, or controller is matched
T2 reversible kernel	closed reversible execution lowers circuit/workload joules for useful bijective, many-to-one, and iterative kernels	optimized irreversible, checkpoint/recompute, compiler elimination, reversible pebbling variants	history, ancillae, output copy, retained state, throughput replication, or final reset is external
T3 adiabatic crossover	a measured operating region beats conventional CMOS at equal service	ordinary, clock/power-gated, DVFS, and near-threshold circuits at matched process/layout	power clock, leakage, interconnect, capacity, or error removes the crossover

TRACK	PREDICTION THAT MAY SURVIVE	STRONG NULL	HARD RETIREMENT CONDITION
T4 retention frontier	a memory tier reduces lifetime retrieval energy at required retention and error	SRAM/DRAM/nonvolatile, recomputation, and tiering appropriate to the horizon	refresh, ECC, silent loss, endurance, reserve, or replacement erases the gain
T5 analog closure	an analog/physical path lowers wall-plug accepted-task energy at required precision	digital mixed precision and matched low-precision/stochastic paths	conversion, calibration, host, drift, shift, or tail-quality cost removes the gain
T6 information engine	net joint work remains favorable with the whole feedback loop inside the boundary	open loop, predictive control, and randomized action at matched sensing/actuation	gain exists only around the plant or depends on uncharged records/control
T7 TUR scope	a registered physical current satisfies an applicable precision–dissipation bound and constrains task-relevant behavior	finite-time/transient variants, hidden-state/non-Markov models, predictive empirical null	process assumptions fail, entropy production is unidentifiable, or task relevance is absent
T8 modularity/mismatch	modules save energy after cross-boundary correlation, traffic, reset, calibration, and prior shift	joint implementation and module system with shared sufficient statistics	advantage assumes the deployment prior or disappears under correlation/shift
T9 locality	sparse/conditional execution lowers IT energy after all movement and capacity terms	dense optimized, structured sparsity, compiler tiling/cache/data reuse	saved arithmetic is offset by bytes, metadata, sync, imbalance, conversion, or idle
T10 facility	workload improvement reduces synchronized facility energy per accepted outcome	matched randomized facility blocks or calibrated side-by-side system	result comes from TDP, generic PUE, short interval, or unstable allocation
T11 lifecycle	operational savings repay incremental embodied burden within supported service life	deployed general hardware, software optimization, and shared specialized service	break-even requires unsupported yield, utilization, demand, lifetime, or displaced-hardware credit
T12 full stack	one candidate-backed composition Pareto-improves quality, latency, risk, capacity, energy, and lifecycle under uncertainty	complete ordinary stack plus boundary and mechanism ablations	no Pareto gain survives held-out regimes, support gates, and required sensitivities

All tracks use the same accepted-task contract and preserve failed devices, rejected requests, timeouts, retries, uncorrectable errors, silent corruption, abstention, idle capacity, maintenance, and replacement in their denominators. Confirmation groups withhold physical devices, fabrication cohorts, waveform and duration regimes, target errors, temperatures, retention horizons, workload families, hierarchy patterns, controller versions, sites, seasons, lifecycle cases, and future time.

The common result vector is

$$\mathbf{Y} = (f_{\text{acc}}, Q, L_{0.50}, L_{0.99}, \rho, e^{\text{IT}}, e^{\text{fac}}, e^{\text{life}}, E^{\text{err}}, C^{\text{cap}}, G, W, M),$$

where f_{acc} is accepted fraction [dimensionless], Q is task quality [task-native unit], $L_{0.50}$ and $L_{0.99}$ are median and 99th percentile latency [s], ρ is risk [failure/request], the three e terms are energy intensity [J/accepted outcome], E^{err} is error-consequence energy [J], C^{cap} is provisioned capacity [device s], G is greenhouse-gas inventory [kg CO₂e], W is water inventory [m³], and M is a material/labor vector in declared native units.

A physical-efficiency claim survives only when its simultaneous uncertainty region is no worse on every hard-gated coordinate and strictly better on at least one preregistered primary coordinate against the strongest compatible null across required sensitivity cases. Passing supports only the existing candidate scope named in F-010. Failure identifies the boundary that produced the apparent saving and retires the wider claim.

Sparse prediction and adaptive compute

Scope

This chapter specifies the online control path that decides, for each event, whether to reuse an existing prediction, exit, invoke another module, read memory, acquire another observation, or escalate. Its objective is not minimum FLOPs. It is minimum measured resource use subject to explicit quality, calibration, deadline, and safety constraints.

The runtime has three coupled but distinct jobs:

1. estimate what changed and what remains uncertain;
2. choose the next computation or observation by expected decision value; and
3. allocate physical capacity without making the fast task path unstable.

Training curricula, long-term consolidation, structural pruning, and skill compilation are handled in other chapters. Their products may be invoked here, but this runtime does not silently rewrite slow model state.

Biological observation

The relevant evidence separates demand detection, distributed context, local response, and physical supply.

Activity is selective and regulated

Neural signaling operates under a strict metabolic budget, which constrains the fraction of cortex that can be strongly active at once (C-001). In the *Drosophila* mushroom body, feedback inhibition keeps odor codes sparse and decorrelated; disrupting that loop impairs discrimination between similar learned odors (C-025). These results support selective activity and active inhibition. They do not identify top- k routing, a pruning ratio, or a digital energy saving.

Prediction, uncertainty, and action are different signals

Hierarchical prediction with feed-forward residuals accounts for selected visual-cortical response properties (C-005). Active whisker sensing also shows why a passive residual is incomplete: self-generated sensor mechanics affect the observed neural code (C-022). A residual says that a prediction was wrong. It does not by itself say whether the error is irreducible noise, model ignorance, task-relevant novelty, or a reason to buy another observation.

Broad context is interpreted locally and temporally

Context-dependent disinhibition can change gain in a selected cortical population (C-020). More direct causal work on neuromodulation shows three constraints on a possible artificial abstraction:

- antagonistic signals can sustain different distributed behavioral modes (C-046);
- the frequency and duration of one broad signal can produce different and non-monotonic system responses (C-047); and

- receptor identity and location can turn the same source into transient or sustained local responses (C-048).

The engineering candidate is therefore a rate-limited context broadcast whose receivers own their gains, thresholds, and temporal filters. It is not a central command containing a route for every module. Its evidence, limits, and conventional analogues are developed in the neurodevelopment and global-control audit.

Demand and physical supply can be locally coupled

Presynaptic ATP production can be recruited by local activity (C-049); mitochondrial docking changes resource placement and local synaptic dynamics (C-050); and neural activity can recruit an adjacent vascular response in the studied rodent preparations (C-051). The transferable constraint is modest: measure demand near work, account for placement, and let an infrastructure layer adjust supply. The neurovascular paper's model-derived percentage is not an architecture constant, and none of these cellular results provides a conversion from ATP to accelerator joules.

Aggregate headroom is likewise not evidence of delivery or useful uptake at the support where demand occurs; local deficit, arrival, consumption, and export remain distinct records (C-1488).

Taken together, the observations motivate a layered controller: fast local prediction and gating, a narrow context path, and a slower resource plane. The architecture still has to beat standard estimation, decision, routing, and control methods.

Proposed AI translation

Separate the quantities before making a gate

Let y_t be an observation in its declared sensor units and $\hat{y}_t|_{t-1}$ the prediction available before observing it. Define the raw residual

$$r_t = y_t - \hat{y}_t|_{t-1}.$$

If the predictor supplies a calibrated residual covariance S_t , the normalized innovation is

$$\nu_t = r_t^T S_t^{-1} r_t.$$

r_t has the units of y_t ; ν_t is dimensionless. A large r_t with a large expected S_t may be unsurprising. Conversely, a small residual in a safety-critical variable may justify action. The runtime therefore carries separate fields for:

QUANTITY	MEANING	UNIT OR REPRESENTATION
residual r_t	observed minus predicted state	observation units
normalized innovation ν_t	residual relative to predicted covariance	dimensionless
aleatoric uncertainty	expected irreducible variability	distribution in output units
epistemic uncertainty	model uncertainty or support deficit	calibrated probability or ensemble statistic
task risk	cost of acting or exiting incorrectly	probability and consequence in task units
expected acquisition value $\Delta U_t(a)$	expected decision improvement from action a	declared utility units
estimated energy $\hat{E}_t(a)$	energy for acquisition and downstream work	joules
estimated latency $\hat{L}_t(a)$	completion time including queueing	seconds
traffic $\hat{B}_t(a)$	memory or network movement	bytes
context rate Rb	transmitted global operating context	bits per second

No implementation may collapse all of these into “confidence” and retain the same claim. Residual magnitude, likelihood, expected value, and risk answer different questions. Predictive regulation adds the same separation between forecast value, feedback correction, integrated action exposure, reserve debit, and later outcome (C-1494).

Relative sensing is a symmetry, not a normalization label

The same separation applies when a gate should respond to change relative to a background. For a positive input $u(t)$ and reference $r(t)$, a log ratio

$$z(t) = \ln \frac{u(t)}{r(t)}$$

is invariant to the common transformation $(u, r) \mapsto (pu, pr)$. That algebra does not establish how r is acquired, how long it remains valid, or whether the downstream target is itself scale-invariant. Exact fold-change detection is stronger: the complete output trajectory must remain the same after the input history and its adapted initial state are scaled together (C-1540). Exact adaptation or equal peaks cannot substitute for that test.

Biological evidence shows why the interface must be recorded. Bacterial chemotaxis connects finite-regime intracellular rescaling to active migration (C-1542); nuclear RelA fold change predicts three downstream transcripts in one cell system (C-1544); yet dynamic extracellular cytokine protocols support a rectified dose-change statistic at the upstream cytokine-to-NF- κ B interface (C-1545). The pathway name therefore cannot select the statistic.

The runtime carries a typed candidate-statistic record

$$\mathcal{V}_t^{(\text{rel})} = [u_t, u_t - r_t, \dot{u}_t, \ln(u_t/r_t), \tau_t^{(\text{ref})}, \chi_t^{(\text{sup})}] ,$$

where $\tau_t^{(\text{ref})}$ is reference age in seconds and $\chi_t^{(\text{sup})}$ is a typed support/validity state. A selector may route a ratio-qualified task through the relative channel, but the record's fields retain their individual units and types; it is not an ordinary numeric vector. An absolute load, dose, damage, or safety target retains a calibrated absolute path. Near zero, after sign changes, under clipping, or with stale reference state, the log-ratio path must abstain or fall back.

This design must beat explicit log-ratio, difference, future-free streaming normalization, state-space and compact recurrent nulls while charging reference updates, selector execution, fallback and state writes. The exact model, interface firewall, scale-orbit plot, and trajectory discrepancy are derived in [Interface-qualified scale symmetry](#); the reciprocal ten-track test is Fixture F-026.

Command, interface work, transport, and observation are different stages

An adaptive gate often mistakes its command for the work that actually crossed an interface. Electrochemistry supplies a sharp counterexample. In one common charge-transfer form,

$$j = j_0 \left[\exp\left(\frac{\alpha_a F \eta}{R_g T}\right) - \exp\left(-\frac{\alpha_c F \eta}{R_g T}\right) \right] ,$$

where j and j_0 are current densities in amperes per square metre, η is overpotential in volts, α_a and α_c are dimensionless transfer coefficients, F is the Faraday constant in coulombs per mole, R_g is the molar gas constant in joules per mole-kelvin, and T is absolute temperature in kelvin. Even when this local closure is correct, bulk transport, double-layer state, and series loss still separate commanded voltage from realized flux and terminal observation (C-1530).

The corresponding artificial contract records a five-stage chain:

1. the controller issues a command under declared authority;
2. receiver-local state converts the command into a driving force;
3. an interface transfers bounded useful work;

4. internal transport accepts, delays, or saturates that transfer; and
5. an observation operator reports a noisy, delayed projection of the result.

A candidate may compress this chain only after an intervention shows that the omitted stage is conditionally redundant. Otherwise, a low command error can coexist with wrong realized service. A validity check may reject an observation before interpretation, but passing such a check does not identify a unique mechanism (C-1532). Likewise, a low prediction residual cannot give physical semantics to unidentifiable latent parameters (C-1537).

Observation is affected by action

The acquisition model must also represent how sensing and response change the future observation channel. Operational telemetry can be delayed, incomplete, pooled, or policy-coupled: robust residual methods remain strong nulls (C-124); current counts can omit events that occurred but have not yet arrived (C-126); pooled signals can trade attribution for coverage (C-127); and a fast behavioral proxy can drift with platform, attention, or policy (C-128).

More importantly, an action can alter both hidden state and telemetry. Isolation may reduce propagation while also removing the isolated component's messages; throttling can suppress an error count by suppressing all traffic; a warning can change user behavior; and a new sensor policy changes which events can be observed. The controller therefore versions coverage, delay, ascertainment, observation-model state, and every sensing or response action. A drop in the post-action residual is not independent recovery evidence (C-131).

Candidate 007 tests joint process/observation estimation against residual CUSUM/GLR, nowcasting, maximum coverage, value-of-information sampling, and delay-aware POMDP or model-predictive control. It must use true data vintages, preserve subgroup coverage, estimate counterfactual observation paths, and charge investigation and action capacity as well as compute.

Price a menu of acquisitions

At event t , the system considers a finite action menu $\mathcal{A}_t$:

```
exit | reuse cached state | run local refinement | invoke expert
    | read episodic/factual memory | acquire sensor action | call tool
    | escalate to a larger model or human
```

For acquisition a , estimate the expected improvement in decision utility $\Delta U_t(a)$ and its full energy, latency, and traffic costs. A Lagrangian form of the selection rule is

$$a_t^* = \arg \max_{a \in \mathcal{A}_t} \left[\Delta U_t(a) - \lambda_E \hat{E}_t(a) - \lambda_L \hat{L}_t(a) - \lambda_B \hat{B}_t(a) \right],$$

subject to hard task constraints such as

$$\Pr(\text{unsafe outcome} \mid a, \mathcal{I}_t) \leq \delta, \quad \hat{L}_t(a) \leq L_{\max}.$$

$\mathcal{I}_t$ is the information legally available when event t is decided, $\delta \in [0, 1]$ is the preregistered maximum conditional probability of an unsafe outcome, and $L_{\max}$ is the latency ceiling in seconds. λ_E , λ_L , and λ_B have units utility/joule, utility/second, and utility/byte respectively, converting the three costs into the declared utility scale. When those conversions cannot be defended, use constrained multi-objective selection and report a Pareto surface instead of inventing a scalar score. An action with high information gain but no expected decision benefit receives no credit.

The next figure uses a hypothetical action ledger to expose that rule. Point positions, costs, probabilities, and the highlighted action are analytical illustrations only; they are not measured values or a recommended controller.

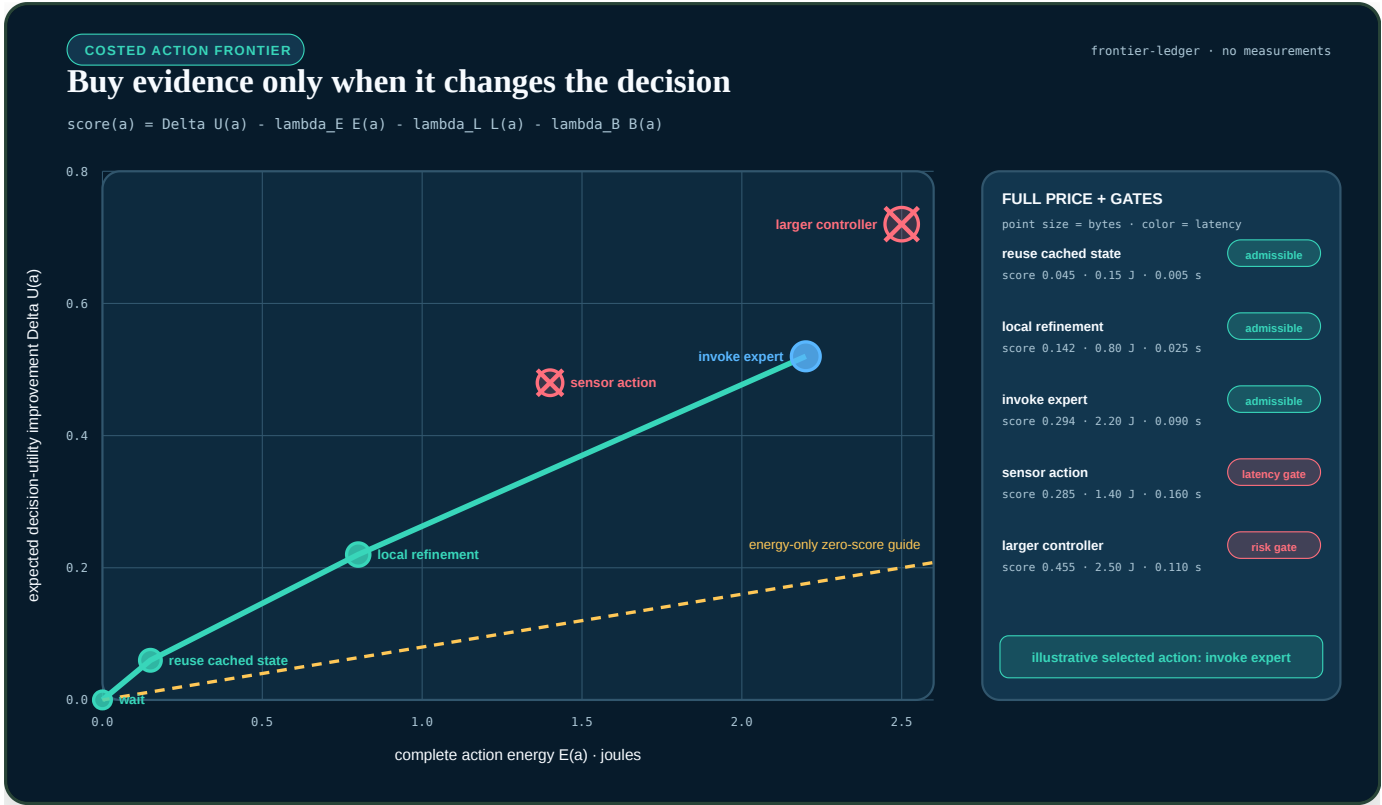


FIGURE 29 — Surprise or information gain alone does not buy another computation or measurement: an action first survives hard risk and latency gates, then competes on decision-value gain against its complete energy, delay, and traffic cost.

Three runtime timescales

LOOP	TYPICAL CADENCE	STATE OWNED	PERMITTED ACTIONS
event loop	per sensor event or token	prediction, residual, uncertainty, risk, route	no-op, exit, acquire, route, escalate
context loop	one or more explicitly rate-limited cadences	small broadcast plus receiver-local filtered state	change local gain, threshold, mode, or compute eligibility
resource loop	slower than task execution and bounded by actuation delay	queues, utilization, temperature, energy, bandwidth prices	adjust local capacity, apply backpressure, reserve or release routes

The event loop cannot raise its own power or bandwidth allocation. It requests capacity through the resource plane. The resource plane cannot change task semantics or fabricate confidence; it changes supply and prices. The context loop cannot transmit a per-module command table under the name “broadcast.” Equal mean state does not collapse delayed local regulation, anticipatory action, and slow structural change into one loop (C-1492, C-1494, C-1496).

A timescale label is not a closure certificate

Calling a loop *fast*, *slow*, or *coarse* describes its schedule. It does not show that the state exposed to that loop is sufficient to predict its own future. Eliminating an unobserved variable can leave history, an unresolved initial-condition term, or lift-dependent evolution behind. The runtime may use a reduced path only after one of the following four contracts is made literal; otherwise it retains or restores the fuller state.

PROJECTION CAN MOVE OMITTED STATE INTO MEMORY

For the dimensioned linear example

$$\dot{x} = -\alpha x + \beta y, \quad \dot{y} = \gamma x - \lambda y,$$

x and y share unit U , while $\alpha, \beta, \gamma, \lambda$ have unit s^{-1} . Eliminating y does not generally give a history-free equation:

$$\dot{x}(t) = -\alpha x(t) + \beta e^{-\lambda t} y(0) + \int_0^t \beta \gamma e^{-\lambda(t-s)} x(s) ds.$$

The first term is instantaneous, the second retains unresolved initial state, and the third is memory. The kernel $K(\tau) = \beta \gamma e^{-\lambda \tau}$ has unit s^{-2} , so its integral has unit $U s^{-1}$ like $\dot{x}$. The general projection result also contains an orthogonal-dynamics term; calling it “noise” does not make it independent, Gaussian, or negligible (C-1526). A finite history window H seconds therefore needs a measured tail-error boundary, not a convenient buffer size.

For the single analytical kernel $K(\tau) = K_0 e^{-\tau/\tau_m}$, with memory time τ_m seconds, the normalized mass omitted after retaining H seconds is $R(H) = e^{-H/\tau_m}$. The figure is an exact illustration of that one kernel; it is not evidence that an artificial workload has exponential or finite memory.

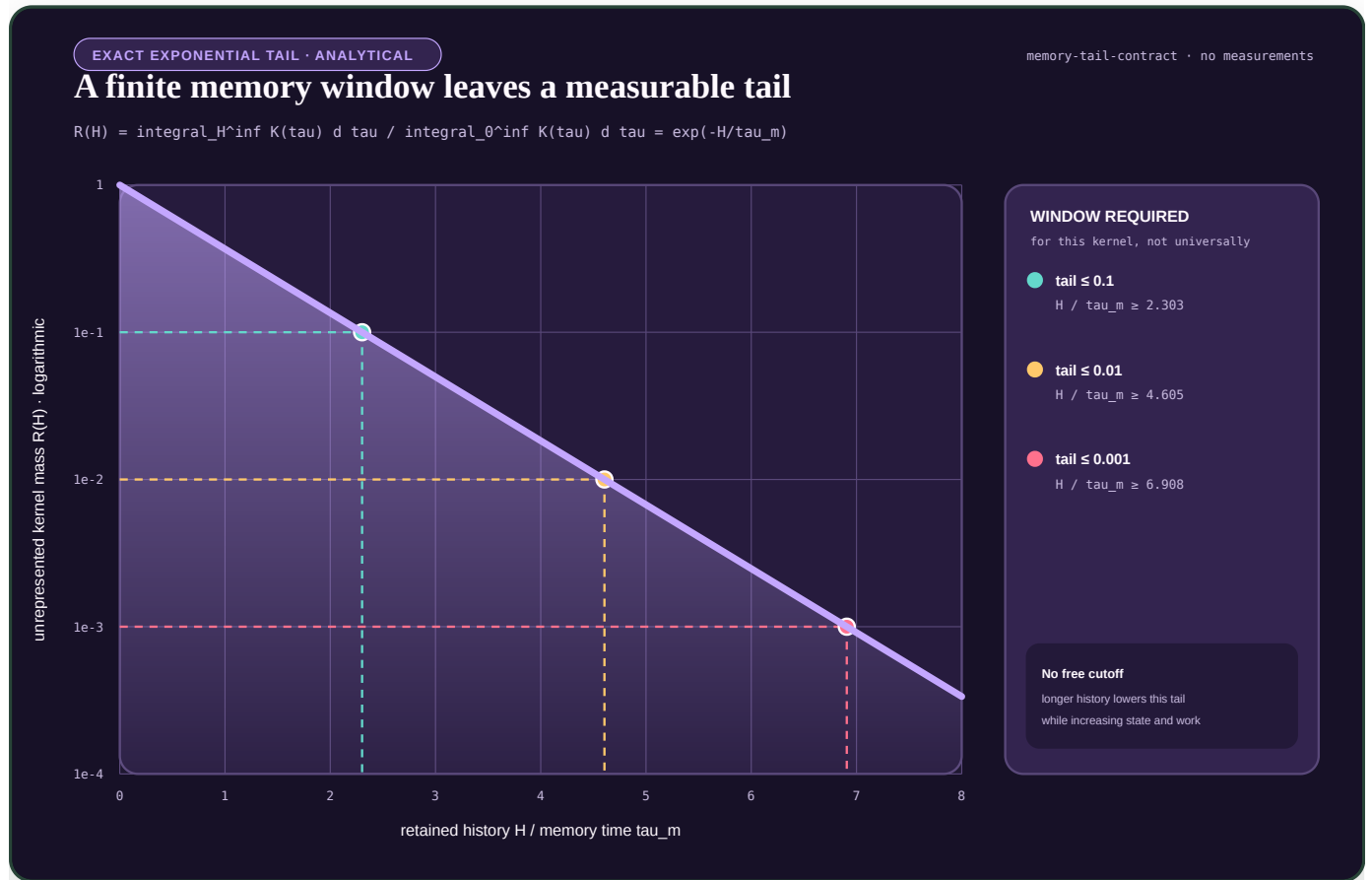


FIGURE 30 — Analytical illustration of the exact normalized tail left by truncating one exponential memory kernel; the horizontal coordinate H divided by τ_m and the tail mass are dimensionless, and no workload measurements are shown.

A SLOW MANIFOLD HAS A GEOMETRIC BOUNDARY

A named fast/slow split also does not establish a slow manifold. In the dimensionless fold normal form

$$\varepsilon x' = y - x^2, \quad y' = -1,$$

prime denotes differentiation with respect to $\theta = t/\tau_0$, where the declared reference time τ_0 is measured in seconds, and ε is the dimensionless fast/slow timescale ratio. On the attracting critical branch $x^*(y) = \sqrt{y}$, the dimensionless normal spectral margin is $\gamma N(y) = 2\sqrt{y}$. It tends to zero as $y \rightarrow 0^+$, while the sensitivity $|dx^*/dy| = 1/(2\sqrt{y})$ diverges. Ordinary

normal-hyperbolic persistence is therefore qualified only on a declared compact region whose margin stays away from zero; it does not continue through the fold by naming the route “slow” (C-1527).

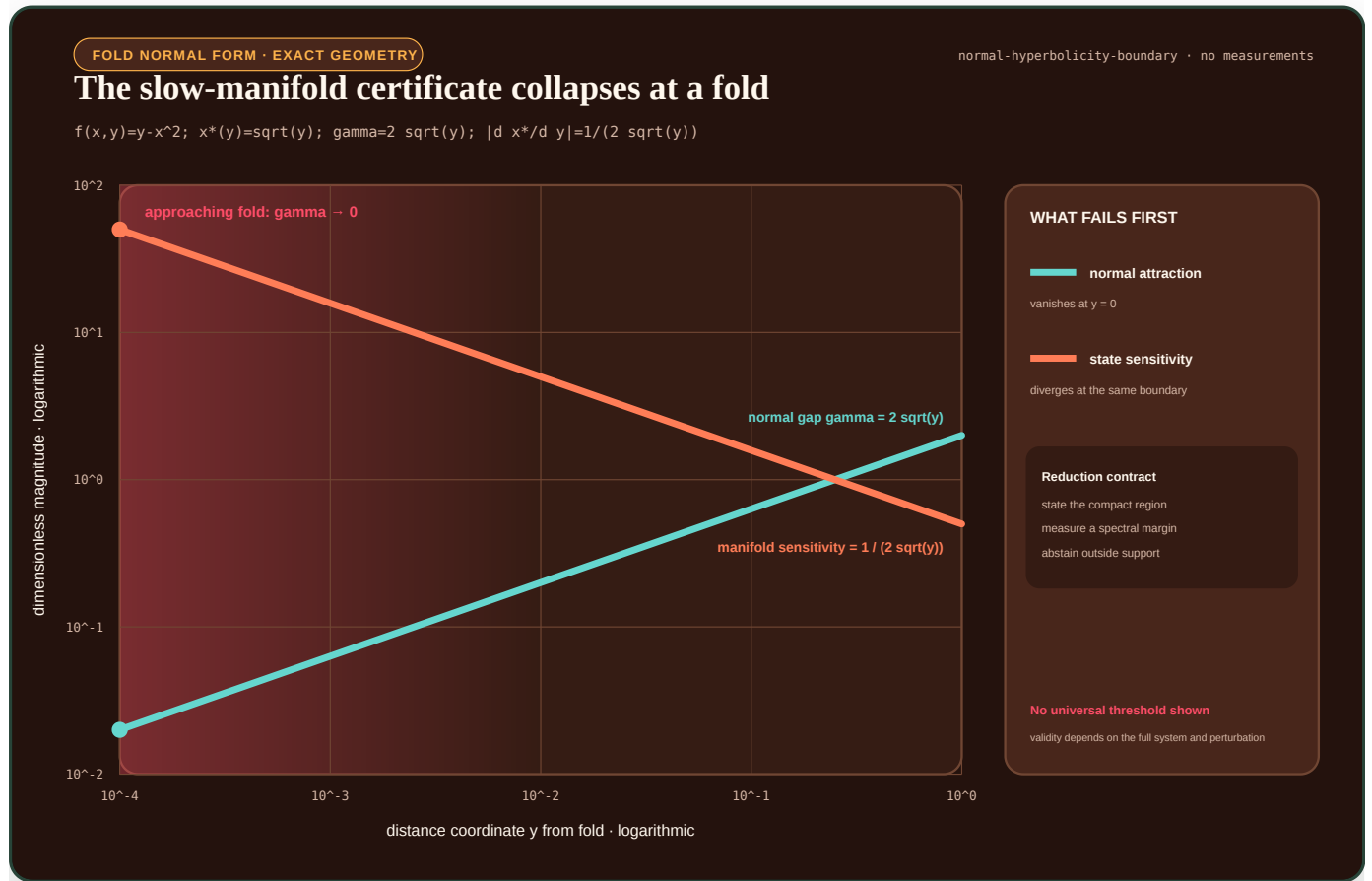


FIGURE 31 — Analytical illustration of the dimensionless fold normal form: normal attraction tends to zero and slow-branch sensitivity diverges as the dimensionless distance coordinate y approaches the fold; this is exact geometry for the displayed normal form, not a measured or universal online threshold.

COARSE COMPUTATION MUST EXPOSE HOW DETAIL RETURNS

Two different execution contracts cover cases in which a coarse state remains useful without pretending that fine detail vanished:

1. **Local micro-query contract.** A heterogeneous multiscale method declares a compression $Q_c : u \mapsto U$, a reconstruction $R(U, \xi) \mapsto u$, and the consistency residual $\|Q_c R(U, \xi) - U\|$. Here u and U retain their native fine- and coarse-state units and ξ indexes admissible unresolved detail. When the macro update needs an unavailable datum such as a flux, it runs a bounded local microproblem to estimate that datum. Microcell support, boundary treatment, relaxation, sampling, failed solves, coefficient queries, iterations, bytes, and reconstruction error all remain in the cost and uncertainty ledger (C-1528). “Local” is not synonymous with “cheaper.”
2. **Lift--heal--evolve--restrict contract.** With fine propagator Φ_T^f over burst time T seconds, the lift-specific coarse map is $\Phi_T^c(U; \xi) = Q_c \Phi_T^f(R(U, \xi))$. Several admissible lifts with the same U are evolved through a declared healing time t_h seconds before restriction. Post-healing disagreement is measured in the native norm of U . If materially different lifts still give different coarse derivatives or rollouts, the proposed coarse variables are not closed at that state and horizon; the route must abstain, add state, or fall back to fine evolution while charging all attempted work (C-1529).

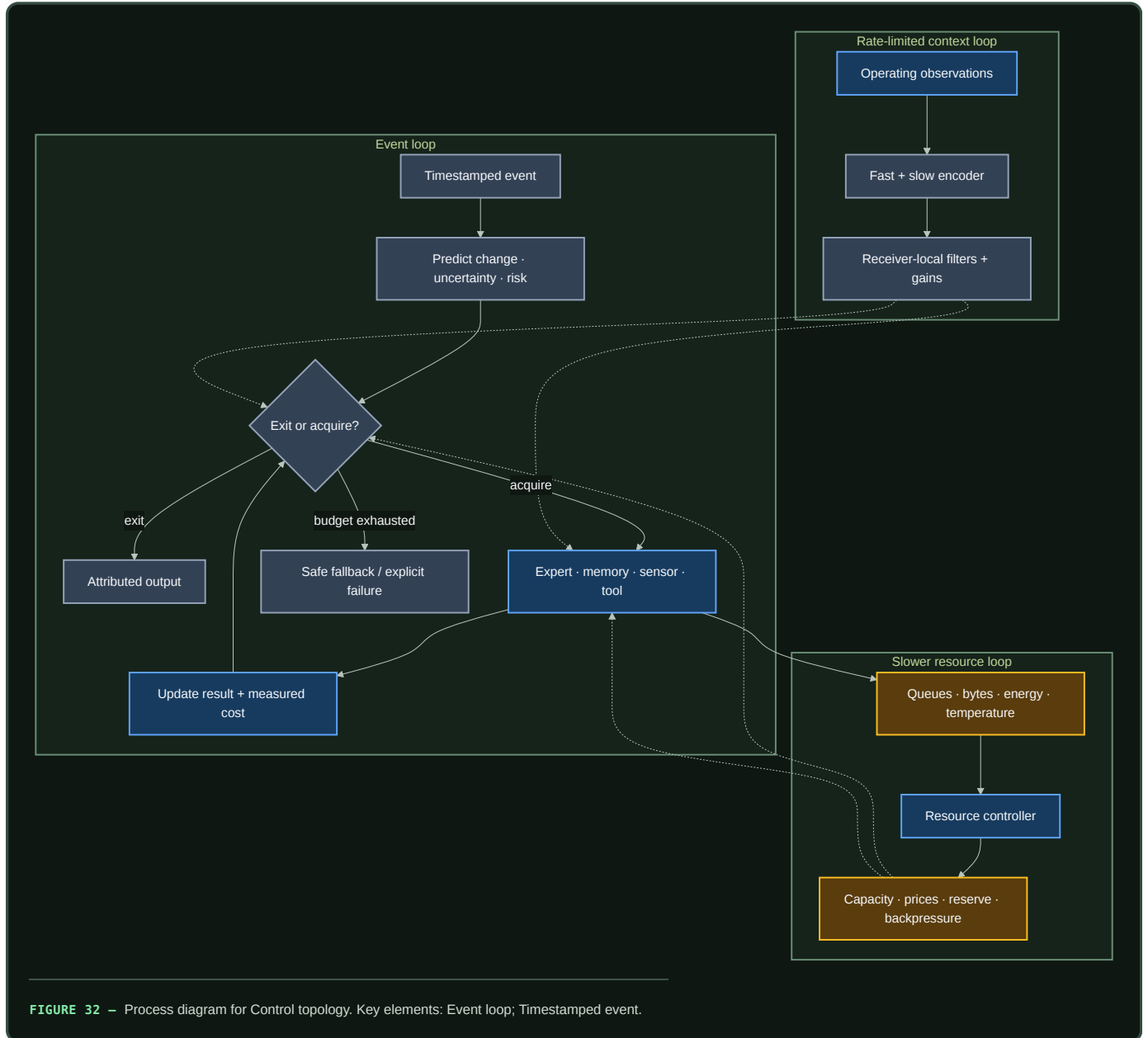
The complete equations, dimensional checks, and non-overlapping compute, traffic, and microstep ledgers are in the multiscale reduction contract. Fixture F-024 registers matched-information tests against Markov, full-state, analytic homogenization, identified coarse-state, and

continuously fine nulls. `NO_RESULT`: the fixture is pre-implementation; neither plot contains measurements, and this section establishes no accuracy, compute, traffic, energy, or readiness result.

Concrete per-event sequence

1. **Timestamp and classify the event.** Record source, modality, freshness, deadline, and risk class. Missing timestamps make prediction error and queue latency ambiguous.
2. **Run a local change test.** Compare the event with the last accepted local state using a sensor-noise model. Below-threshold change creates a no-op or cached-prediction candidate, not an automatic exit.
3. **Predict before incorporating the event.** Produce $\hat{y}_t|_{t-1}$, residual covariance S_t , and the task outputs available at the cheapest depth.
4. **Decompose mismatch.** Compute r_t and ν_t ; update a separate sequential change statistic for persistent drift; estimate epistemic uncertainty and task consequence.
5. **Test the cheapest valid exit.** Exit only if calibration for the event's risk stratum satisfies its error bound and no mandatory provenance, freshness, or tool check remains.
6. **Update the context receivers.** If a fast or slow broadcast update is due, each subscribed module applies its own causal filter, bounded gain, and local-state-dependent threshold. The received bitstream and local response are logged.
7. **Value the remaining acquisitions.** Estimate $\Delta U_t(a)$ and full costs for the available experts, memories, sensors, tools, and escalation paths. Reject actions whose expected value does not exceed their priced cost or whose completion misses the deadline.
8. **Route under capacity constraints.** Select the smallest useful set of modules. Include queue delay, expert load, communication, and state placement. Open reserve routes before tail latency collapses only if the measured congestion rule warrants it; the ant result in C-035 is a lead, not the routing algorithm.
9. **Execute with local backpressure.** A module consumes its local compute, memory, and bandwidth budget. Exhaustion triggers deferral, a cheaper approximation, or escalation; it never silently drops a safety check.
10. **Reconcile prediction and cost.** Update output, uncertainty, and the measured resource ledger. If the risk bound is unmet and deadline and budget remain, return to acquisition valuation. The loop has a fixed maximum iteration count.
11. **Emit an attributable result.** Return prediction or action together with confidence calibration version, routes used, evidence provenance, elapsed time, and measured or allocated energy. Runtime telemetry enters later maintenance analysis; it does not directly rewrite slow weights.

Control topology



Authority follows information and remaining capability

Power-grid protection supplies a hard engineering version of local reflex and global escalation. Local digital relays can act quickly inside a declared zone (C-186), while wide-area schemes add broader but delayed and failure-prone evidence (C-190). Fast response is still constrained by current, headroom, energy, duration, and interacting controls (C-196). The regional/global physiology boundary sharpens the same rule: a useful local correction must still expose shared resistance, exported load, trigger prevalence, and fallback cost (C-1493).

The held translation assigns controller i an admissible action set

$$\mathcal{U}_i(t) = \mathcal{E}_i(\tau_i(t), q_i(t), m_i(t), b_i(t), c_i(t)), \quad u_i(t) \in \mathcal{U}_i(t),$$

where τ_i is observation age in seconds, q_i is integrity state, m_i is operating-mode state, b_i is remaining physical or compute headroom, c_i is coordination availability, $\mathcal{E}_i$ is a certified set-valued map, and u_i is the proposed action. The vector components retain their own units, timestamps, uncertainty, and provenance; the notation does not make them interchangeable.

Stale evidence, lost integrity, exhausted reserve, or lost coordination should shrink authority toward an independently enforced fallback. A wider action requires validated handoff and a checked postcondition. The candidate loses if adaptive protection, gain scheduling, constrained control, barrier functions, or runtime assurance reproduce the same frontier. [Candidate 012](#) tests this under saturation, hidden failure, attack, communication loss, partition, second events, and staged recovery.

Context is a constrained interface

The context candidate sends a small quantized code at declared fast and slow rates. Receivers expand it locally through distinct filters and bounded transfer functions. This is useful only when receiver-local structure reconstructs module-specific responses more cheaply than transmitting them.

[Candidate 002](#) tests that proposition against instantaneous and shared-filter FiLM, memoryless and recurrent gates, bottleneck and standard global tokens, low-rank hypernetworks, fixed multirate receiver banks, and gain-scheduled supervisory control. Its primary channel is two 8-bit components at different cadences, with an average logical rate of 825 bit s^{-1} . That number is an experiment setting, not a proposed universal bus rate.

A bounded interface still needs a back-action test

For the same upstream input, initial state and random seed, characterize the producer first without a consumer and then after attaching one. Keep producer trajectory distortion separate from delivery latency, consumer fidelity, shared-resource contention and any intended feedback. A connection is not harmful merely because it consumes work, and a producer is not insulated when its consumer receives no usable service.

The biological evidence supplies several scoped warning cases: downstream binding pools can alter transcriptional transients, signalling targets can change the operating regime or bandwidth of an upstream cycle, and fast intermediate processes can attenuate some load effects (C-1550–C-1555). The systems translation starts with stronger ordinary controls:

1. an ideal immutable-copy path with no finite-interface backpressure;
2. a tuned bounded ring buffer and explicit backpressure;
3. asynchronous snapshots, copy-on-write and expiry;
4. load-aware admission and resource reservation;
5. unrelated CPU or memory contention without a producer connection; and
6. explicit filtering or intended feedback where temporal shaping is useful.

Only connection-specific distortion that survives those controls is called a retroactivity analogue. Fixture F-027 and its [mathematical contract](#) freeze that comparison. The first executable slice is a public diagnostic and has no performance or energy authority.

Deficit travels; feasibility remains local

Plant nitrogen acquisition supplies a precise three-part allocation loop: nitrogen-starved roots send CEP deficit signals upward, shoot-derived CEPD signals return global context, and high-affinity uptake increases only where a root also encounters nitrate (C-207). This is not evidence for a new controller. Backpressure and primal–dual allocation are the strongest nulls.

For module i , let $d_i(t)$ be unmet task demand per second, $a_i(t) \in [0, 1]$ be local capability availability, $u_i(t)$ be allocated compute-seconds per second or watts, and $B(t)$ be the total budget in the same unit as u_i :

$$\sum_{i=1}^n u_i(t) \leq B(t), \quad u_i(t) = 0 \text{ when } a_i(t) = 0.$$

The possible advantage is communication structure: compress deficits upward, return a small scarcity code downward, and keep the expensive feasibility decision at the receiver. It should lose when complete state is fresh, capability is uniform, or backpressure already carries the required

deficit and feasibility information more cheaply. C-1488 therefore keeps aggregate supply, local arrival, usable uptake, and export separate: a global balance cannot certify receiver feasibility. Candidate 013 tests moving demand and resource patches, delay, topology churn, strategic over-reporting, reversible allocation, slow growth, and second events.

Electrochemical depletion adds a geometry-sensitive failure case. A healthy global mean can conceal a receiving boundary whose local carrier supply is approaching zero; once the support assumptions change, positive feedback can amplify a protrusion rather than merely serve more demand (C-1536). The artificial analogue must therefore bind every growth or allocation action to a local support estimate and expose an `unknown` or fallback state when that estimate is not observable. Ordinary queue, backpressure, and robust load controllers remain the nulls, and the local estimator's sensing, messages, throttling, and false alarms are charged.

Engineering null models

Biological language is removed when an established method explains the result at the same cost.

PROPOSED ROLE	NULL MODEL THAT MUST BE INCLUDED	RESULT IF NULL MATCHES
detect unexpected input	calibrated Kalman innovation or likelihood residual	call it residual gating
distinguish a persistent change	Page/CUSUM accumulate–reset detector	call it change detection
decide whether to buy more information	one-step expected value of sample information	call it value-of-information control
skip depth on easy cases	calibrated intermediate classifier / early exit	call it adaptive depth
select capacity	top- <i>k</i> mixture-of-experts with load balancing	call it conditional routing
stabilize utilization or queue load	tuned PI and Kelly-style primal/dual allocation	call it feedback resource control
distribute operating context	FILM, recurrent gates, learned global tokens, and hypernetworks	retain the simplest winning conditioner
separate fast and slow control	tuned multirate or gain-scheduled supervisory controller	call it multirate control
place recurring state near work	static placement, LRU/cache policy, and NUMA-aware scheduling	call it caching or placement

The equations and unit requirements for innovation, value of information, CUSUM, allocation, and energy are recorded in the [engineering-analogue audit](#).

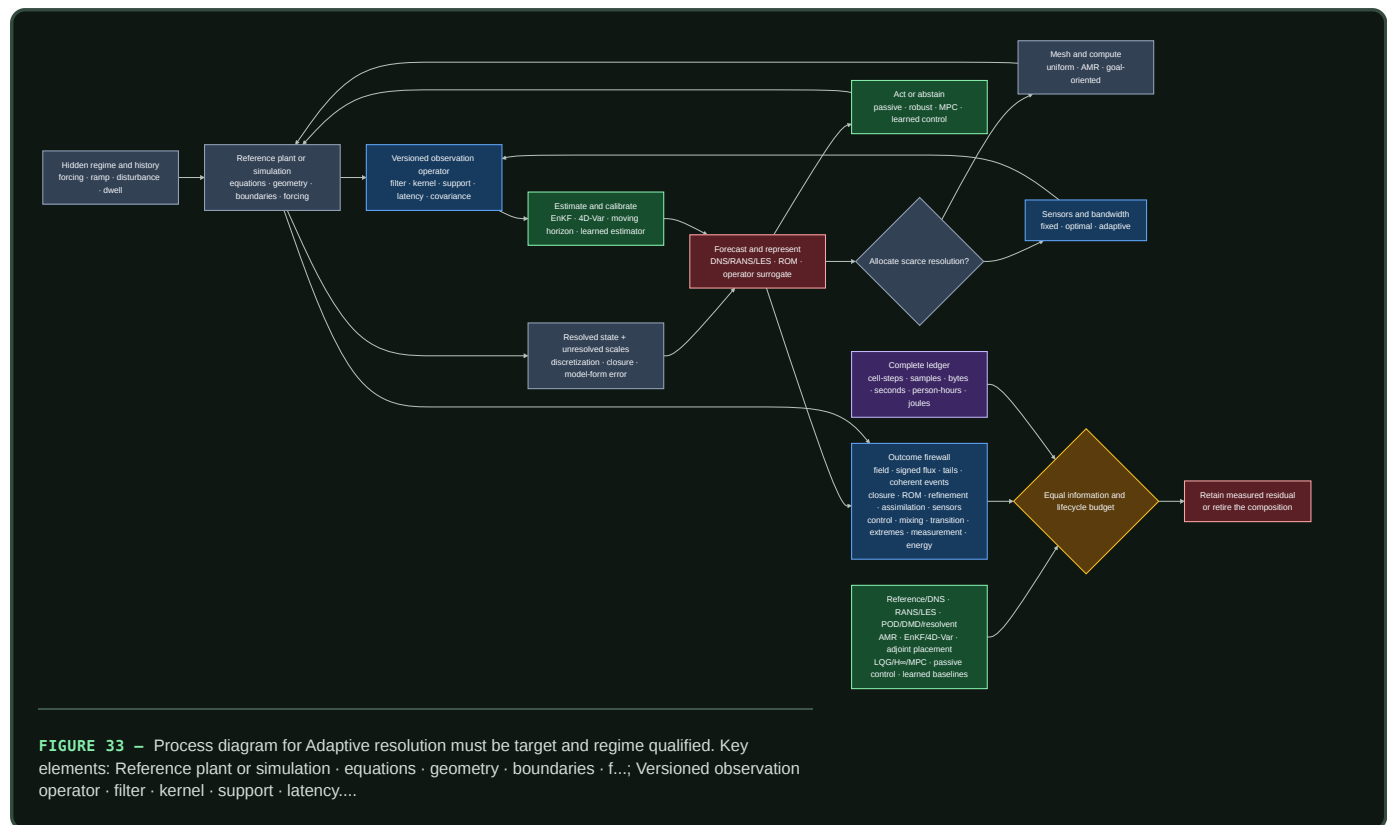
Adaptive resolution must be target- and regime-qualified

Fluid dynamics is a hostile test for “allocate more compute where prediction error is high.” The governing equations can be known while the useful state is still limited by unresolved scales, boundary and forcing uncertainty, discretization, closure error, partial observation, chaotic amplification, regime change, intermittency, and rare extremes. The audited evidence in C-881–C-925 therefore tightens the allocator contract.

An allocation signal is valid only relative to a declared target and measurement identity:

- **variable and support:** which physical or latent variable, spatial region, temporal window, and averaging kernel define the target;
- **filter and detector:** which scale, threshold, reference frame, event detector, and observation operator define a residual or coherent event;
- **signed relation:** whether energy, information, error, or another invariant should move toward larger or smaller scales under the current regime;
- **model boundary:** which part is resolved dynamics, closure, discretization, observation error, and out-of-support model-form discrepancy;
- **quantity of interest:** whether the reduced state must preserve average field reconstruction, control authority, transition timing, mixing, or an extreme tail; and
- **regime and history:** geometry, forcing, ramp direction, dwell, disturbance, prior occupancy, sensor drift, actuator condition, and current uncertainty.

A mean residual can be small while the signed scale flux is wrong, the tail is miscalibrated, or the rare event that matters is missed. A high-energy reduced basis can discard a weak direction that controls transition or actuation. A mesh can use fewer cells at one instant while spending more on regridding, subcycling, synchronization, load imbalance, data transfer, and failed solves. These are different failures and stay separate in the outcome vector.



Editable source: [regime-qualified-flow-inference-control.mmd](#).

Fixture F-005 crosses ten adversarial tracks: signed multiscale transfer, closure portability, target-qualified reduced state, adaptive resolution, assimilation, sensor placement, closed-loop control, mixing, path-dependent transition, and extreme prediction. Its decisive comparator is a complete composition of reference simulation, classical RANS/LES closures, POD/DMD/resolvent or balanced reduction, goal-oriented AMR, EnKF/4D-Var and moving-horizon estimation, adjoint/Fisher/Gramian sensor placement, robust/MPC/passive control, and learned operators or policies.

The flow mathematics binds every result to geometry, equation, boundary, forcing, solver, grid, measurement, filter, detector, data lineage, actuator, target, regime, and resource identity. An adaptive mechanism earns credit only when it improves its preregistered target beyond that complete stack without reversing flux, losing stability, uncovering uncertainty, missing the sealed natural tail, or producing negative net lifecycle energy. Otherwise the ordinary method remains and the proposed composition is retired.

Efficiency mechanism

Sparse work has four independent levers

The runtime can save resources by:

1. suppressing unchanged events;
2. exiting at a cheaper validated depth;
3. activating fewer modules or memory paths; and
4. avoiding movement by keeping repeated work near its state.

These levers are not interchangeable. An early exit can reduce arithmetic while leaving input and cache traffic unchanged. Sparse experts can reduce active weights while increasing all-to-all communication. Local placement can reduce bytes moved without reducing operations.

For module gates $g_i(x) \in [0, 1]$ and declared operation costs $c_i(x)$, define the operation-weighted active fraction

$$\rho_{\text{ops}}(x) = \frac{\sum_i g_i(x) c_i(x)}{\sum_i c_i(x)}.$$

Also report byte-weighted activity ρ_{bytes} from measured memory and network traffic. Neither quantity is an energy estimate.

Charge every control path

Over the same measurement interval,

$$E_{\text{adaptive}} = E_{\text{predict}} + E_{\text{gate}} + E_{\text{route}} + E_{\text{active}} + E_{\text{memory}} + E_{\text{network}} + E_{\text{context}} + E_{\text{control}} + E_{\text{idle}}.$$

Every term is in joules. The runtime produces an efficiency gain only when

$$E_{\text{adaptive}} < E_{\text{dense}}$$

at matched quality, risk, latency class, batch opportunity, and hardware boundary. Report board, node, and facility energy separately when measured; do not infer one from another.

State movement and controller work must also amortize. For a placement, compilation, or cache migration completed N times, let E_{runtime} be incremental runtime energy in joules per completed use; let E_{discover} , E_{migrate} , and E_{validate} be one-time energies in joules; and let N be a dimensionless completed-use count. Then

$$\bar{E}_{\text{use}} = E_{\text{runtime}} + \frac{E_{\text{discover}} + E_{\text{migrate}} + E_{\text{validate}}}{N}.$$

$\bar{E}_{\text{use}}$ is joules per completed use. If N is not observed, report the break-even reuse count rather than claiming a saving.

Resource supply is a feedback problem

For one optimization instance, choose a single rate basis—requests per second, tokens per second, or bits per second—and let x_i be module i 's allocation in that unit. Let $U_i(x_i)$ be a nondecreasing concave utility normalized to a common dimensionless scale, let R be a dimensionless resource-incidence matrix, and let each component of c be capacity in the same chosen rate unit. A conventional allocation null is

$$\max_{x \geq 0} \sum_i U_i(x_i) \quad \text{subject to} \quad Rx \leq c.$$

The slower resource plane observes queues, deadlines, utilization, energy, and temperature, then adjusts capacity or prices. Any learned controller must beat a tuned PI or primal/dual implementation on settling time, overshoot, constraint violations, tail latency, and joules per control update. “Local metabolism” is not a substitute for that comparison. Interacting loops with gain and delay must also be tested for oscillation and stability rather than accepted from their mean allocation (C-1492). Regime-dependent mediator supply must retain starvation, drag, leakage, failure, and pumping terms (C-1499); mean load also cannot clear a state-delay instability such as stick-slip (C-1501).

Evidence status

COMPONENT	CURRENT SUPPORT	STATUS FOR THIS ARCHITECTURE
strict biological activity budget	C-001	established biological constraint; digital magnitude unassigned
sparse conditional capacity	C-003	established AI mechanism; end-to-end benefit workload-dependent
input-dependent early exit	C-004	established on evaluated BERT tasks; risk-stratified generality open
cortical residual hierarchy	C-005	plausible engineering lead, not a universal brain objective
event-driven on-chip learning	C-015	feasible on one published substrate; superiority not established
slower activity stabilization	C-018	established in cultured neurons; artificial set point open
context-dependent local gain	C-020	established for the studied mouse circuit; routing value open
active sensing mechanics	C-022	established for the whisker preparation; acquisition policy open
inhibitory sparse discrimination	C-025	established for the fly circuit and task; hardware benefit open
congestion-triggered reserve route	C-035	established in the ant experiment; conventional routing remains the null
multirate broadcast and local decoding	C-046–C-048	biological observations established; AI mechanism is Candidate 002
local demand, placement, and supply	C-049–C-051	cellular observations established; artificial resource plane untested
sparse, delayed, pooled, and policy-coupled surveillance	C-122–C-131	scoped epidemiological and statistical evidence; AI-system translation is Candidate 007
joint process/observation estimation with action provenance	C-132	speculative composition against POMDP, detection, nowcasting, and value-of-information nulls
regime-qualified closure, reduction, assimilation, refinement, and control	C-881–C-925	scoped fluid evidence and formal relations; integrated allocator remains Fixture F-005

The integrated runtime is speculative until its gates are tested separately and then recombined under one measurement boundary. A combined win cannot identify which control path caused it.

Speculative extensions

Residual queues instead of global lockstep

Each module could own a queue of unresolved residuals and wake only when local work, context, or a deadline changes its priority. The comparator is an efficient batched scheduler with the same queue and dispatch overhead. The extension is abandoned if asynchronous launches reduce utilization or increase p95 latency enough to erase skipped work.

Predictive placement

Frequently reused expert weights, key–value state, and memories could move toward their expected consumers. The hypothesis is about reuse-distance and migration amortization, not artificial mitochondria. Static placement, LRU, and NUMA-aware policies remain required baselines; every migration is charged in bytes, seconds, and joules.

Controlled recovery probes

A maintenance observer could inject small bounded perturbations and measure recovery time, overshoot, and restoring margin before ordinary quality metrics fail. This is specified in [Candidate 003](#). Recovery diagnostics do not themselves stabilize the runtime, and their probe energy and task disturbance remain part of the cost.

Cross-modal acquisition brokerage

A common valuation layer could choose between deeper internal computation, a memory lookup, another sensor view, a physical action, a tool call, or a human query. Each option needs a calibrated outcome model and a common decision contract. Entropy reduction alone is not sufficient because an observation can be surprising yet irrelevant to the pending action.

Failure modes

FAILURE	OBSERVABLE SIGNATURE	REJECTION OR CONTAINMENT RULE
residual–uncertainty conflation	high-noise inputs always buy more compute without improving decisions	compare against normalized innovation and EVSI; reject residual-only gate
overconfident early exit	average accuracy holds while rare or shifted strata fail calibration	impose risk-stratified exit bounds and safe fallback
router collapse	a few experts saturate, queues and p95 latency rise, reserve capacity idles	compare load-balanced MoE and constrained allocation; reject unstable router
granularity below hardware break-even	active FLOPs fall but kernel count, bytes, latency, or joules rise	coarsen gates or return to dense fused execution
broadcast becomes a hidden router	context rate, token width, or receiver traffic scales with module count	enforce quantization, cadence, and logged bit/byte budgets
receiver saturation or mode lock	stronger context produces non-monotonic failure, unrecoverable hysteresis, or uniform module response	bound gains, test impulse responses, retain redundant fallback
positive feedback between routing and supply	busy modules receive more capacity, attract more traffic, and monopolize service	separate router and resource objectives; test step, burst, and delay stability
controller oscillation	periodic queue, power, or route changes with excess settling time	include anti-windup and delay-aware PI/primal-dual nulls; reject learned loop if dominated
resource-proxy gaming	predicted “useful work” rises while task utility per joule falls	reconcile estimates against measured outcome and energy
state-migration thrash	repeated placement changes exceed saved memory traffic	require hysteresis and observed break-even reuse count
shortcut no-op	unchanged-input gate suppresses slow but important drift	maintain persistent-change statistic and scheduled freshness checks
accounting boundary leak	reported savings omit host, network, idle, context, or control energy	withhold efficiency claim until the missing boundary is measured
mean-fit tail failure	average field error falls while signed flux, transition, or extreme-event calibration worsens	retain the full outcome firewall; reject pooled-score improvement
closure/numerics cancellation	a learned residual wins only on one solver/grid and degrades under refinement	separate closure, discretization, and model-form support; test hidden solver and grid lineages
adaptive-resolution bookkeeping leak	cell count falls while regrid, rejected steps, transfers, imbalance, or sensing dominate	compare complete work, wall time, bytes, and joules at equal target error
control saving without net saving	task drag or loss falls but actuation, sensing, compute, auxiliary, installation, or maintenance erase it	report service-interval net energy and reject non-positive benefit

Measurable predictions

These are experiment commitments. Directional improvements require paired uncertainty intervals and a predeclared practical margin; a lower theoretical operation count is not a pass.

ID	INTERVENTION AND COMPARATOR	PRIMARY MEASUREMENTS	PREDICTION AND FAILURE BOUNDARY
AC-01	calibrated early exit versus fixed depth and uncalibrated confidence threshold	error and calibration by risk stratum; layers/item; J/item; p95 ms	lower J/item at equivalent high-risk error and calibration; reject if gains come from degraded rare-event performance
AC-02	residual/uncertainty/EVSI acquisition policy versus raw residual threshold, Kalman innovation, CUSUM, and one-step EVSI	downstream utility; acquisitions/item; false escalations/item; bytes/item; J/item	candidate must move the utility–energy–latency frontier beyond the strongest composed null; otherwise use the null
AC-03	sparse expert and memory routing versus dense execution and load-balanced top- <i>k</i> MoE	active operations; memory and network bytes; queue depth; p95 ms; J/item	physical bytes and joules must fall with active operations at matched quality; a FLOP-only reduction fails
AC-04	Candidate 002 multiscale receiver versus FiLM, GRU gate, global token, hypernetwork, and multirate control	task error; impulse-kernel NRMSE; bit/s; bytes/step; J/step; p95 ms	require at least 5% task and 10% kernel improvement over the best equal-interface simple baseline, then task equivalence to the higher-bandwidth method with at least 25% fewer context bytes or 5% less energy and no more than 5% p95-latency increase
AC-05	local resource plane versus tuned PI, Kelly/primal-dual, and centralized fixed allocation	settling s; overshoot %; violations/s; utilization %; J/control update; J/item	learned/local control must improve a quality–tail-latency–energy frontier under burst and delay; equality means merge into conventional control
AC-06	predictive placement versus static placement, LRU, and NUMA-aware scheduling	migration bytes; cache misses/item; p95 ms; J/item; break-even reuse count	placement wins only after migration and validation amortize within observed reuse; otherwise retain the conventional policy
AC-07	difficulty-controlled input sets with equal length but varied ambiguity, relevance, and risk	acquired work/item; J/item; task utility; calibration	work should track expected decision value and risk, not length or irrelevant noise; extra work without utility improvement falsifies the allocator
AC-08	distribution shift and rare-event stress with every adaptive gate enabled	worst-stratum error; expected calibration error; missed-hazard probability; safe-fallback rate; J/item	adaptive savings must survive the declared risk bounds; a favorable average with a worse safety tail fails

ID	INTERVENTION AND COMPARATOR	PRIMARY MEASUREMENTS	PREDICTION AND FAILURE BOUNDARY
AC-09	Fixture F-005 complete regime-qualified composition versus its strongest classical/learned closure, ROM, AMR, assimilation, sensor, and controller stack	target-native field/flux/tail/event errors; calibration; stability; complete cell-steps, bytes, person-hours, and J/run	require a preregistered target improvement under hidden regime, solver, grid, observation, and hardware changes without degrading signed flux, sealed-tail calibration, stability, or net lifecycle energy; a tie retires the composition

Mechanism ablations are interpreted selectively. Removing the exit gate should increase depth without changing routing identity; removing sparse routing should increase active modules and traffic; removing the context fast path should selectively damage transient response; removing the slow path should selectively damage sustained response; freezing resource control should worsen burst recovery rather than semantic accuracy. If every ablation merely lowers capacity and hurts everything, the architecture has not isolated its claimed control loops.

Fast memory, replay, and consolidation

Experience should change behavior immediately without receiving immediate permission to rewrite stable capability.

Scope

Memory is a lifecycle: capture an event, preserve its origin, decide whether it deserves more work, test an integration, then retain, transform, externalize, weaken, or delete it. The objective is rapid adaptation without granting every surprising event permission to rewrite stable capability.

One memory lifecycle

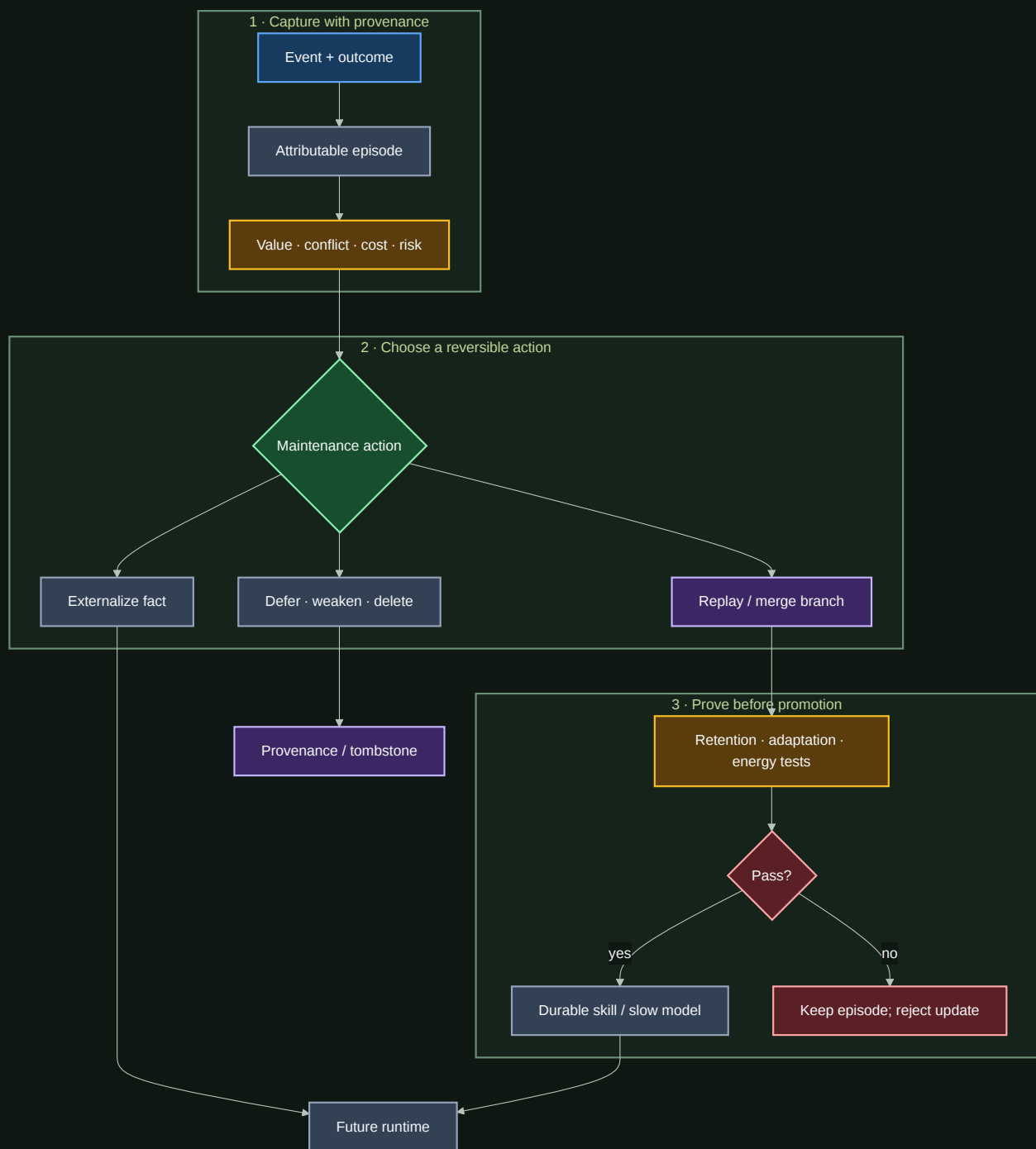


FIGURE 34 — Process diagram for One memory lifecycle. Key elements: 1 · Capture with provenance; Event + outcome.

Editable source: `../assets/diagrams/memory-lifecycle.mmd`.

The lifecycle separates two decisions often collapsed into “learning”:

- 1. What should be remembered now?** Capture is fast, attributable, and reversible.
- 2. What should change the durable system?** Consolidation is selective, tested, and budgeted.

Biological observation

Complementary Learning Systems theory describes interacting fast hippocampal and slower cortical learning processes (C-008). The useful pattern is not only different storage speeds. It is controlled transfer between a rapidly changing record of experience and structure whose value depends on remaining stable.

The evidence also makes replay a selection problem:

- disrupting replay from a selected hippocampal assembly can selectively impair the associated rodent spatial memory (C-036);
- replay allocation varies with reward, learning, familiarity, and memory weakness rather than following one universal priority (C-037);
- existing relational structure can accelerate integration of compatible associations (C-038);
- retrieval can make an established memory temporarily update-sensitive (C-039), although the exact human prediction-error gate remains disputed (C-040); and
- forgetting can be actively regulated by neural and glial mechanisms in specific preparations (C-041, C-042).

Together they motivate a maintenance controller that allocates limited work across capture, replay, integration, protection, and forgetting.

Proposed AI translation

1. Separate stores by write authority

STORE	UPDATE RATE	CONTENT	NORMAL MUTATION
Working state	every event	active context, goals, predictions	overwritten freely
Episodic store	rapid	sourced trajectories, outcomes, errors	append, expire, redact
Slow model	controlled	reusable representations and skills	validated consolidation
Factual store	independent	mutable, attributable propositions	explicit versioned update

The separation is about authority, not hardware. A vector store, database, recurrent state, adapter, and model weights may share a device while obeying different write policies. One logical store may also span devices when locality or retention requires it.

Memory mode, lifetime, and retrieval are separate axes

Plant systems expose useful counterexamples to treating every durable effect as one scalar “memory strength.” In Arabidopsis vernalization, a quantitative whole-organism response can arise from the fraction of loci or cells in a binary state rather than an analogue value stored in every unit (C-1516). The same state can persist through somatic growth yet be actively reset at a lifecycle boundary (C-1517). In heat-stress memory, the acute writer, persistent trace, and later reinduction are causally separable (C-1518). Regenerative preparedness can also be localized while its required reserve remains an explicit cost (C-1519).

An artificial memory record therefore declares at least five independent fields:

FIELD	QUESTION	REQUIRED COMPARISON
representation	analogue value, discrete state, population fraction, trace, or external record?	equal-state filter, quantizer, latch population, cache, or database
write event	which observation and authority create or revise it?	always-write, threshold, change-point, and versioned branch
maintenance lifetime	what keeps the state valid, and for how long?	TTL, decay, replay, checkpoint, and explicit refresh
retrieval gate	which recurrence, query, or context lets it influence action?	recurrent state, keyed retrieval, and calibrated classifier
reset boundary	which evidence authorizes weakening, deletion, or reinitialization?	no reset, fixed reset, evidence-gated reset, and ordinary adaptation

A relative-sensing reference is one concrete instance of this lifecycle. In the cited EGF/HGF system, background-dependent surface-receptor abundance is consistent with a ligand-specific stored reference, while production, delivery, turnover and recovery maintain it (C-1548). The artificial translation therefore stores channel identity, reference value, age, support, writer, reset and maintenance cost together. A current ratio without those fields is not a complete memory record.

This reference state competes directly with an exponential moving average, state-space estimator and compact recurrent state. It survives only when its channel specificity and task value remain after missing observations, cross-channel stimulation, stale state, forced reset and accelerated turnover, and after every reference write and fallback is charged. The lifecycle track is RSD-T09.

Hysteresis is state, not reset authority

A minimal binary hysteretic state makes the history dependence explicit. For dimensionless input u_t , retained state $m_t \in \{0, 1\}$, and ordered thresholds $\theta_{\text{off}} < \theta_{\text{on}}$,

$$m_{t+1} = \begin{cases} 1, & u_t \geq \theta_{\text{on}}, \\ 0, & u_t \leq \theta_{\text{off}}, \\ m_t, & \theta_{\text{off}} < u_t < \theta_{\text{on}}. \end{cases}$$

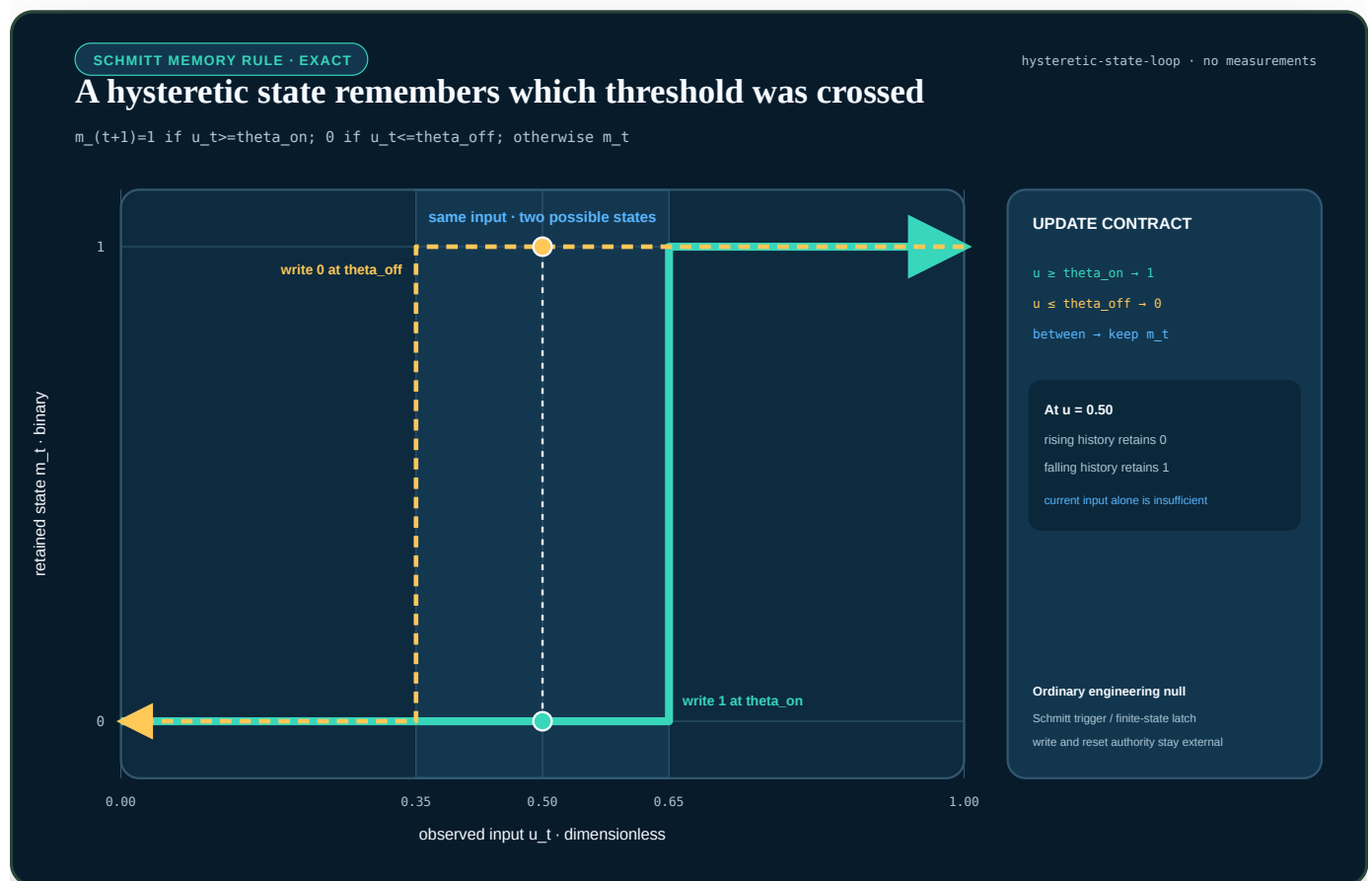


FIGURE 35 — An exact hysteresis loop for a binary Schmitt rule: inside the band, the same current input retains either zero or one according to the previously crossed threshold.

Inside the band, u_t alone does not identify m_{t+1} ; the prior state is a necessary input. This is useful precisely because it supplies a strong ordinary null: a proposed population memory must beat a Schmitt trigger, quantized accumulator, or finite-state latch under equal state, random-number, write, reset, and error budgets. The rule also keeps two questions separate. It defines how state persists, but it does not decide which authenticated event is allowed to reset that state at a lifecycle boundary.

History dependence also enters through acquisition order. Different final performance after the same set of tasks can reflect unequal age or exposure, ordinary catastrophic interference, curriculum selection, finite-capacity pre-emption, shared-state modification, facilitation, or irreversible lock-in. These are not interchangeable explanations. The **history-conditioned modular-succession audit** requires the same task multiset and eligible module identity set, exogenous presentations and task-local update ceilings, plus matched capacity, optimizer, evaluator and lifecycle budgets before estimating an order effect. Realized module lifecycle state remains an outcome. Endogenous routed acceptance is measured and equalized only in the exposure-cut cells. Its mathematical contract then crosses the candidate carriers as randomized interventions and kills the translation when canonical replay, scheduling, replay/EWC/OGD, or a search-budget-charged curriculum is non-inferior.

Two electrochemical results sharpen the test. First, eliminating a distributed diffusion field can yield a memory kernel whose apparent power law changes when a finite boundary becomes visible; the memory horizon is therefore a property of the support and observation band, not a free architectural constant (C-1531). Second, insertion systems can expose the same scalar occupancy with different terminal responses because the particle population followed different paths (C-1538). A single direction bit, finite-state latch, compact recurrent state, and balanced state-space realization are the required nulls before retaining a larger history or population state.

For a linear diffusion field $\partial_{tc} = D \partial_{zz}c$ on length L , the characteristic time

$$\tau_D = \frac{L^2}{D}$$

has units of seconds because L is measured in metres and the diffusion coefficient D in square metres per second. A proposed memory window shorter than the supported fraction of τ_D must reveal its truncation error; a window much longer than required must pay for retained state and movement. The full ECM-T02 and ECM-T09 contracts test both cases without assuming that an electrochemical kernel is the best software realization.

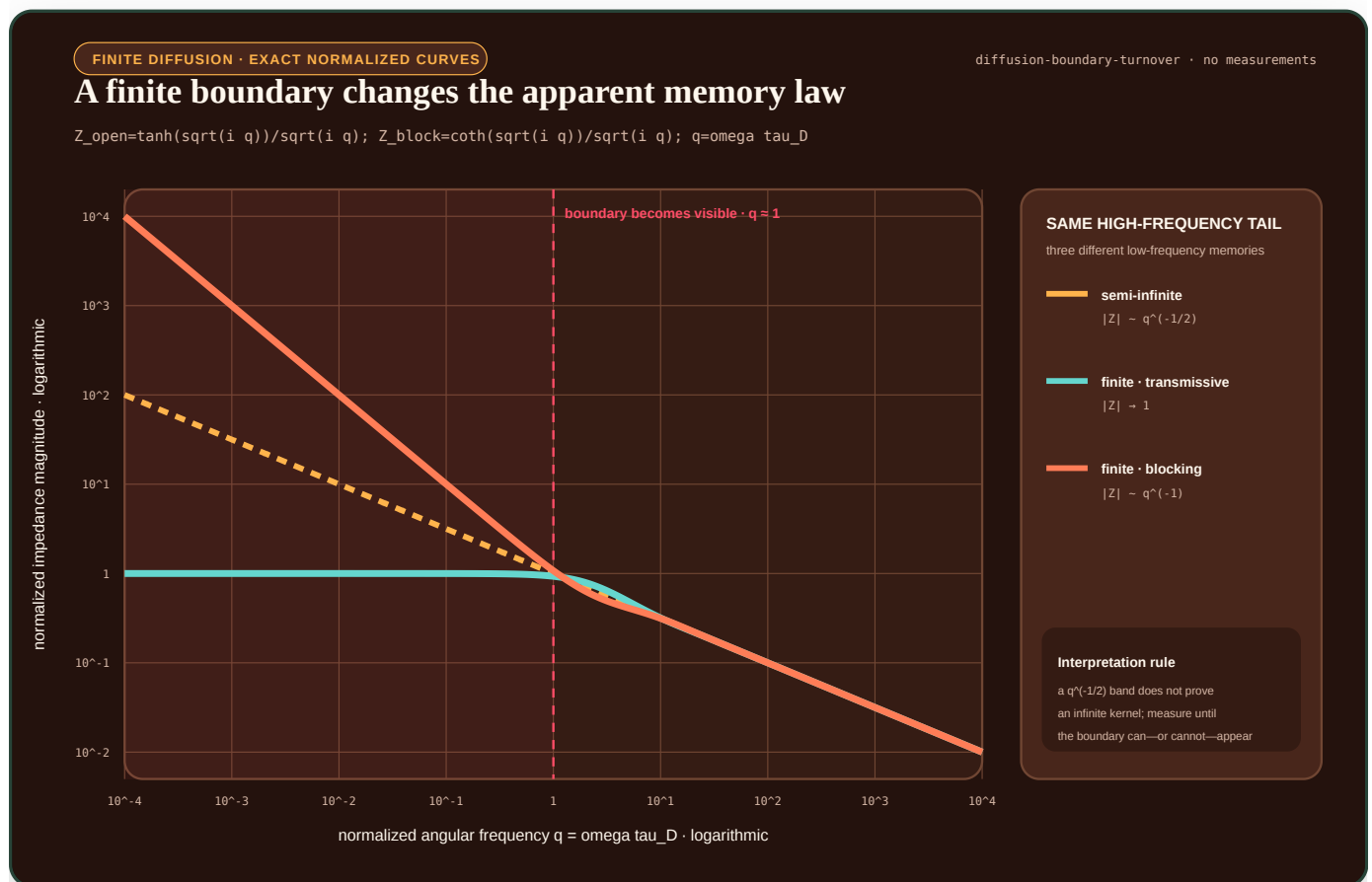


FIGURE 36 — Three normalized diffusion impedances share a high-frequency tail but split into transmissive, blocking, and semi-infinite low-frequency behavior when the boundary becomes observable.

With $q = \omega\tau D$, the plotted normalized finite-boundary forms are $\tanh(\sqrt{iq})/\sqrt{iq}$ and $\coth(\sqrt{iq})/\sqrt{iq}$. Their magnitudes approach different low-frequency limits even though both can resemble the semi-infinite $1/\sqrt{iq}$ law over a higher-frequency band. The visible turnover is why a fitted fractional-looking kernel cannot establish infinite memory by itself.

These fields expose three shortcuts to review: encoding duration with unjustified precision, calling incidental decay a designed reset, and treating trace correlation as the sole memory carrier. They also make reserve bytes, random state, calibration, refresh, reset, false retrieval, and recovery visible in the lifecycle ledger.

Plant signalling adds a second constraint: remembered state and current context may arrive on different routes. A cheap common alarm can require typed context before action (C-1520); route identity can be conditionally informative but must survive tag, capacity, maintenance, and common-mode accounting (C-1521). Active growth can create the evidence that gates later structural commitment (C-1522), while current exportable resource can condition future interface density only after its developmental delay and confounders are tested (C-1523). Boundary sensing plus a routed event wave remains a plausible composition, not an automatic advantage (C-1524), and a coupled light/temperature/history sensor is useful only inside an identifiable calibration envelope (C-1525).

Fixture F-023 turns these distinctions into ten preimplementation tests. It currently has `NO_RESULT`; no biology-derived efficiency factor is assumed.

2. Capture evidence before abstracting it

Each episode records enough context to explain a later update:

- observation and relevant prior state;
- action, answer, or intervention;
- outcome and uncertainty;
- data and tool provenance;
- active modules and retrieved memories;
- physical telemetry; and
- privacy, retention, and safety constraints.

The episode need not contain every hidden activation. It must preserve enough attributable evidence to reproduce, challenge, or reverse the lesson later derived from it.

3. Allocate the maintenance budget

At a maintenance window, the controller estimates what each candidate action could improve and what it would cost. Novelty, reward, uncertainty, familiarity, conflict, interference, and schema fit are features—not interchangeable definitions of importance.

The first controller should earn its complexity against ordinary policies:

1. uniform reservoir replay;
2. recency;
3. loss- or TD-error priority;
4. interference priority;
5. schema-fit priority; and
6. the proposed multi-signal lifecycle policy.

Every method receives the same episode bytes, replay examples, optimizer updates, wall time, and energy boundary. This makes scheduling policy—not extra maintenance—the independent variable.

The following single-item price envelope uses hypothetical gains and costs to make the selection equation visible. It is not a fitted scheduler or an empirical ranking of the listed actions.

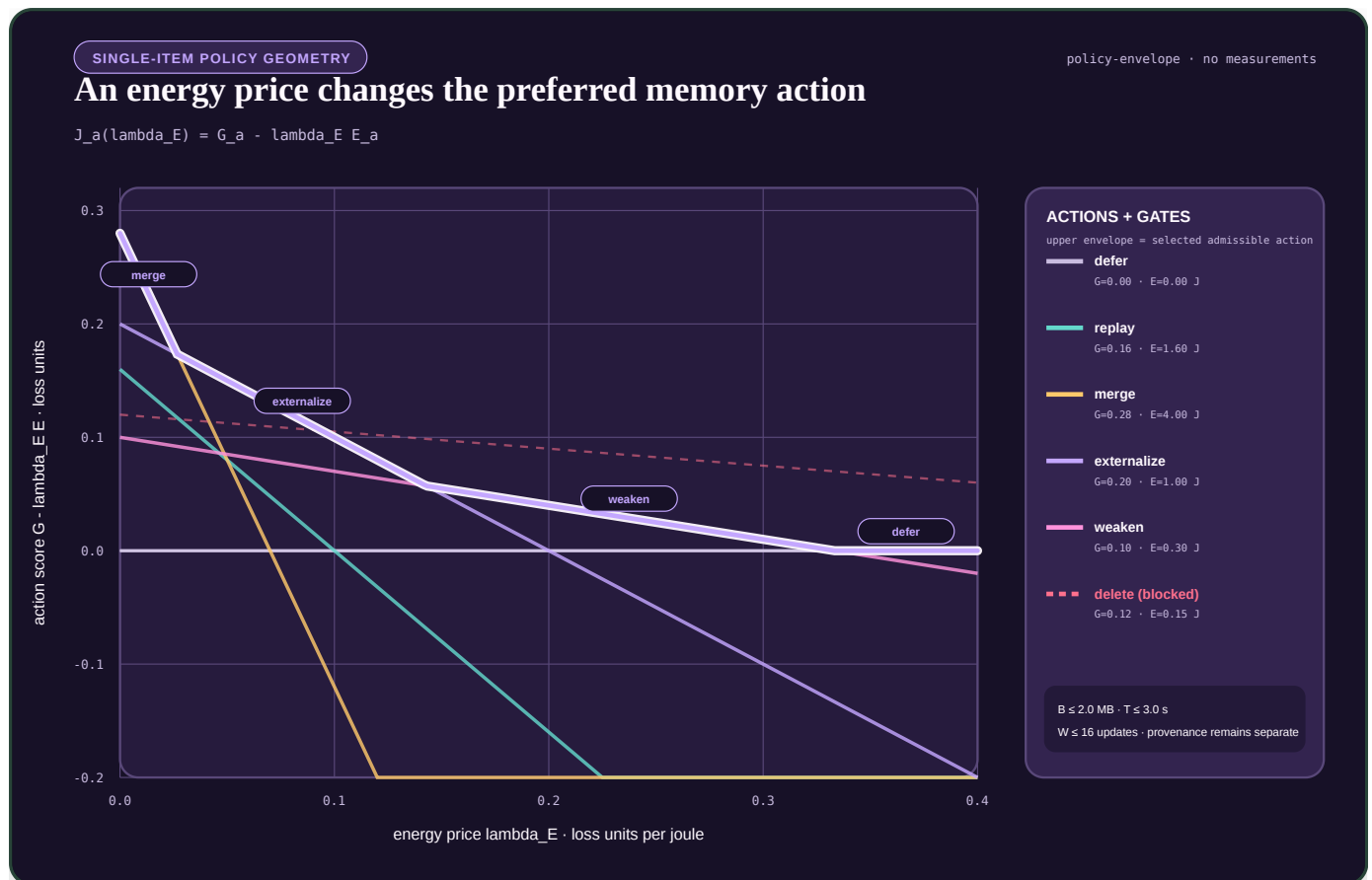


FIGURE 37 – For one memory item, each admissible action defines a score line G minus $\lambda_E E$; raising the energy price can switch the selected action, while byte, time, update, provenance, and destructive-loss gates can remove an otherwise high-scoring choice.

4. Branch, test, then promote

Retrieval or replay opens a versioned candidate branch. It never makes the durable state writable by itself. The branch may modify a local adapter, module, route, representation, or factual record and is then tested for:

- retention of protected historical capability;
- acquisition of the proposed new capability;
- calibration and rare-case behavior;
- provenance and conflict handling;
- measured energy, bytes moved, and latency; and
- reversibility after rejection.

Replay and Elastic Weight Consolidation are distinct comparison mechanisms (C-009, C-010); neither is privileged as the final protection policy.

A passed branch can merge into the slow model, remain provisional, or enter the maturity and structural-consolidation lifecycle. A failed branch returns the episode with its negative result attached so the same invalid integration is not proposed indefinitely.

5. Forgetting is an action

Unbounded retention consumes storage, retrieval bandwidth, replay capacity, and attention. The controller therefore distinguishes:

ACTION	EFFECT	REQUIRED SAFEGUARD
Defer	preserve without more work	future reconsideration rule
Replay	spend work to test or strengthen	equal replay budget

ACTION	EFFECT	REQUIRED SAFEGUARD
Merge	compress compatible state	provenance survives compression
Externalize	move mutable information out of weights	source and version retained
Weaken	reduce retrieval or routing influence	rare-case regression probes
Delete	remove active state	reconstructable source or explicit retention exception

Weakening is separate from deletion because many obsolete or interfering memories should first lose influence while evidence accumulates. A tombstone or provenance record prevents deleted state from becoming an unexplained absence.

Storage contracts precede memory policy

Persistent state has several independent contracts:

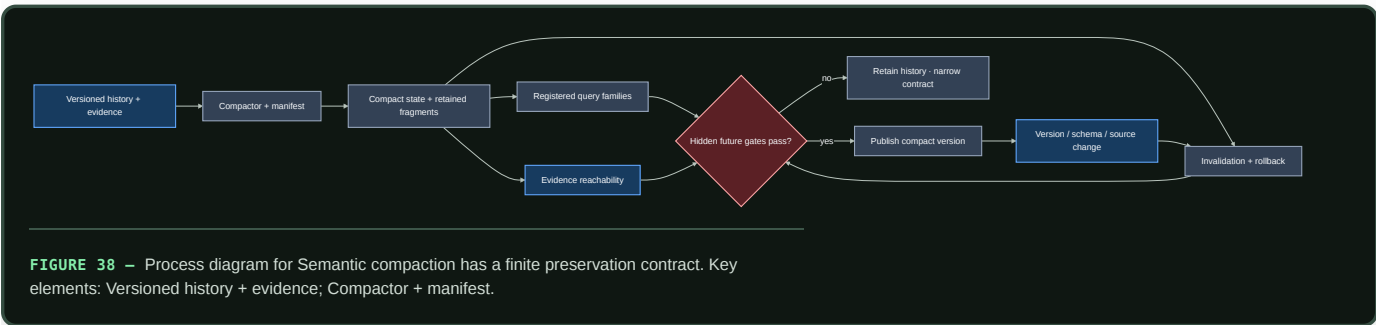
CONTRACT	QUESTION	DOES NOT ESTABLISH
atomicity and recovery	which in-scope effects commit or abort after failure?	isolation, outside-effect reversal, or truth
isolation	which concurrent histories are visible?	real-time order or durability
durability	what survives acknowledged completion under the fault model?	correctness or indefinite retention
version visibility	which snapshot or temporal coordinate answers a read?	serializability or current-world truth
replication/coding	which exact bytes or log survive named faults?	independent judgment or semantic diversity
indexing/caching	where is a candidate copy found cheaply?	importance, authority, or source-of-truth status
retention/reclamation	which roots, readers, holds, and horizons keep state live?	future irrelevance or epistemic worth

These boundaries are established by storage theory and systems evidence in C-325–C-338. Versioning alone does not make a history serializable; a durable statement can remain false; an agreed log can preserve a bad command; and event replay can change meaning when handlers, schemas, configuration, nondeterministic inputs, or outside effects are not version-compatible.

Valid time and system time remain separate. The first records when a proposition is asserted to hold in the modeled world; the second records when the store held that version (C-337). Capture, receipt, processing, correction, and supersession may add more clocks.

Semantic compaction has a finite preservation contract

Physical compaction preserves a declared storage view; it is not automatically semantic consolidation (C-334, C-341). A semantic compactor may replace history H with $Z = C_\phi(H)$ only under registered query, evidence, uncertainty, rollback, invalidation, deletion, and failure-degradation obligations.

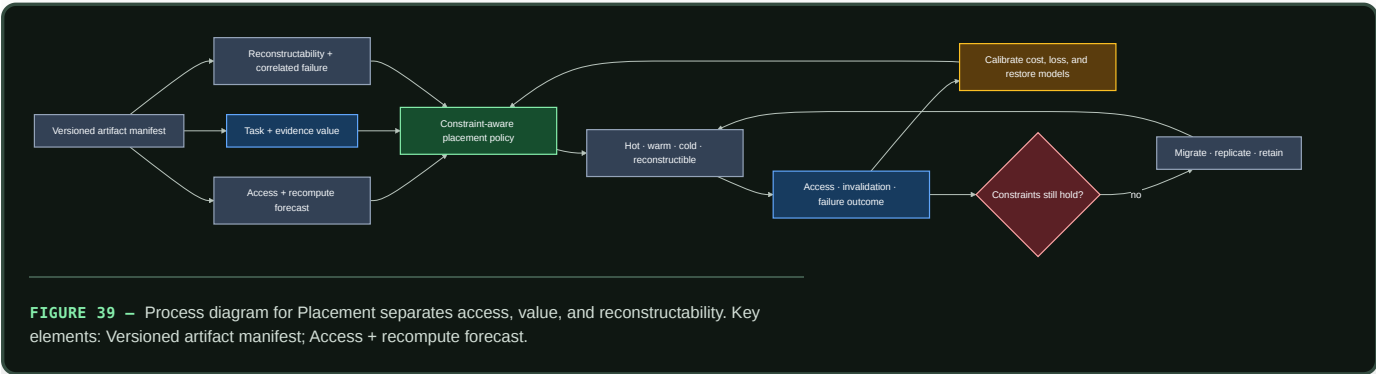


Editable source: semantic-compaction-contract.mmd.

Candidate 017 freezes the compactor before generating hidden future queries and invalidation requests. It compares with indexed history, snapshots plus suffix logs, materialized views, key compaction, lossless compression, extractive evidence summaries, and cold archives. Unsupported questions must be exposed, not filled with invented detail.

Placement separates access, value, and reconstructability

Recency and frequency are strong baselines, not definitions of importance (C-333). A rare cold artifact may still block a safety audit, source invalidation, rollback, or recovery. Conversely, a hot derived view may be cheap to reconstruct. The placement controller therefore keeps access forecast, recomputation cost, task loss, evidence value, staleness, and correlated-failure reconstructability as separately calibrated quantities.

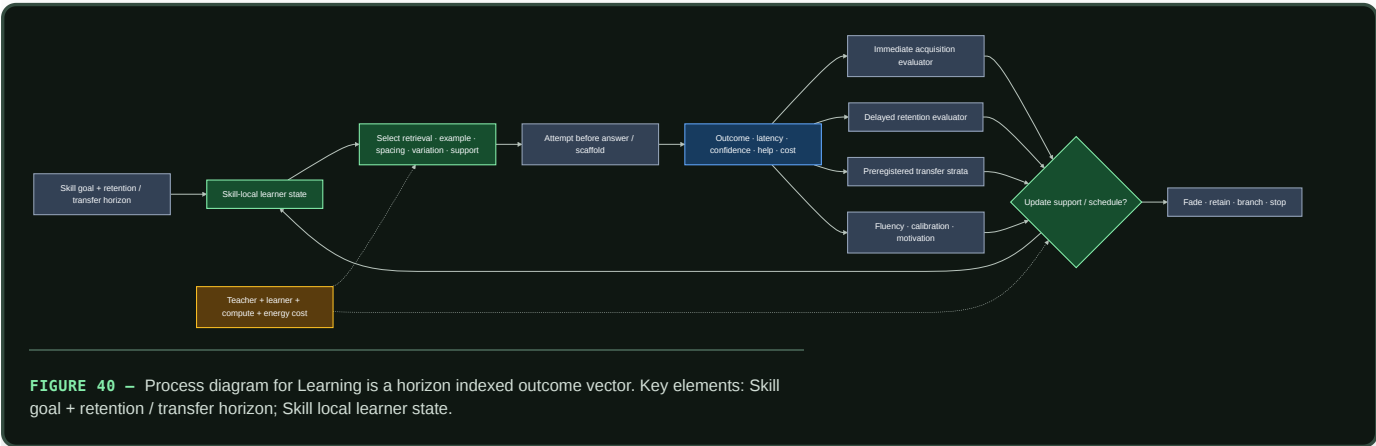


Editable source: [value-aware-artifact-tiering.mmd](#).

Candidate 018 must beat LRU/LFU, ARC, W-TinyLFU, size- and miss-cost-aware caching, economic tiering, static optimization, and coded fault-domain placement without future labels. Physical bytes, movement, metadata, endurance, p99 latency, policy work, privacy, and failure costs remain inside the boundary.

Learning is a horizon-indexed outcome vector

Immediate task success does not identify durable learning. A curriculum event can improve acquisition while weakening delayed retention, novel transfer, fluency, calibration, motivation, or total efficiency. The evaluator therefore keeps those outputs separate and makes retention horizon and transfer distance part of the claim. The evidence boundaries are C-627–C-658.



Editable source: [horizon-qualified-learning.mmd](#).

For each skill, retain current acquisition estimate, retention estimates at declared horizons, transfer by preregistered stratum, fluency/error frontier, calibration, support/scaffold state, confusability/context, intervention history, state version, and uncertainty. Difficulty is useful only when the attempt is processed successfully enough to generate information; failure is not beneficial merely because it was hard.

Candidate 004 must compare the adaptive sequence with tuned spacing, ordinary knowledge tracing, fixed fading, hard-example mining, and fixed curricula at equal attempt, feedback, time, storage, compute, and energy. Candidate 019 must compare interactive teaching with a versioned artifact plus tests, then measure retention and novel transfer after actual learner/model turnover while charging both teacher and learner effort. The [learning-outcome mathematics](#) defines the common evaluator.

6. Follow one event through the lifecycle

Suppose a tool returns a surprising result. Working state can use it immediately, while an attributable episode preserves the action, outcome, uncertainty, source, and active path. The slow model does not change yet. Related and conflicting outcomes accumulate until a maintenance window judges the episode worth replaying. Replay opens a local branch and tests it against protected history.

What happens next depends on the content. A reusable operation may become a skill or provisional module. A mutable proposition belongs in factual memory. Noise loses influence or expires under the retention policy. Later outcomes measure whether that choice was correct and recalibrate the scheduler. This is the feedback that turns a storage hierarchy into a lifecycle.

Efficiency mechanism

Online operation avoids immediate full-model gradient updates. Maintenance spends that work later and selectively, where expected future value justifies the physical cost.

For N_{served} events between maintenance windows, amortized energy per served event is

$$\bar{E}_{\text{event}} = E_{\text{online}} + \frac{E_{\text{maintenance}}}{N_{\text{served}}},$$

where E_{online} is runtime energy per event in joules and $E_{\text{maintenance}}$ includes scheduler, replay, memory movement, optimization, validation, and recovery energy in joules. A system saves energy only when the second term remains below the online work avoided by delayed and selective integration.

The full constrained action model—including joule, byte, second, and optimizer update budgets—is defined in `../math/memory-lifecycle.md`.

Evidence status

MECHANISM	EVIDENCE	CURRENT STATUS
Fast/slow learning split	C-008	established theory and supporting results; system translation incomplete
Interference protection	C-009	demonstrated in scoped sequential tasks
Replay for machine consolidation	C-010	plausible mechanism with task-specific evidence
Content-specific replay	C-036	established in the measured rodent intervention
Multi-signal replay allocation	C-037	established constituent observations; unified policy experimental
Schema-sensitive integration	C-038	established in scoped learning conditions
Retrieval-induced update window	C-039, C-040	lability established narrowly; exact human mismatch gate disputed
Active forgetting	C-041, C-042	established in scoped interventions; safe AI policy untested
Storage and temporal contracts	C-325–C-338	mature engineered mechanisms; mandatory nulls and vocabulary
Semantic compaction and value-aware tiering	C-339–C-342	held residual experiments under Candidates 017 and 018
Complete lifecycle controller	none	speculative synthesis

Speculative extensions

- Generate counterfactual variants around high-value episodes rather than only replaying recorded inputs.
- Learn expected knowledge gain per joule while retaining hard resource and safety bounds.
- Perform module-local consolidation first and synchronize globally only when a cross-module invariant changes.

- Use exact checkpoints to test several consolidation outcomes in parallel and retain the cheapest one that passes.
 - Learn memory placement jointly with replay policy so frequently paired state becomes physically local.
-

Failure modes

- Replay amplifies biased, adversarial, or privacy-sensitive episodes.
 - The scheduler starves quiet, rare, or safety-critical memories.
 - A correlated shortcut is mistaken for schema compatibility.
 - Generated replay drifts away from the environment.
 - Retrieval becomes an adversarial write primitive.
 - Weakening or deletion removes evidence later required for recovery or audit.
 - A compact summary passes familiar queries but loses evidence, invalidation, rollback, deletion, or rare hidden-future obligations.
 - A placement policy calls access frequency “importance,” leaks future value, or treats correlated replicas as independently reconstructible.
 - Scheduler scans and telemetry consume the saved maintenance budget.
 - The episodic store becomes an unbounded duplicate of the training corpus.
-

Measurable predictions

1. Fast attributable memory reduces adaptation latency without increasing protected slow-model regression.
2. A multi-signal scheduler improves the retention–adaptation–energy frontier beyond uniform, recency, loss-priority, and interference-priority baselines at equal replay count and bytes moved.
3. Schema-compatible episodes require fewer optimizer updates to integrate than violations while shortcut-controlled transfer remains unchanged or improves.
4. Explicit weakening reduces obsolete-memory intrusions without exceeding the declared rare-case deletion bound.
5. Separating mutable propositions from reusable skills reduces correction cost and unsupported factual carryover.
6. Maintenance energy amortized per served event remains below the online training work it replaces.
7. Registered semantic compaction reduces physical bytes without reducing hidden-query, evidence, rollback, or invalidation coverage below threshold.
8. Value/reconstructability features improve shifted-workload artifact placement beyond strong cache and storage policies after migration cost.

Maturity, grokking, and reversible structural consolidation

Scope

Define how a useful but plastic structure becomes protected, cheaper to run, and eventually eligible for pruning—without mistaking age, low training loss, or a grokking curve for proof of maturity.

Maturity is a lifecycle state, not a compliment. A mature module has earned a narrower update surface because its behavior is understood inside a declared validation envelope. That protection must remain reversible when the envelope changes or the module becomes brittle.

Biological observation

Development does not produce a final circuit by training every connection at a constant rate forever. In the studied mouse retinogeniculate preparation, relative activity and complement signaling affected microglial engulfment and retention of developing inputs (C-043). The useful abstraction is a slow maintenance process that helps refine structure under local evidence; it is not a universal deletion rule.

Mature circuits can also carry local constraints on further change. Targeted interventions on extracellular structure and a cholinergic brake reopened specific forms of adult visual-cortex plasticity (C-044, C-045). Protection and plasticity are therefore not opposite endpoints. They can be different operating states of the same structure, controlled locally and revisited under an explicit intervention.

Two results from ecological systems sharpen the control problem. Recovery from small perturbations slowed as one cyanobacterial microcosm approached a controlled tipping point (C-058); warning statistics also appeared during a manipulated whole-lake food-web transition (C-059). A system may look acceptable at rest while its restoring dynamics weaken. Maturity therefore cannot be certified by steady-state accuracy alone.

Nor is resilience one number. Greater species richness improved temporal stability but reduced resistance to warming in a large ciliate-microcosm experiment (C-060). In a separate microbiome analysis, functional redundancy was associated with resistance to newcomer engraftment (C-057). Stability, resistance, recovery, and capacity to admit a better replacement can move in different directions.

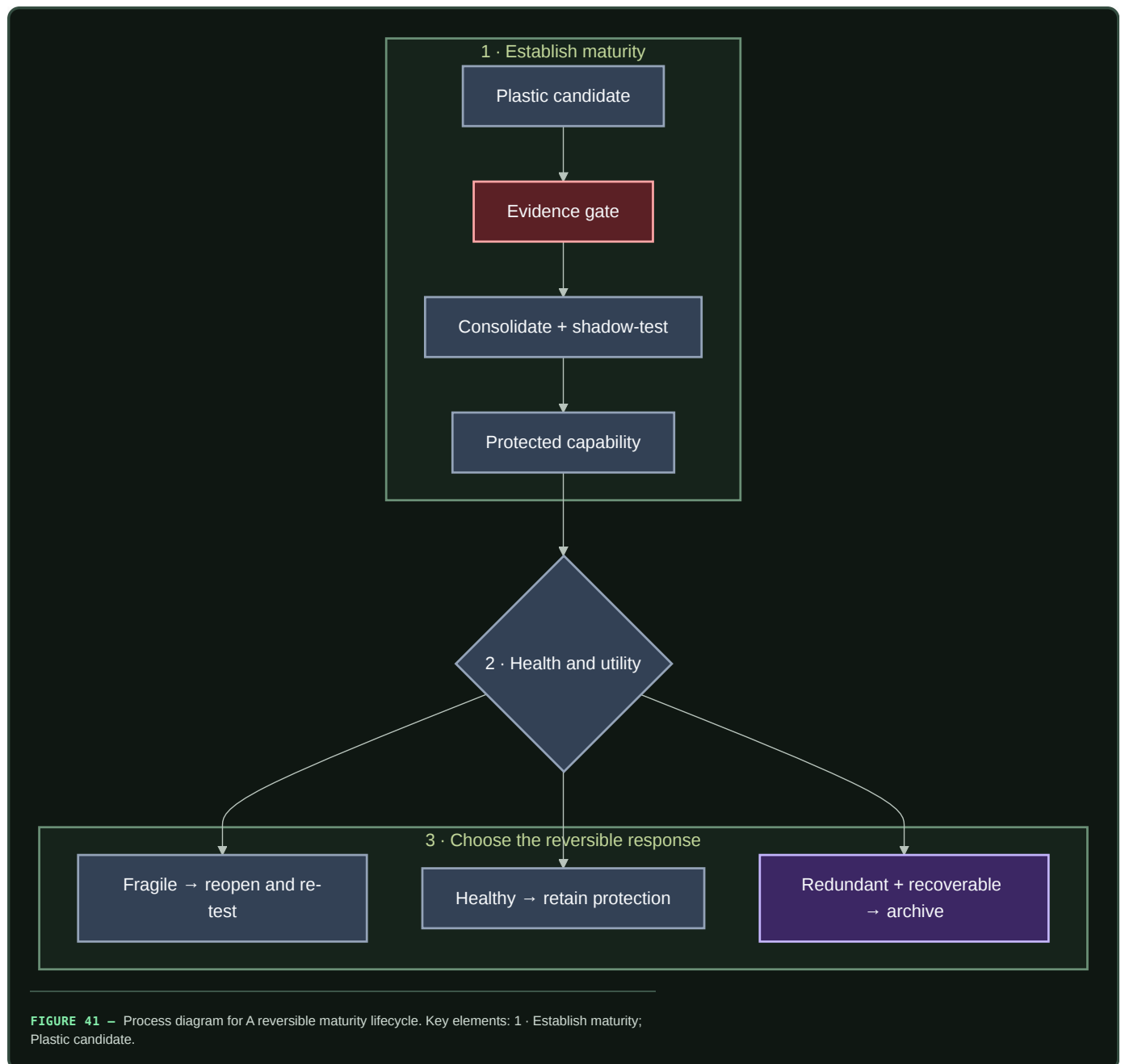
This makes grokking a useful phenomenon but a poor gate. Delayed generalization has been observed in bounded algorithmic settings, yet the claim that extended training reliably reveals the underlying rule and certifies readiness to prune is disputed (C-011). The system needs evidence about what a structure does, how uniquely it contributes, how it fails, and how it recovers.

Proposed AI translation

A reversible maturity lifecycle

Every structurally changeable module, route, memory transform, or compiled path has an explicit lifecycle state:

1. **Candidate:** highly plastic, attributable to its training episodes, and cheap to discard.
2. **Consolidating:** replayed against related, conflicting, rare, and intervention cases while its unique contribution is measured.
3. **Protected:** update rate and writable surface are reduced; structural changes require a versioned branch and shadow evaluation.
4. **Reopened:** a copy-on-write branch receives bounded adaptation while the protected version remains available for comparison and rollback.
5. **Retiring:** traffic is drained only after another path covers the required behavior and the physical execution graph can actually be compacted.
6. **Archived or removed:** provenance, validation envelope, and a reconstructable checkpoint remain for the declared retention period; hot execution state is released.



Editable source: [../assets/diagrams/maturity-fragility-cycle.mmd](#).

The lifecycle separates three decisions often collapsed into “pruning”:

- **protect**: this path is useful and should stop drifting;
- **reopen**: this path no longer responds adequately inside its required envelope; and
- **retire**: this path is no longer uniquely useful and can be removed without making the system irrecoverable.

Protection is not retirement. A path may be mature precisely because it is important enough to preserve.

Maturity is a vector gate

For module i , maintain a maturity record

$$\mathbf{m}_i = (Q_i, U_i, S_i, F_i, A_i, C_i, P_i),$$

where:

- Q_i is quality across the declared validation envelope;
- U_i is unique causal contribution under ablation and rerouting;
- S_i is behavioral and routing stability across time and environments;
- F_i is fragility measured from recovery and margin estimates;
- A_i is adaptation and newcomer-acceptance behavior under controlled shift;
- C_i is full-lifecycle physical cost; and
- P_i is provenance and rollback completeness.

The record remains a vector. A weighted score may rank candidates for review, but it cannot hide a failed safety, provenance, recovery, or rollback constraint. Thresholds are workload-specific and include uncertainty intervals; there is no universal maturity age, pruning percentage, or recovery constant.

Protection requires evidence that the module is useful, stable, attributable, and recoverable. Retirement reverses one condition: its *unique* contribution must be low because another tested path covers its role. Low weight magnitude or low average routing frequency is not enough.

Structural consolidation before deletion

When several routes repeatedly implement the same stable computation, the maintenance plane first tries to make their shared work explicit:

1. identify the recurrent subgraph and its boundary contract;
2. build a compact candidate through structured pruning, distillation, compilation, quantization, fusion, or relocation;
3. replay both common and conflicting cases through old and new graphs;
4. intervene on each source module to measure residual unique behavior;
5. shadow the compact path under live-like traffic;
6. drain old routes gradually while keeping rollback state; and
7. release tensors, optimizer state, routing entries, and communication only after the observation window passes.

Iterative pruning can reveal competitive sparse subnetworks in its tested settings (C-012). Here it is one operator inside the lifecycle, not the lifecycle policy itself. Magnitude pruning is a baseline; causal coverage and end-to-end physical savings decide promotion.

Reopening without overwriting the canonical path

A protected path is reopened only after a persistent signal, such as:

- calibrated error or intervention failure outside its historical variance;
- recurring novelty that existing candidates cannot absorb without interference;
- loss of recovery margin despite acceptable steady-state quality;
- repeated rollback or escalation around the same boundary; or
- evidence that protected redundancy prevents a superior newcomer from receiving a fair evaluation.

Reopening creates a branch. It does not make the canonical path globally writable. The branch receives a declared update, data, compute, and duration budget; the protected version continues on control traffic. The branch becomes canonical only after retention, calibration, intervention, cost, and recovery tests. Otherwise the branch is discarded and the trigger is retained as an unresolved event.

Recovery dynamics as a maturity signal

Snapshots answer whether a module is currently inside its envelope. Recovery dynamics ask how strongly it returns after a small displacement. The maintenance plane should begin with passive fluctuation analysis. If the state is a shadow or replica and the service budget permits it, a bounded probe can estimate return time or a local stability margin.

The proposed Stage-1 test is Candidate 003. It compares bounded recovery probes with SLO dashboards, queueing headroom, change detection, passive autoregressive estimates, and standard active system identification under equal budgets. If ordinary system identification performs as well, the system should use it; the design requirement is visibility into restoring dynamics, not a special biological estimator.

For the Candidate 003 linear simulator, the recovery-time curve is analytical. The ten-second line is its declared Stage-1 engineering threshold, not a biological constant, production alarm, or failure-time prediction.

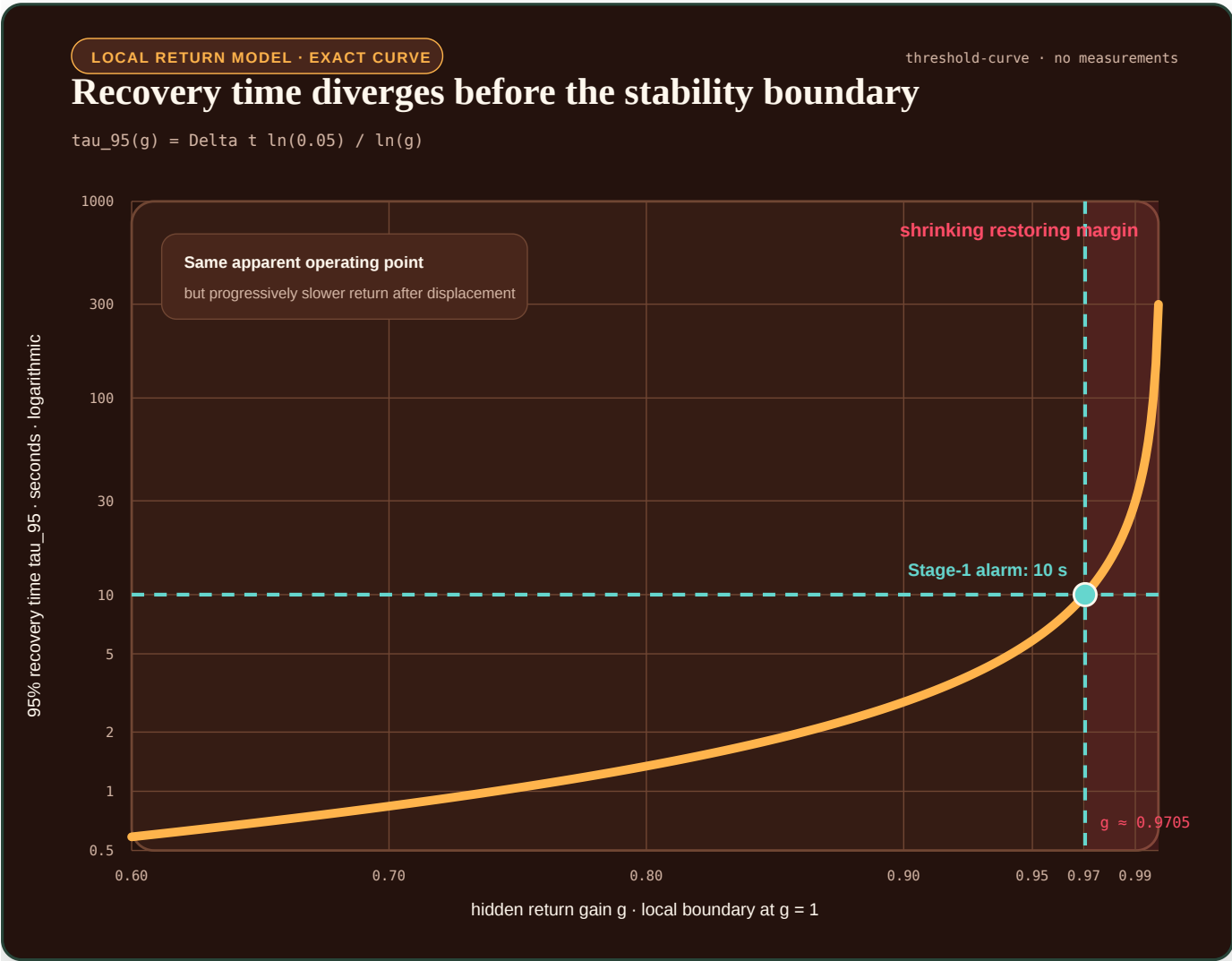


FIGURE 42 – The operating point can look unchanged while recovery time diverges as the local return gain approaches one; the ten-second line is a Stage-1 engineering threshold rather than a biological constant or predicted failure date.

Recovery is only one axis. A maturity record should expose at least:

DIMENSION	QUESTION	EXAMPLE MEASURE
temporal stability	Does behavior fluctuate under a stationary regime?	quality variance per event
acute resistance	How far does quality fall during a bounded perturbation?	maximum quality loss, fraction
recovery	How quickly and completely does behavior return?	return time, s or events

DIMENSION	QUESTION	EXAMPLE MEASURE
adaptability	How much work is required to learn a valid new regime?	updates, examples, and J
newcomer acceptance	Can a better candidate receive traffic and prove itself?	time to useful routing share, s
rollback	Can the prior behavior be restored after a failed change?	success fraction and lost work
reserve	What capacity remains for unexpected demand or repair?	bytes, W, routing slots, or RE/s

These quantities must be reported separately before any aggregate resilience score is computed.

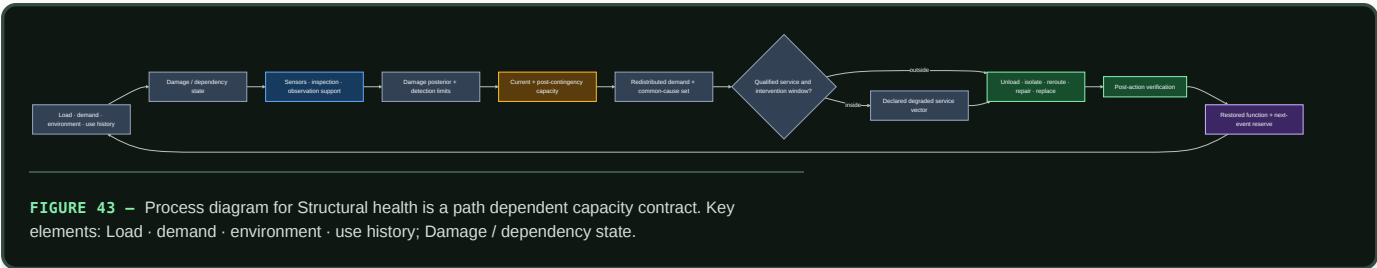
Structural health is a path-dependent capacity contract

A module or route can continue producing output after pruning, damage, drift, or overload while its remaining margin and next-event tolerance collapse. “Healthy,” “redundant,” “gracefully degraded,” and “reserve” are therefore not single scalar states.

For every structurally consolidated asset, record:

1. damage or dependency posterior and cumulative load/use history;
2. observation method, calibration, support, and detection limits;
3. present capacity and post-contingency capacity in native units;
4. surviving load/routing paths, common causes, and redistributed demand;
5. permitted degraded-service vector, affected strata, and exposure duration;
6. remaining time to a constraint boundary and available intervention;
7. unload, isolation, reroute, repair, replacement, or rollback action;
8. post-action verification; and
9. restored next-event reserve rather than merely returned output.

Mechanism-qualified wear is a direct falsifier of endpoint-only health: equal cycle count or terminal wear volume can hide different damage histories, repair states, and remaining service (C-1500).



Editable source: [residual-capacity-contract.mmd](#).

The mechanics and network evidence in C-481–C-499 establish mature nulls and measurement boundaries, not an AI effect size. The combined contract in C-500 remains an experimental schema shared by Candidates 005, 012, and 014. It is rejected if ordinary mechanics/reliability, asset management, fault handling, and network assignment match it.

Efficiency mechanism

Maturity can reduce recurring work in four places:

- protected modules need fewer parameter writes, optimizer states, and global synchronization events;
- stable routing can use smaller decision surfaces and better placement;
- recurring subgraphs can become fused, quantized, compiled, or cached paths; and
- retirement can release whole tensors, memory pages, routing entries, and network transfers.

The benefit is physical only when the runtime uses the new structure. Zero weights inside a dense kernel and dormant experts that are still loaded or synchronized do not count as structural consolidation.

For a proposed consolidation serving N future events, let $E_{\text{run,after}}$ and $E_{\text{run,before}}$ be steady runtime energy in joules per served event. Let $E_{\text{consolidate}}$, E_{validate} , E_{probe} , and E_{migrate} be one-time energies in joules, let $\mathbb{E}[E_{\text{recovery}}]$ be expected recovery energy in joules including failed branches, and let N be a dimensionless future-event count. Compare amortized energy per event:

$$\bar{E}_{\text{after}} = E_{\text{run,after}} + \frac{E_{\text{consolidate}} + E_{\text{validate}} + E_{\text{probe}} + E_{\text{migrate}} + \mathbb{E}[E_{\text{recovery}}]}{N},$$

against $\bar{E}_{\text{before}} = E_{\text{run,before}}$. The one-time numerator divided by N and both runtime terms are joules per event, so the comparison is dimensionally closed. Storage, data movement, tail latency, quality, calibration, and risk remain separate constraints rather than being silently converted into energy.

A consolidation advances only if its observation horizon is long enough that $\bar{E}_{\text{after}} < \bar{E}_{\text{before}}$ and it improves or preserves the declared quality–risk–latency–resilience frontier. The expected recovery term must include failed branches and rollback, not only successful releases.

Evidence status

ELEMENT	STATUS	WHAT IT SUPPORTS HERE
delayed generalization as universal maturity certificate (C-011)	disputed	grokking cannot be the gate
competitive sparse subnetworks (C-012)	established in tested settings	staged pruning is a valid operator, not a universal policy
activity-sensitive developmental refinement (C-043)	established in the cited preparation	a slower maintenance process can participate in structural contraction
reopening mature plasticity (C-044, C-045)	established in the cited preparations	protection can be local and reversible under intervention
redundancy and newcomer engraftment (C-057)	plausible association	stability may obstruct beneficial replacement
recovery warning signals (C-058, C-059)	established in scoped ecological systems	restoring dynamics are worth testing as a fragility signal
multidimensional stability tradeoff (C-060)	established in the cited microcosms	resilience dimensions must remain separate
mechanics, residual capacity, damage tolerance, and redistributed network load (C-481–C-499)	established or plausible in scoped engineering models	mature nulls for structural health, reserve, and recovery
path-dependent residual-capacity contract (C-500)	speculative synthesis	cross-layer fault-injection schema only
complete digital lifecycle controller	speculative synthesis	requires isolated and composed experiments

Speculative extensions

- Learn a lifecycle policy from logged promotion, rollback, and recovery outcomes while retaining hard provenance and safety constraints.
- Preserve cheap seed capacity outside the hot graph so retirement does not eliminate the ability to specialize under a future regime.
- Distill a coalition into a compact composite module, then retain the sources in a cold checkpoint until the composite survives a full recurrence cycle.
- Let recovery margin determine degrees of protection: lower update rate, narrower writable interfaces, or more stringent branch validation.
- Use capability-gap analysis to decide whether a failing mature module should reopen, be complemented by a newcomer, or retire.
- Reuse mature relational structure as a prior for faster consolidation while routing violations to a longer validation path.

Failure modes

- **False maturity:** a shortcut is stable on average and is protected before compositional, rare-event, or intervention tests expose it.
- **Brittle maturity:** quality remains inside its SLO while return time grows and the structure loses restoring margin.
- **Maturity monopoly:** protected redundant modules absorb all traffic and prevent a better newcomer from establishing evidence.
- **Rare-function erasure:** average routing and magnitude tests delete a path whose unique role appears only in a low-frequency or safety-critical regime.
- **Reopening thrash:** noisy triggers repeatedly create branches, retraining cost, and routing churn without a durable regime change.
- **Rollback rot:** checkpoints exist but dependencies, data schemas, or routing contracts have changed enough that restoration no longer works.
- **Cosmetic sparsity:** parameter count falls while bytes moved, kernel work, synchronization, and wall energy do not.
- **Maintenance inversion:** replay, probes, shadow traffic, migration, and regression testing consume more energy than mature execution saves.
- **Scalar resilience:** one score declares success while acute resistance, recovery, adaptability, or newcomer acceptance has deteriorated.
- **Causal misattribution:** a correlated low-usage path is pruned even though it stabilizes another module or handles a hidden confound.
- **Hidden reserve loss:** current output remains acceptable while damage, redistributed demand, common causes, or depleted intervention margin remove post-contingency capacity and next-event reserve.

Measurable predictions

1. A vector maturity gate using causal contribution and cross-context tests preserves rare and intervention performance better than fixed schedules, magnitude pruning, or training-loss thresholds at matched retained capacity.
2. Protected modules require fewer update joules and suffer less interference than continuously plastic modules, while branch-based reopening reaches a valid new regime with less regression than full unfreezing.
3. Recovery features identify some hidden loss of stability margin before steady-state SLOs. They remain in the architecture only if they add value beyond passive monitoring and standard system identification at matched probe and compute cost.
4. Structured consolidation lowers parameter bytes resident, bytes moved per event, communication, and measured joules together. Parameter reduction without those physical changes is rejected.
5. Increasing redundancy improves some stability dimensions while degrading adaptation or newcomer admission in at least one controlled regime; the raw resilience vector reveals the tradeoff that a scalar score hides.
6. Copy-on-write reopening plus rollback reduces catastrophic update loss relative to in-place adaptation after monitoring, checkpoint, and shadow costs are counted.
7. Consolidation produces a net lifecycle energy benefit only above a measurable reuse horizon N ; below that horizon, leaving the computation plastic or interpreted is cheaper.
8. Residual-capacity fields predict unsafe second-event or redistributed-load failures beyond ordinary output, utilization, health, and redundancy scores; otherwise the cross-layer schema is removed.

Hardening, reflex paths, and factual memory

Repeated transformations, reusable skills, and mutable propositions require different execution, update, and recovery contracts.

Scope

This chapter defines the boundary between five runtime outcomes:

- 1. a **compiled reflex path** executes a narrow, qualified transformation;
- 2. a **reusable skill** handles variable situations without embedding volatile propositions in its parameters;
- 3. **versioned factual memory** supplies mutable claims with provenance and validity state;
- 4. **escalation** sends an unresolved or high-risk event to a more capable model, tool, or human;
- and
- 5. **rollback** restores a last-known-good path or record after invalidation.

Hardening is a promotion decision, not a synonym for freezing. Every promoted artifact retains an applicability envelope, version, owner, regression set, physical cost record, invalidation policy, and recovery target.

Runtime paths and recovery

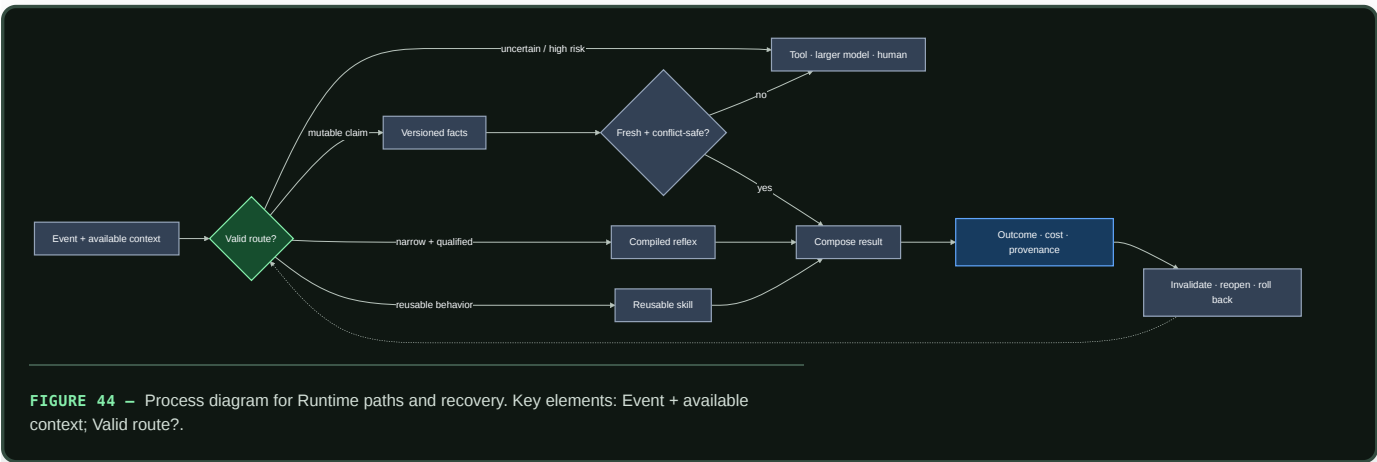


FIGURE 44 — Process diagram for Runtime paths and recovery. Key elements: Event + available context; Valid route?.

Editable source: `../assets/diagrams/hardening-memory-paths.mmd`.

The gate selects authority, not merely compute. A reflex may execute but cannot silently update a fact. A retrieved record may inform an answer but cannot rewrite a skill. An escalated result enters durable state only through the ordinary memory and consolidation lifecycle.

Biological observation

Biological and computational evidence supports separating rapid acquisition from slower integration (C-008). Retrieval can also make an established memory temporarily update-sensitive in a scoped preparation (C-039), while prediction error is disputed as a precise general trigger for that transition (C-040). Separate intervention studies show that mature constraints can be reopened under specific conditions (C-044, C-045).

The transferable requirements are multiple update timescales, guarded promotion, local reopening, and recoverable versions. The compiled path, factual schema, and escalation protocol below are engineering mechanisms tested against conventional systems.

Proposed AI translation

Five distinct contracts

PATH	STORES	MAY DO	MUST NOT DO
Compiled reflex	a narrow transformation, guard, and version	return a bounded result or action at low dispatch cost	answer outside its envelope or contain independently mutable facts
Reusable skill	a representation, policy, or parameterized transformation	generalize and compose across qualified contexts	present volatile propositions as current without retrieval
Factual memory	typed propositions and source relations	retrieve, supersede, dispute, revoke, or expire records	become true merely because retrieval ranked it highly
Escalation	no durable knowledge by itself	obtain more evidence or computation under a declared budget	bypass provenance, access, or promotion rules
Rollback	last-known-good versions, tombstones, and recovery metadata	restore routing and reconstruct prior state	erase the failed version or its affected-output trace

Compiled reflex paths

A compiled reflex is a versioned executable graph h with:

- an input and output schema;
- a validity predicate $G_h(x, \mathcal{I}_t) \in \{0, 1\}$;
- a declared applicability envelope $\mathcal{A}_h$;
- a deterministic or bounded-stochastic execution contract;
- a protected regression and adversarial set;
- a fallback route and last-known-good version; and
- a manifest for precision, kernels, placement, dependencies, and measured physical cost.

Here x is one event, $\mathcal{I}_t$ is the information available at decision time t , G_h is dimensionless, and $\mathcal{A}_h$ names the permitted input, environment, dependency, and risk strata. The runtime may dispatch h only when $G_h = 1$ and every dependency version remains valid. Guard evaluation is part of the path's latency, energy, and error budget.

Suitable candidates include parsing a fixed protocol, executing a stable local control law, applying a verified transform, or serving a repeatedly observed low-risk subgraph. Ordinary code, rules engines, memoization, and compiler optimization remain the first alternatives. A learned compiler is useful only if its generated guard and path outperform those alternatives under the same coverage and recovery requirements.

Compilation across physics

A mature reflex can also be stored in geometry, compliance, an analog transfer function, a physical reservoir, or a reprogrammable material state. Demonstrated components include passive task-specific dynamics (C-112), soft-body and physical-reservoir memory (C-113, C-114), mechanical logic and physical learning (C-115, C-116), and local material repair (C-120). These observations extend the set of possible deployment substrates; they do not bypass the qualification gates.

A physical pattern is classified before it is credited with computation:

CLASS	WHAT CHANGES STATE	EXTERNAL BOUNDARY	LEARNING CLAIM
passive	relaxation or fluctuation under fixed energy landscape and boundaries	preparation, geometry, thermal/gravity/elastic conditions	none
driven fixed	continuous fuel, light, field, vibration, gravity, or flow under fixed interactions	all drive generation and dissipation	none
feedback controlled	explicit observation is converted into intervention	sensor, compute, communication, actuator, and drive	none unless the policy updates
adaptive	outcome evidence changes policy, interaction, morphology, or objective	all preceding costs plus training, validation, memory, and reset	testable

This separation follows C-463–C-479. Flocking, phase separation, defect motion, jamming, or assembly can be useful physical operations; visible order alone does not supply a task, evaluator, or policy update.

The held systems candidate is a rewritable physical path whose input and output remain locally coupled to the environment, while a versioned digital shadow preserves its specification, protected tests, calibration envelope, fallback, and output trace. The path is admitted only while health probes remain inside that envelope. Drift or damage returns authority to the digital path before a new substrate state is programmed and shadow-validated. Static surface texture must therefore survive workload-envelope, manufacture, inflexibility, and reversal tests (C-1503); a locally adaptive interface additionally carries constituent inventory, activation/recovery, hysteresis, depletion, substrate health, and fallback (C-1504).

This path competes first with tuned passive mechanics, analog control, and FPGA/ASIC implementation, not only with an inefficient general model. Its lifecycle boundary includes design, fabrication, programming, drive, conversion, readout, reset, calibration, maintenance, fallback, repair, failed devices, and retirement. Candidate 006 tests whether a measured conversion, transport, recurrence, or command path is actually removed and whether break-even occurs before the qualified substrate lifetime ends.

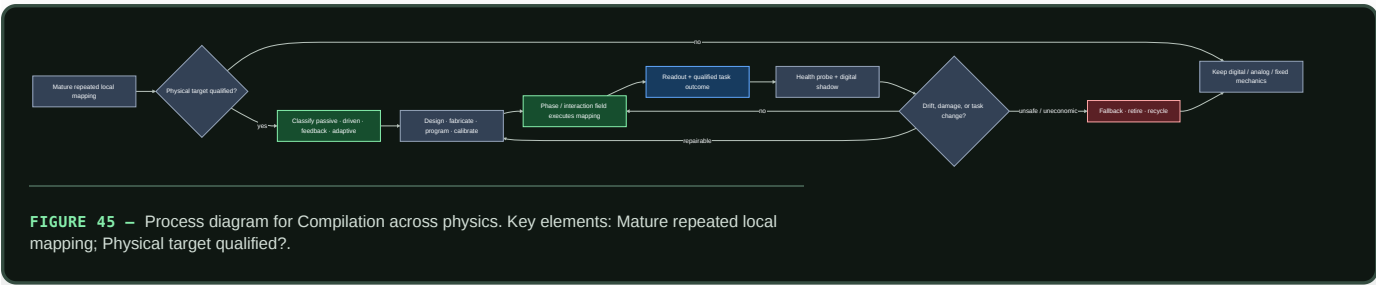


FIGURE 45 – Process diagram for Compilation across physics. Key elements: Mature repeated local mapping; Physical target qualified?.

Editable source: phase-field-compilation.mmd.

Candidate 006 now tests C-480 as one physical- phase specialization. Particle-scale power cannot win the comparison; the boundary includes fuel or field generation, sensing, control, transduction, readout, reset, fabrication yield, health probes, shadow, fallback, and retirement.

Reusable skills

A reusable skill is broader than a reflex. It accepts variable inputs, may consult current context or memory, and is evaluated for transfer outside the episodes that created it. It normally remains a slow-model module, adapter, tool policy, or callable subgraph.

Skill qualification asks whether the artifact preserves a reusable relation or operation. A proposition such as a price, office holder, software version, or medical recommendation is not a skill: its truth can change while the method used to retrieve, compare, or explain it remains valid. If a skill emits a mutable proposition, the output record must identify which factual version supplied it.

A skill may later receive structured pruning or quantization. Iterative pruning is evidence for competitive sparse subnetworks only in its tested settings (C-012). Ternary-weight language models are a plausible candidate under C-013; promotion still depends on end-to-end quality, risk, latency, bytes, and joules on the project workload.

Qualification gates

Promotion is conjunctive: a candidate fails when any hard gate fails.

GATE	REQUIRED RECORD	REJECT WHEN
Semantic class	reflex, skill, or factual record with one owner	the artifact mixes a stable transform with independently mutable claims
Applicability	schemas, $\mathcal{A}_h$, guard G_h , dependency versions	the guard cannot abstain before an out-of-envelope execution
Quality and risk	metrics by common, rare, safety, and shift stratum	average quality hides a stratum outside its tolerance
Causal contribution	ablation, reroute, and ordinary-code comparison	a cache, rule, smaller model, or router explains the gain
Physical accounting	guard, dispatch, execution, movement, idle, build, validation, and recovery costs	savings exist only in FLOPs or omit lifecycle work
Reversibility	immutable candidate version, atomic route switch, last-known-good target, rollback drill	prior behavior cannot be restored inside the recovery envelope
Provenance	source episodes, code/data versions, tests, approver, and artifact digest	the artifact or its qualification result cannot be reconstructed

For candidate path h and baseline b , let Q_h and Q_b be task quality in one declared unit, R_h and R_b be risk in one declared unit, and ϵ_Q and ϵ_R be preregistered tolerances in those respective units. Let $pFA_{h,h}$ be the dimensionless fraction of out-of-envelope events incorrectly admitted by G_h , and let αFA be its maximum permitted value. Qualification requires

$$Q_h \geq Q_b - \epsilon_Q, \quad R_h \leq R_b + \epsilon_R, \quad pFA_{h,h} \leq \alpha FA.$$

These conditions are evaluated by stratum as well as in aggregate. A guard that rejects almost everything is exposed by reporting its coverage $c_h = N_{G_h=1}/N_{\text{offered}}$, a dimensionless fraction. Coverage is a result, not a target inferred after testing; $N_{G_h=1}$ and N_{offered} are event counts.

Reversible verification before commitment

Kinetic proofreading shows that recognition and commitment can be separated by driven intermediate states with discriminatory rejection and reset (C-159). Its speed, error, and dissipation costs form a model-specific frontier rather than a universal accuracy multiplier (C-160). The systems translation therefore has four hard requirements:

1. temporary execution remains inside a declared rollback boundary;
2. the later verifier adds conditional information or a distinct detector;
3. rejected attempts, reset, delay, and provenance remain in the cost ledger;
4. irreversible authority is withheld until commitment.

For observations $z_1, \dots, z_t$, the strongest statistical null conditions on the evidence already seen:

$$L_t = \sum_{i=1}^t \log \frac{p(z_i \mid R, z_{<i})}{p(z_i \mid W, z_{<i})}.$$

Here L_t is the dimensionless cumulative log-likelihood ratio, R and W denote correct/safe and wrong/unsafe hypotheses, and $z_{<i}$ is prior evidence. Ignoring that conditioning turns correlated rechecks into false confidence.

Candidate 010 tests reversible execution and risk-conditioned verification against this sequential test, calibrated cascades, abstention, retries, redundant verifiers, and error-detecting codes. It must tie or lose when the later stage is only a correlated copy or when reset leaks irreversible effects.

Graded assurance envelopes

Qualification records must state what kind of assurance each result provides. The classes are not interchangeable:

ASSURANCE CLASS	SUPPORTS	DOES NOT ESTABLISH
type, refinement, or proof	a named property under declared semantics and trusted base	termination, unspecified behavior, security, task quality, or truth
effect description	operations the model may perform under the analysis	authority to perform them or their correctness
capability grant	enforced authority inside a complete-mediation boundary	intent, competence, or safe outcome
empirical evaluation	behavior on declared data, environment, slices, and uncertainty	untested distributions or future versions
runtime monitor	a verdict over observed events under one temporal formula	unobserved channels or arbitrary future behavior
provenance	artifact identity and derivation path	source truth or claim entailment
transaction or compensation	recovery of participating state or a declared compensating action	reversal of time, disclosure, physical effects, or third-party actions

These boundaries are established in scoped programming-language and systems results: type soundness (C-145), effects versus capabilities (C-148), runtime-monitor scope (C-152), transactional rollback limits (C-154), and provenance without truth (C-156).

The held synthesis binds every assurance record to the same module version, artifact digest, dependency graph, state migration, authority policy, monitor schema, evidence set, and invalidation triggers. A dependency change rechecks only its affected cone, but stale assurance escaping to production and unnecessary rechecks are both measured. Candidate 009 compares this envelope against a complete conventional stack of typed APIs, sandbox/IAM, CI and static analysis, runtime policy monitoring, lineage, canaries, transactions, schema migration, and build-system invalidation.

Compromise-bounded authority and recovery

The security contract adds an adversary and trust boundary without collapsing distinct stages. Authentication establishes a scoped protocol property; authorization grants an action; detection classifies telemetry; containment blocks covered future use; and recovery re-establishes declared invariants from a tested root. None substitutes for the next (C-250, C-262, C-265).

For capability class j , let g_j and r_j be its grant and effective revocation times in seconds, and let w_j be a declared dimensionless severity weight. Authority exposure is

$$X_A = \sum_j w_j \max(0, r_j - g_j),$$

with unit weighted-capability-seconds. The weights and individual intervals remain visible because a single broad destructive capability is not equivalent to many harmless reads. Sensitivity to plausible weights is reported.

Nominal credential lifetime is not the revocation result. If t_{comp} is the bounded compromise time and t_{last} is the last acceptance at every covered enforcement point, then

$$W_{\text{rev}} = \max(0, t_{\text{last}} - t_{\text{comp}})$$

is revocation exposure in seconds. Sessions, caches, delegation, offline verifiers, propagation delay, clock rollback, and missing acknowledgements are part of the measurement (C-260). The incident record keeps four clocks separately: compromise interval, detection, effective

containment, and independently validated recovery.

The held profile binds principal and workload identity, capability scope, credential/key/attestation epoch, revocation freshness, approval-domain independence, observation age, adversary model, compromise horizon, and clean- root evidence to the same versioned artifact. It survives only if Candidates 009 and 012 reduce harm or secure recovery time beyond mature short-lived IAM and a rehearsed reimage–rotate–validate workflow at equal lifecycle cost.

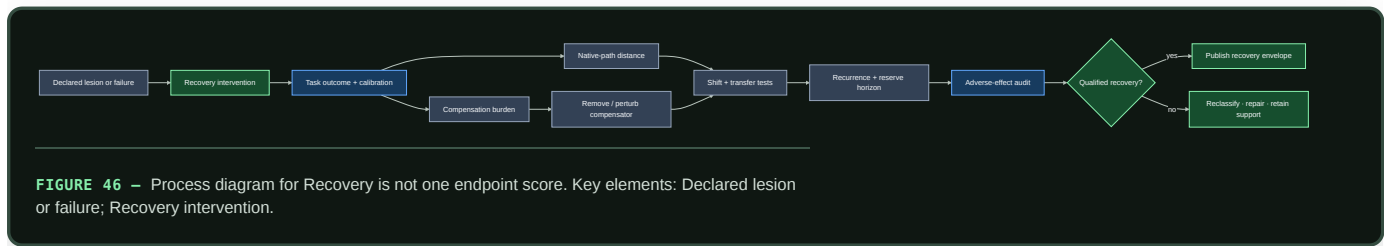
Recovery is not one endpoint score

After a lesion or fault, the same task outcome can come from restored capability, a larger fallback, a different route, a tool, a cache, human intervention, or a more permissive environment. Rehabilitation evidence makes that underidentification explicit (C-316–C-324).

Let z be a declared post-recovery internal, routing, or behavioral state and z_{ref} a justified reference. Report

$$D_{\text{native}} = d(z, z_{\text{ref}}), \quad \mathbf{B}_{\text{comp}} = (E_{\text{extra}}, M_{\text{extra}}, L_{\text{extra}}, H_{\text{extra}}, P_{\text{fragile}}),$$

where D_{native} has the declared metric's unit or is explicitly normalized; extra energy is joules; memory is bytes; latency and human support are seconds; and P_{fragile} is the dimensionless probability of failure under a named support-removal or shift test. These axes are not merged unless conversion weights are declared in advance.



Editable source: [compensation-aware-recovery.mmd](#).

Compensation is not a failure when it is robust, affordable, and declared. The failure is to call it restoration, hide recurring support, or deploy outside the context that makes it work. Conversely, an ordinary cheap workaround may dominate native restoration. Candidate 005 therefore compares the complete accounting method with conventional multi-objective robustness, fault injection, failover, checkpoint restoration, and functional reconstruction.

Versioned factual memory

A factual record r contains at least

$$r = (k, v, u, s, v_s, t_{\text{obs}}, [t_{\text{from}}, t_{\text{to}}], v_r, \pi, \sigma),$$

where:

- k is a typed key or subject–predicate identifier;
- v is the typed value and u is its declared unit, or `none` for a unitless value;
- s is the source identifier and v_s its source version;
- t_{obs} is the observation timestamp;
- $[t_{\text{from}}, t_{\text{to}})$ is the asserted validity interval;
- v_r is the immutable record version;
- π is the access, retention, and jurisdiction policy; and
- $\sigma \in \{\text{active}, \text{superseded}, \text{revoked}, \text{disputed}\}$ is record status.

Timestamps use UTC with declared resolution. Differences between timestamps are reported in seconds. A record may also carry source-supplied confidence or a calibrated probability, always dimensionless and never substituted for source identity or conflict handling.

Let t_{check} be the last successful source check and let τ_k be the maximum unchecked age for key class k in seconds. At query time t , the freshness gate is

$$F(r, t) = 1[\sigma = \text{active}]1[t_{\text{from}} \leq t < t_{\text{to}}]1[0 \leq t - t_{\text{check}} \leq \tau_k].$$

$F(r, t)$ is dimensionless. The domain policy fixes τ_k before evaluation; an unbounded value is permitted only when the domain explicitly defines the record as non-expiring. Freshness means that the record passed its time and status policy, not that its proposition is correct.

For a key k , the conflict set $\mathcal{C}_k(t)$ contains active, policy- admissible records whose values cannot simultaneously hold at time t . Resolution may use an explicit source-authority rule, a time rule, or a domain-specific adjudicator. The system preserves losing records and the resolution trace. When no preregistered rule applies, the retrieval path abstains and escalates rather than averaging incompatible values.

Retrieval-augmented generation establishes that parametric generation can be combined with inspectable and replaceable non-parametric memory on evaluated knowledge-intensive tasks (C-014). This chapter adds version, freshness, conflict, and lifecycle accounting as requirements to test, not as evidence that retrieval is automatically correct.

Invalidation and rollback

Invalidation is triggered by any of the following observable events:

- an input-schema, dependency, tool, hardware, or source version changes;
- a validity interval or freshness allowance expires;
- a protected regression, shift probe, calibration check, or outcome fails;
- an authoritative source revokes or supersedes a record;
- a new admissible record creates an unresolved conflict; or
- guard false admissions, fallbacks, or escalations exceed their declared control limits.

The response depends on the artifact:

ARTIFACT	IMMEDIATE ACTION	DURABLE ACTION
Compiled reflex	atomically route new events to the last-known-good path	preserve failed binary, manifest, traces, and invalidation cause
Reusable skill	freeze the active version and open a copy-on-write branch	replay, validate, reconsolidate, or retire through maintenance
Factual record	remove the version from active retrieval and append a tombstone or dispute edge	retain prior values and rebuild an index version without destructive overwrite
Escalation policy	fall back to the conservative route and cap further delegated work	recalibrate on logged false admission, miss, cost, and outcome data

Every served output records the path, model, dependency, and factual versions that affected it. This makes the affected-output set enumerable after a defect. Report rollback time T_{rb} in seconds, rollback energy E_{rb} in joules at the declared boundary, lost or corrected events N_{loss} as a count, and restoration success as a dimensionless fraction.

Severity-ordered containment and triage

Rollback is one response to one fault contract. A stateful modular system also needs to keep the following actions distinct:

1. **Sense:** collect health evidence without treating a detector output as fault ground truth.
2. **Contain:** throttle inputs, revoke an interface, quarantine a route, or freeze a version to limit spread before diagnosis completes.

3. **Triage:** choose retry, local repair, selective reconstruction, restart, retirement, or replacement from the available evidence and declared cost.
4. **Verify:** test the affected behavior, protected rare behavior, provenance, and adjacent modules before restoring authority.
5. **Escalate:** move to a more destructive action only when verification fails or stronger evidence makes delay unsafe.
6. **Replenish:** restore validated capacity when a component is retired, rather than allowing maintenance to clean the system into capacity collapse.

The biological audit supplies scoped examples of fast load shedding (C-087), repair-versus-degradation triage (C-090), tag-dependent compartment routing (C-091), selective extraction (C-092), repair before removal (C-094), and removal coupled to replacement (C-095, C-096). The order is conditional: a rapidly spreading irrecoverable fault may require immediate replacement, while a local reversible fault should not trigger a global rebuild.

The engineering value of the composition remains unproven. Circuit breakers, taint tracking, checkpoints, replica failover, microreboots, scrubbing, rejuvenation, and Bayesian repair/replace policies are mandatory comparators. [Candidate 005](#) tests the staged policy across locality, observability, repairability, and correlated detector error while charging sensing, reserve, copying, replacement, verification, downtime, and collateral loss.

Escalation is a budgeted route

Escalation activates when the local guard abstains, required facts are stale or conflicted, the event enters a protected risk stratum, or the available path cannot meet its quality contract. The escalation target may be a larger model, a deterministic tool, an authoritative data source, or a human.

The request carries the triggering uncertainty, attempted path and versions, relevant evidence, allowed data disclosure, deadline, and remaining energy or financial budget. Its result returns with provenance and observed cost. Escalation does not grant write authority; durable change still requires a versioned maintenance action.

Efficiency mechanism

Per-event and lifecycle accounting

For path version h serving N qualified events, define amortized energy

$$e_h(N) = e_{\text{guard}} + e_{\text{dispatch}} + e_{\text{execute},h} + \frac{E_{\text{build},h} + E_{\text{validate},h} + E_{\text{migrate},h} + \mathbb{E}[E_{\text{rollback},h}]}{N}.$$

Lowercase e terms are measured joules per qualified event. Capital E terms are one-time joules at the same device, node, cluster, or facility boundary; $\mathbb{E}[E_{\text{rollback},h}]$ includes failed promotions weighted by their observed or preregistered probability. N is a dimensionless event count, so $e_h(N)$ is joules per qualified event. If the deployment horizon is unknown, report the break-even event count rather than assuming amortization.

For event x , end-to-end latency is

$$\ell_h(x) = \ell_{\text{guard}}(x) + \ell_{\text{dispatch}}(x) + \ell_{\text{execute},h}(x),$$

with every ℓ term in seconds. Report the empirical end-to-end percentile $L_{h,p}$ for a declared percentile $p \in (0, 1)$ rather than adding separately measured component percentiles. Report physical traffic B_h in bytes per qualified event across each named boundary. Energy, latency, and bytes remain separate results.

The execution term expands by path:

PATH	COSTS THAT MUST BE VISIBLE
Compiled reflex	guard, dispatch, code and state loads, kernel execution, cache residency, precision conversion
Reusable skill	routing, parameter and activation movement, inference, memory/tool calls, synchronization
Factual memory	query encoding, index access, retrieval, reranking, freshness/conflict checks, evidence bytes, composition
Escalation	failed local attempt, serialization, network, remote or tool execution, waiting time; human time reported separately
Rollback reserve	retained versions, manifests, index generations, route switch, replay, correction, recovery validation

A candidate advances only when quality and risk remain qualified and it improves the preregistered energy–latency–traffic frontier after all listed costs are included. Skipped model operations alone are insufficient.

Provenance and freshness measurements

For N_f outputs that use factual memory and N_p of those outputs carrying a complete record-to-source trace, provenance coverage is

$$C_{\text{prov}} = \frac{N_p}{N_f}.$$

C_{prov} is dimensionless. Also report source-check age in seconds, stale-use and unresolved-conflict rates as fractions, index and evidence bytes, joules per lookup and update, and p50/p95 lookup latency in seconds.

If a source changes at t_{source} and the first correctly served version is available at t_{serve} , correction latency is

$$T_{\text{corr}} = t_{\text{serve}} - t_{\text{source}}$$

in seconds. Record source polling or event-delivery cost alongside this value; instant correction purchased by continuous high-cost polling is not free.

Strongest null models

All nulls receive the same task stream, factual sources, safety policy, hardware opportunity, and lifecycle horizon.

ID	NULL	REQUIRED COMPARISON
N0	full capable model on every event	tests whether any dispatch hierarchy beats unconditional execution
N1	calibrated early exit or static small/large cascade	tests ordinary adaptive depth under C-004
N2	conventional rules engine, compiler, or memoization cache with TTL	tests whether a compiled reflex adds more than established software practice
N3	separately trained smaller, distilled, or quantized model	tests whether path specialization beats a simpler fixed deployment
N4	parametric-only knowledge with scheduled fine-tuning	tests factual correction cost, carryover, and provenance
N5	strong hybrid-search RAG with reranking, citations, and ordinary freshness filters	tests whether version/conflict machinery improves the factual frontier
N6	in-place skill update with checkpoint restore	tests copy-on-write promotion and rollback overhead against standard recovery
N7	calibrated risk/confidence threshold using the same escalation target	tests whether the structured guard adds value beyond a scalar threshold
N8	encoded redundancy, replica/quorum recovery, integrity scrub, or self-stabilizing legitimate-set repair	tests whether established exact or rule-encoded reconstruction dominates learned recovery under the declared fault model
N9	circuit breaker, static isolation, microrestart, rejuvenation, or Bayesian repair/replace policy	tests whether ordered containment and triage add value beyond mature fault-management composition

If ordinary code or a standard data system matches the candidate, keep the conventional mechanism and retain only the qualification and accounting contract.

Evidence status

MECHANISM	EVIDENCE	STATUS FOR THIS CHAPTER
adaptive early exit	C-004	established on evaluated BERT tasks; shift and rare-risk gating open
fast acquisition versus slow integration	C-008	plausible architectural separation; exact tiers unvalidated
structured consolidation and pruning	C-012	established in scoped experiments; physical saving not automatic
ternary-weight models	C-013	plausible preprint, technical-report, and open-kernel evidence; matched project workload and measured lifecycle frontier remain unresolved
parametric plus non-parametric memory	C-014	established on evaluated RAG tasks; freshness and conflict remain open
retrieval-sensitive updating	C-039, C-040	scoped observation; precise general mismatch gate disputed
reversible mature constraint	C-044, C-045	scoped biological intervention; digital rollback contract experimental
exact restore, encoded repair, failure detection, and integrity scrubbing	C-079–C-085	established engineering nulls under explicit fault and cost models
constraint-guided functional reconstruction	C-086	speculative residual candidate for semantic or capability loss without a clean exact state
load shedding, tagged routing, selective extraction, repair/removal ordering, and replacement feedback	C-087–C-096	scoped cellular mechanisms; the composed systems policy remains a held candidate
morphology, physical reservoirs, mechanical memory/logic, local assembly, and material healing	C-112–C-120	established in scoped substrates; end-to-end advantage over passive, analog, FPGA/ASIC, and digital nulls remains workload-specific
reversible physical skill compilation	C-121	speculative lifecycle systems hypothesis tested by Candidate 006
soft/active-matter physical order and external feedback	C-463–C-479	established or plausible in scoped physical systems; no audited task-success learning
phase/interaction-field skill compilation	C-480	speculative Candidate-006 specialization
types, contracts, effects, capabilities, proof checking, static analysis, runtime monitoring, transactions, hot update, and provenance	C-145–C-156	established scoped assurance classes with explicit trusted bases and invalidation boundaries
versioned graded assurance envelopes	C-157	speculative systems composition tested by Candidate 009
automatic reflex discovery and qualification	none	speculative until it beats the null models above

Speculative extensions

- Train a compiler to propose a small executable graph, applicability predicate, counterexample set, and rollback manifest as one candidate artifact.
- Learn source-check schedules from update hazard and consequence while keeping hard maximum ages for protected domains.
- Compile stable *relations* learned across factual versions while leaving the current proposition external and attributable.
- Use signed provenance graphs so a correction can enumerate dependent outputs, cached artifacts, and downstream derived records.
- Co-design path placement and precision with hardware only after the logical qualification gate passes on a conventional substrate.

Failure modes

FAILURE SIGNATURE	OBSERVABLE MEASURE	REQUIRED RESPONSE
guard admits shifted input	$pFA_{,h}$ rises by stratum	disable version; route to fallback; add counterexample
guard rejects most valid traffic	coverage c_h collapses while aggregate quality looks stable	report lost coverage; compare with N1/N7
shortcut becomes a reflex	protected counterfactual or subgroup regression	invalidate and reopen source skill
volatile fact leaks into parameters or compiled code	correction requires retraining or old value persists after record update	demote proposition to factual memory; trace affected outputs
fresh but wrong source record	outcome error despite $F(r, t) = 1$	preserve source trace; dispute/revoke; strengthen authority policy

FAILURE SIGNATURE	OBSERVABLE MEASURE	REQUIRED RESPONSE
unresolved contradiction is silently merged	non-empty $\mathcal{C}_k(t)$ without a resolution trace	abstain and escalate
index and model disagree on active version	served $v_{\mathcal{T}}$ differs from index-generation manifest	atomically roll back index and replay affected queries
promotion thrashes	repeated compile–invalidate cycles, migrations, or route flips	raise evidence horizon; charge churn energy; retain conventional path
quantization hides a rare regression	mean quality holds while protected-stratum R_h rises	reject precision change
rollback is nominal only	T_{rb} , N_{loss} , or recovery tests exceed envelope	block future promotion until recovery is repaired
retrieval savings vanish physically	operations fall while bytes, latency, or joules do not	reject efficiency claim; keep stronger null
escalation becomes an unpriced default	escalation fraction and remote cost rise without risk improvement	recalibrate guard and expose full route cost
ordered phase is credited as intelligence or learning	no independent task signal, evaluator, or outcome-driven update exists	relabel as passive/driven physics and retain it only as a substrate/null
particle-scale power looks excellent while wall energy rises	drive, sensing, compute, transduction, reset, or fabrication dominates	reject the system-efficiency claim and restore the full boundary

Measurable predictions

1. For recurrent, stable, low-risk transformations above a measurable break-even horizon, a qualified compiled path reduces joules per event or end-to-end latency relative to N0–N3 while preserving stratum-level quality, risk, and guard false-admission limits.
2. Reusable skills outperform memoization and rules on held-out compositions, while reflex paths outperform the skill only inside their narrower declared envelope.
3. Separating mutable propositions from skills reduces correction latency, retraining energy, and obsolete-value carryover relative to parametric-only N4 at matched task quality.
4. Version, freshness, and conflict gates reduce stale or contradictory factual use relative to strong N5; the result is rejected if lookup cost erases the quality–risk benefit.
5. A structured applicability guard reduces out-of-envelope false admissions under shift relative to N1 and N7 without achieving the result by rejecting nearly all traffic.
6. Complete provenance increases the fraction of factual outputs whose source and dependent artifacts can be enumerated after a correction, at a measured byte, latency, and energy cost.
7. Atomic route rollback and versioned factual tombstones reduce restoration time and lost events relative to in-place N6 after injected bad promotions and source revocations.
8. Quantized or compiled paths produce an end-to-end physical gain only when dispatch, memory movement, validation, retained rollback state, and failed promotions amortize within the observed deployment horizon.
9. A reversible phase or interaction field advances only if it removes a measured repeated digital/transport path and beats passive mechanics, analog control, FPGA/ASIC, and distributed digital control after complete drive, fabrication, sensing, reset, drift, fallback, and retirement cost.

System synthesis

Intelligence is the coordinated movement of state through perception, prediction, action, memory, adaptation, and maintenance under a finite physical budget.

Scope

The working architecture defines the runtime, adaptation, and maintenance loops. This chapter follows information through those loops, identifies who may write each kind of state, and gives the order in which the integrated system should be assembled.

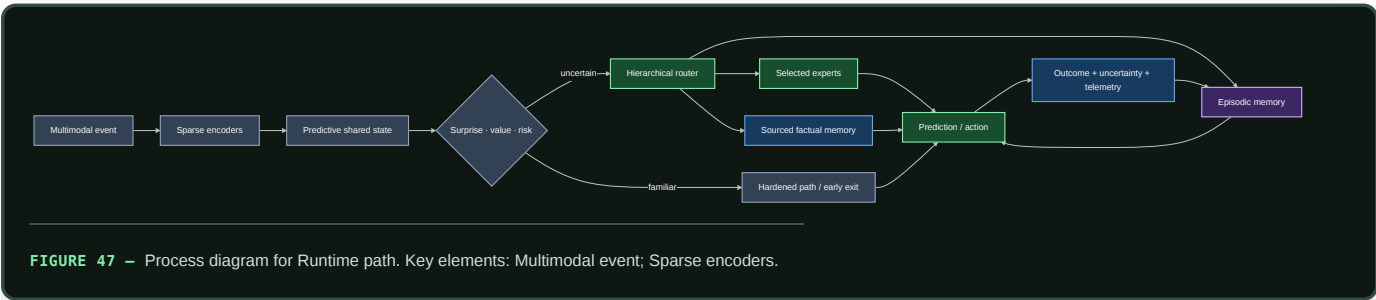
Biological observation

Living intelligence coordinates sensing, action, memory, plasticity, resource supply, and repair across different timescales. Local circuits handle repeated work, broader signals alter priorities, fast traces change behavior before slow structure moves, and maintenance continues during operation.

The transferable principle is **coordinated state ownership**. Reading a result does not grant authority to overwrite its source. Fast behavior, rapid learning, slow consolidation, and physical resource control can therefore interact without becoming one global update rule.

Proposed AI translation

Runtime path



Editable source: `../assets/diagrams/system-runtime.mmd`.

The path has six stages:

- 1. Ground the event.** Encoders align perception with current action, timing, location, and tool state.
- 2. Predict before expanding computation.** The shared state estimates what is likely, what is uncertain, and which goal currently matters.

- 3. **Choose the execution path.** A gate may dispatch a hardened skill, stop at an early exit, retrieve memory, activate selected experts, intervene through a tool, or escalate.
- 4. **Execute only admitted work.** Inactive capacity remains addressable without being loaded and operated on for the event.
- 5. **Measure the result.** Output is paired with observed consequence, calibration, latency, bytes moved, and energy.
- 6. **Capture an episode.** The adaptation loop receives an attributable record rather than an unexplained parameter change.

The runtime preserves both relative and absolute authority

The predictive state may expose absolute level, additive change, derivative, ratio, reference age and support as different fields. A router does not choose one of them from a modality or pathway name. It chooses from the registered task and observation interface: relative sensing can transfer across a multiplicative nuisance scale, while absolute load, dose, damage and capacity remain on calibrated channels.

Every relative route therefore carries the reference identity and age that made it valid, plus a fallback when the input approaches zero, changes sign, clips, saturates or leaves the learned scale support. Removing the absolute channel is a reversible ablation, not a default compression. The robustness--recoverability trade is tested explicitly in RSD-T10 and derived in Interface-qualified scale symmetry.

A subscriber can change its producer

State ownership prevents unauthorized writes, but it does not guarantee that a read or subscription is dynamically neutral. A consumer can pin a producer's slot, hold a reference, block publication, create a gradient path, compete for shared resources, or intentionally feed back. Those cases are distinct from an immutable read and from unrelated system contention (C-1550, C-1558).

Every reusable module therefore carries a connection-sensitivity record:

- 1. isolated input, state and output trajectory;
- 2. attached client identity, topology, coupling and holding time;
- 3. producer deviation at the producer-owned interface;
- 4. delivered, timely and fresh consumer service;
- 5. shared CPU, memory, queue, network and accelerator load;
- 6. insulation copies, buffers, admission, monitoring and recovery work; and
- 7. the support in which the connection remains qualified.

The system inserts no insulator when the immutable/no-back-action control passes. Where material back-action exists, it selects among ordinary buffering, snapshots, weak coupling, resource isolation, replication, explicit feedback and a bounded adaptive insulator. Suppression is not the only admissible goal: registered useful feedback or temporal shaping must remain intact (C-1559). The complete decision is tested in Fixture F-027, not inferred from the biological name.

Authority is split across control planes

The system uses four control planes with different authority:

PLANE	CONTROLS	READS	CANNOT DO ALONE
Task	goals, quality, permitted actions	predictive state, user/tool feedback	hide physical or risk cost
Resource	energy, latency, communication, placement	estimates and measured telemetry	redefine task success
Adaptation	episodes, hypotheses, provisional modules	outcomes, conflict, uncertainty	promote a global slow-model change
Maintenance	replay, merge, weakening, deletion, topology	history, regressions, fragility, lifecycle cost	bypass provenance or release gates

The planes exchange compact declared state. A resource controller can deny an expensive route but cannot declare the cheaper result equally correct. A task controller can request deeper computation but cannot conceal the resulting traffic from the energy ledger. Adaptation can propose; only maintenance can promote a change into protected state. A local corrective action also cannot hide the shared-system cost it exports when the same trigger becomes widespread (C-1493).

A typed-specialist controller remains a held research arm

One speculative implementation makes the coordinating loop explicit without putting a general language model in charge. The controller would own run state, route selection, specialist and memory access, tool authority, budgets, deadlines, candidate arbitration, final output or action, and abstention. Each specialist would receive a bounded typed request and return a typed candidate, calibrated uncertainty or abstention, provenance, and a resource receipt. A specialist could be a small neural model, estimator, solver, conventional program, or tool. A language model would be one optional linguistic specialist, not the controller or an assumed system-wide brain.

Capability boundaries would come from task contracts and demonstrated causal contribution. Howard Gardner's eight multiple-intelligences categories are a possible omission prompt for evaluation design, not eight model slots. Their independence remains empirically contested, and a human ability label does not establish a neural or artificial module boundary. The bounded source audit preserves the proposal, contrary evidence, measurement dispute, and exact limits of that analogy.

The development controller now makes this boundary executable for the frozen CLRS shakedown. Every registered specialist needs a timestamped observation with a finite validity window. Fit distinguishes `measured-fit`, `known-no-fit`, and `unknown`; `unknown` is not a match. A `measured-fit` observation must bind a caller-owned basis identifier, measurement time, and bounded expiry, after which the controller treats it as `unknown`. The local CLRS adapters use an additional construction-only `task-compatible` state rather than claim a per-request resource measurement. Policy and adapter limits still validate each packet. Each frozen task accepts at most two route candidates. Route selection and a one-shot, request-bound capacity reservation occur together. Per-specialist and aggregate queue and active-work limits bound the controller alongside its wait and readiness-retry limits. Eligible waiters are capacity-matched in FIFO order across overlapping routes, so later work can use spare capacity without taking a slot an older waiter can use. It records typed fallback or rejection, rechecks readiness, cancellation and the request deadline inside the serialised state immediately before invocation, and uses specialist identity as the stable tie-break when declared cost and readiness are equal. An otherwise positive recheck must match the recorded readiness and fit evidence; a newer observation produces an explicit fallback instead of authorising an effect under the older binding. This is a construction check of controller mechanics. It neither measures an efficiency gain nor grants the resulting candidate scientific authority; every path remains `NO_RESULT` pending the matched comparisons described below.

The arm must face a capacity-matched general model and a tuned sparse mixture of experts with the same tools, state, training opportunity, routing information, hardware, and outcome access. Count controller work, serialization, repeated context, communication, memory traffic, cold starts, calibration, verification, retries, failed routes, maintenance, replacement, latency tails, and measured energy. If an ordinary controller or router reaches the same quality--risk--latency--energy frontier, the extra specialist boundary leaves the architecture.

Evaluation need not demand that one complete architecture win every reported axis. Before outcomes are opened, define task-quality and protected-risk floors, comparators, uncertainty analysis, and the axes on which a component may earn retention. Keep a component only when an independent ablation shows a reproducible Pareto contribution without crossing those floors; an energy reduction is particularly useful, but cannot buy lower declared quality or hidden risk. Report every regression and interaction because separately useful components need not compose additively.

A non-winning component moves to a versioned retest backlog with its negative result, failed boundary, and strongest comparator. It is not rerun on a timer. Reopening requires a predeclared concrete trigger: a stronger mechanism, a changed workload or physical boundary, or new evidence that changes the tested prediction. The retest then uses the current stronger baseline rather than the one the component previously lost to. This portfolio rule is itself a research hypothesis under OQ-068, not an accepted architecture decision.

Governance is activated by a real authority problem

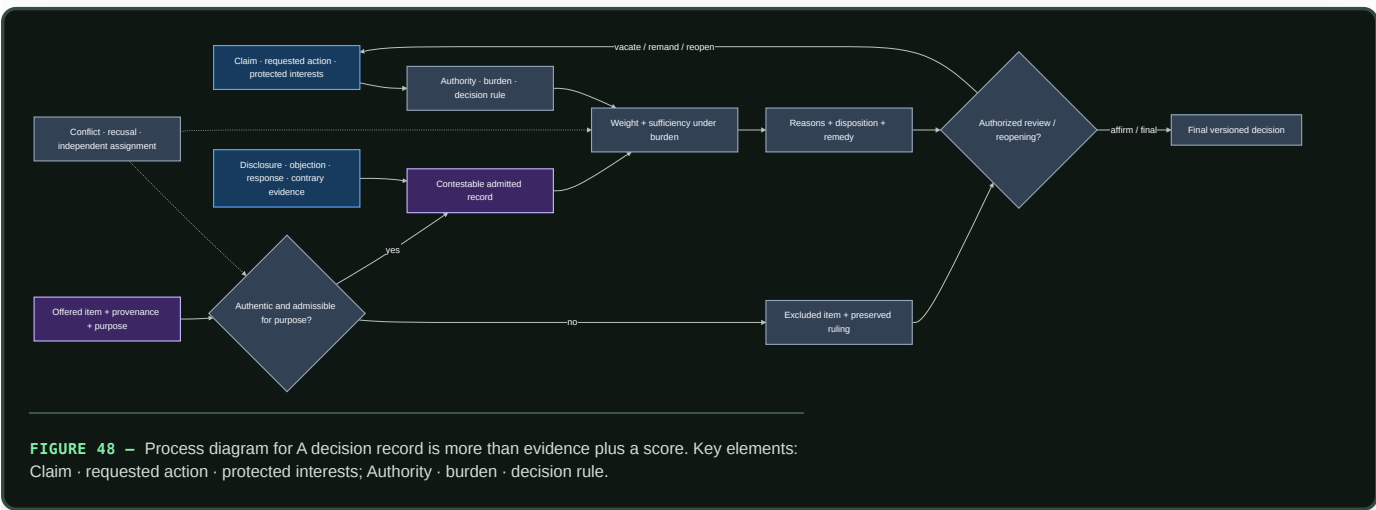
Aggregation rules, delegation, vetoes, constitutions, and participation are not useful decorations for ordinary routing. They enter the architecture only when persistent actors have decision-relevant private information, can benefit from misleading reports, create spillovers across authority boundaries, hold externally authorized protected standing, or must repair higher-order rules without exposing them to ordinary updates.

When that applicability gate passes, the control plane records more than a winning action:

- 1. the authorized objective and non-tradable invariants;
- 2. the actor's identity, scope, evidence, dependencies, and conflicts;
- 3. the proposal set, order, fallback, and omitted-alternative coverage;
- 4. every delegation edge, concentration measure, revocation, and cycle;
- 5. a veto's protected scope, severity, evidence, deadline, appeal, override, and expiry;
- 6. the decision and affected local and external interests; and
- 7. the rule version, amendment path, emergency authority, checked handoff, compatibility result, and rollback target.

A decision record is more than evidence plus a score

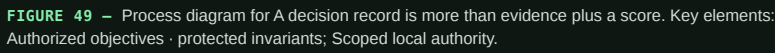
Where an action is contestable and authority-bearing, the system keeps normative authority, empirical evidence, formal inference, and procedural validity as separate predicates. Authenticity, admissibility for a stated purpose, weight, sufficiency under a burden, authorization, review, remedy, and finality are not ranks on one confidence scale.



Editable source: [burden-qualified-decision.mmd](#).

The record binds claim/action, authority, affected interests, proponent, burden, evidence purpose and provenance, disclosure/access, objections, contrary evidence, findings, reasons, decision, protected outcomes, review scope, remedy, successor version, and reopening conditions. It records excluded material so later review can test contamination without silently letting that material influence the decision.

The contestable-decision mathematics keeps error, delay, review, access, human work, energy, and protected procedure as raw axes. Its evidence boundaries are C-679–C-704. The composition must beat typed workflow, provenance, selective prediction, access control, rule graphs, independent review, red-team challenge, and full recomputation. It refines existing assurance, verification, operational learning, observation, convention, and governance candidates; it does not add a new control primitive.

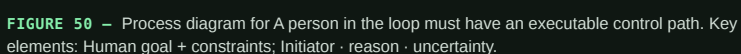


The formal and empirical boundaries in C-368–C-395 forbid several shortcuts: an impossibility theorem does not select a moral objective; stability is not quality; participation is not legitimacy; popular or expert influence is not automatically capture; and an amendment count is not repair. Normative standing and non-tradable harms remain explicit authorized inputs.

A person in the loop must have an executable control path

Each machine-initiated or shared-control transition can therefore carry a recoverable initiative record:

1. initiator, reason, urgency, and calibrated uncertainty;
2. actual mode, pending mode, enabled effects, and response deadline;
3. effective human and machine authority, including silent arbitration;
4. intended action, observed state delta, and provenance;
5. acknowledgement or rejection and whether it changed execution;
6. checkpoint, undo, compensation, or safe-stop boundary;
7. resumption cue, pending goals, and assumptions that may have gone stale;
8. verified recovery result, expiry, and longitudinal-learning link; and
9. accessibility and assistive-technology conditions under which the path was tested.



Human time is part of lifecycle cost. For interruptions $j = 1, \dots, n$,

$$T_{\text{complete}} = T_{\text{active}} + \sum_{j=1}^n (T_{\text{interrupt},j} + T_{\text{resume},j} + T_{\text{rework},j}),$$

with every term measured in seconds under a declared coding rule. Error severity, stress, frustration, workload, training, and person-minutes remain separate outcomes; throughput alone cannot erase them.

The bounded evidence in C-396–C-416 supports specific interruption, mode, automation-bias, recovery, accessibility, and adaptive-interface constraints. The combined record is still speculative, so it refines Candidates 009, 011, 012, and 015 rather than becoming another registry principle.

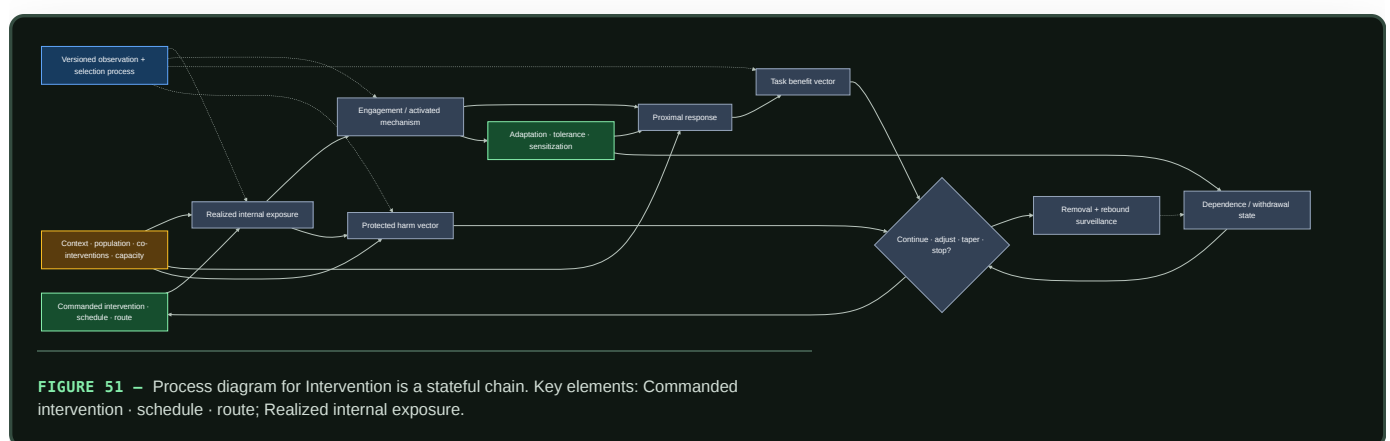
State ownership

STATE	OWNER	NORMAL WRITE PATH	NORMAL READERS
Current predictive context	runtime	every event	selected runtime modules
Recent attributable experience	episodic memory	observed outcome	routing, adaptation, maintenance
Provisional hypotheses and modules	adaptation	bounded generation and trials	shadow runtime and evaluators
Reusable representation and skill	slow model	validated consolidation	encoders, predictors, experts
Mutable proposition	factual memory	sourced versioned update	retrieval router and task modules
Stable repeated transformation	hardened store	promotion pipeline	guarded low-cost dispatch
Resource and risk policy	budget controller	calibrated policy update	every gate and route
Low-dimensional operating context	context bus	rate-limited controller update	subscribed modules through receiver-local filters
Lifecycle and fragility state	maintenance	replay, probes, regressions	admission, reopening, pruning, rollback

The plane table says who controls an operation; the state table says who may write durable state. Each boundary becomes an ablation point: remove it, widen it, delay it, or replace it with a conventional baseline and measure the consequence.

Intervention is a stateful chain

A commanded change is not its realized effect. Training intensity, routing quota, memory injection, tool access, fallback support, pruning pressure, and human escalation can each be delivered incompletely, arrive late, alter later observations, produce different useful and harmful effects, and change the system's future response. The contract therefore keeps this chain explicit: (C-607–C-610, C-616, C-626).



Editable source: state-qualified-intervention.mmd.

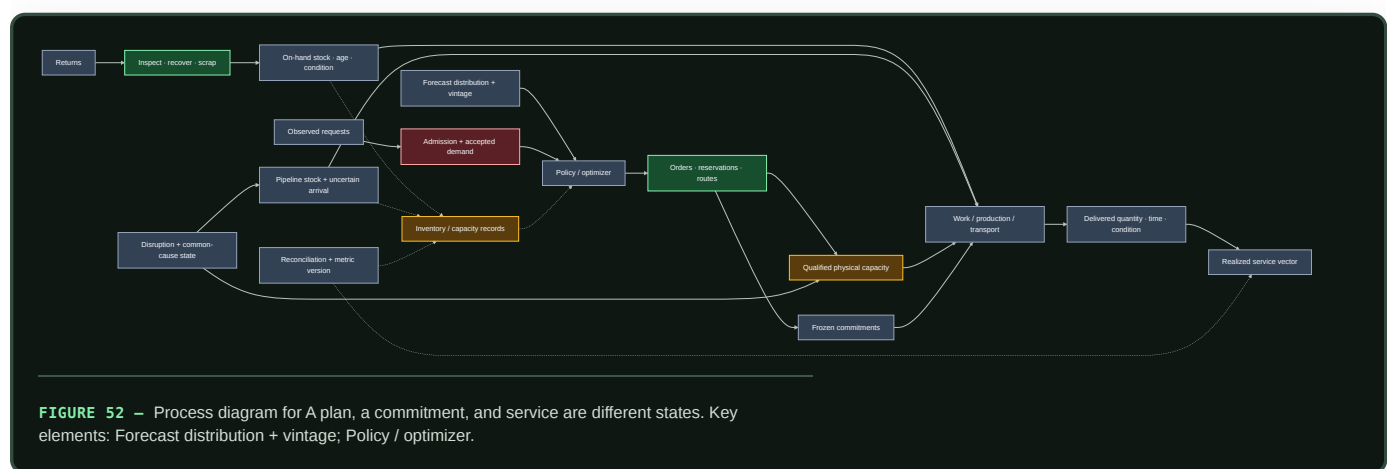
The held system record includes commanded schedule; realized state; mechanism engagement; benefit and protected-harm vectors; adaptation, dependence, and withdrawal state; population/task/context support; observation provenance; uncertainty; reserve; and permitted continuation, adjustment, taper, or stop. Its scientific bounds are recorded in C-611–C-625. It rejects three shortcuts:

1. activation or occupancy is not downstream benefit or safety;
2. equal cumulative input does not imply equal peaks, spacing, state, or recovery; and
3. removing support is a new dynamical intervention, so rollback continues through rebound, recurrence, native-capability, and reserve checks.

The state-qualified intervention mathematics defines the unit-bearing state, authority, interaction, and withdrawal tests. The composition remains a domain track under Candidates 005/007/012/014 and must beat Bayesian state-space estimation, constrained MPC/POMDP, calibrated harm monitoring, and staged decommissioning at equal lifecycle cost.

A plan, a commitment, and service are different states

Resource-bearing execution has an informational path and a physical path. A forecast can change policy; policy can authorize an order, reservation, route, or allocation; none of those creates qualified capacity, moves material, finishes work, or proves service (C-627, C-659–C-678). Filtration, internal recovery, reintroduction, final egress, storage change, and transport work instantiate the same typed-state boundary (C-1491). The complete path is:



This distinction applies when modules allocate physical devices, accelerators, network paths, robot/tool capacity, spares, storage media, human attention, or external services with lead time. Every decision-bearing record preserves:

1. forecast vintage, observed request, admission, and accepted commitment;
2. on-hand, reserved, quarantined, pipeline, and available-to-promise state;
3. qualified capacity, queue, route, setup, common-cause, and lead-time state;
4. dispatched versus delivered work, condition, completeness, and timing;
5. the service metric's denominator, clock, substitutions, and version; and
6. backlog, lost demand, returns, recovery yield, reserve replenishment, and second-event readiness.

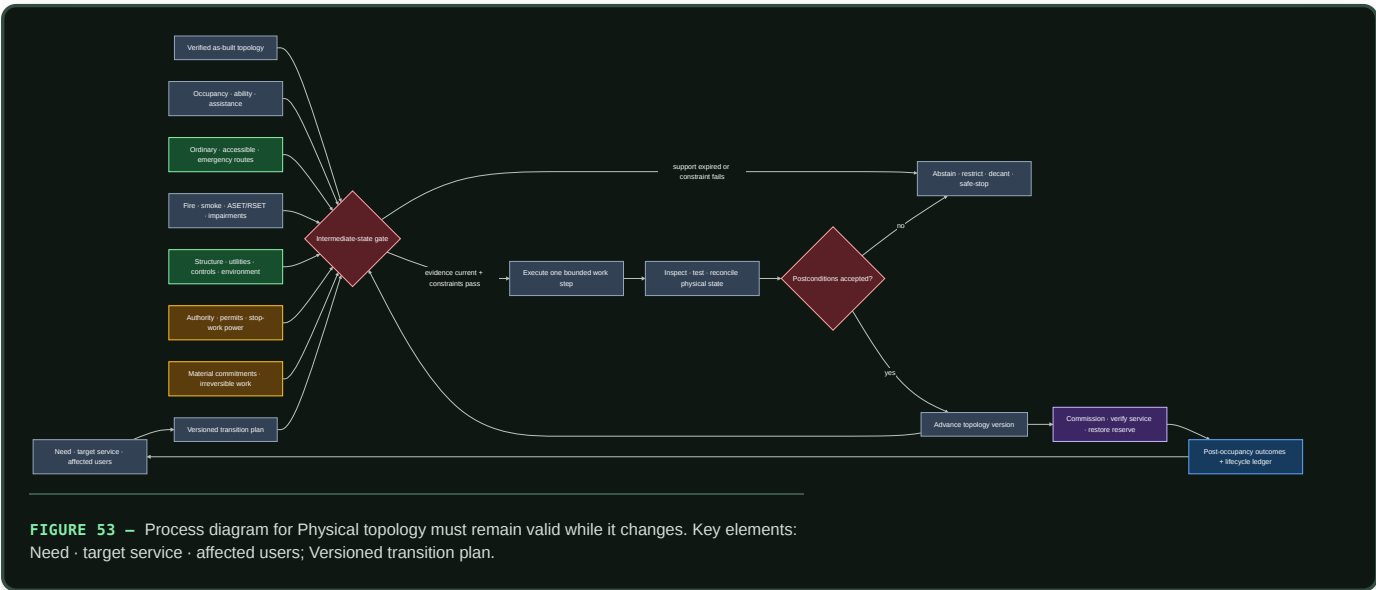
The material/service mathematics makes conservation, service, and recovery units explicit. The held composition must beat event sourcing plus inventory reconciliation, queueing, base-stock and multi-echelon control, stochastic/robust optimization, and receding-horizon planning with frozen commitments. It is an evaluation contract, not a new allocation algorithm.

Physical topology must remain valid while it changes

Logical routing can often switch within milliseconds and retry. Physical topology cannot assume that abstraction when a transition moves barriers, routes, utilities, controls, structure, people, or material. The system must represent the intermediate configurations—not only the before and after graphs.

The built-environment evidence in C-705–C-726 adds a demanding test case. A plan, BIM model, sensor dashboard, command, and verified physical state are separate records. Accessible use, egress, tenability, structure, utilities, environmental service, evidence validity, authority, and

material commitment remain explicit through every work step.



Editable source: [occupancy-qualified-spatial-transition.mmd](#).

This pattern is useful beyond buildings whenever reconfiguration has occupants, physical inventory, slow work, external authority, or irreversible steps. Its core rule is simple: a target state cannot authorize a path whose intermediate states are invalid. Hard constraints remain a conjunction rather than a scalar score; lost service stays visible by affected group; and “rollback” is claimed only when the predecessor is still physically reachable.

The full [transition mathematics](#) and [Candidate 001 stress track](#) compare the composition against ordinary design review, permits, impairment control, configuration management, commissioning, post-occupancy evaluation, and lifecycle asset management at equal budget. If that mature stack ties it, the extra composition is removed.

After an outcome: adaptation proposes, maintenance decides

An outcome can influence the next event through working state or episodic retrieval. Durable changes follow two paths with different authority:

PATH	IMMEDIATE WORK	POSSIBLE RESULT
Adaptation	capture an attributable episode; construct bounded proposals from existing fragments, relations, and abstractions	working-state change, hypothesis, or provisional module
Maintenance	score, replay, branch, compare, and account for lifecycle cost	retain, merge, externalize, weaken, delete, protect, reopen, relocate, prune, quantize, or compile

Variation creates candidates; observed outcomes and controlled interventions decide which survive. Bounded fragility probes remain on shadows or replicas and are charged to the maintenance budget. Rapid reversible regulation and slow structural adaptation remain different update processes whose cumulative cost and reversal recovery must be scored separately (C-1496).

The system may also create learning events: propose a structured hypothesis or latent rollout, choose a bounded intervention, and evaluate the outcome. This closed endogenous curriculum composes the evidence in C-061–C-066; random variation supplies candidates, not validation. Its operations are separated under equal budgets in [Candidate 004](#).

The memory lifecycle governs evidence and retention. The maturity lifecycle governs protection, reopening, and structural consolidation. A proposal remains scoped and attributable until maintenance promotes it.

Every promoted structural action records:

1. a declared trigger;

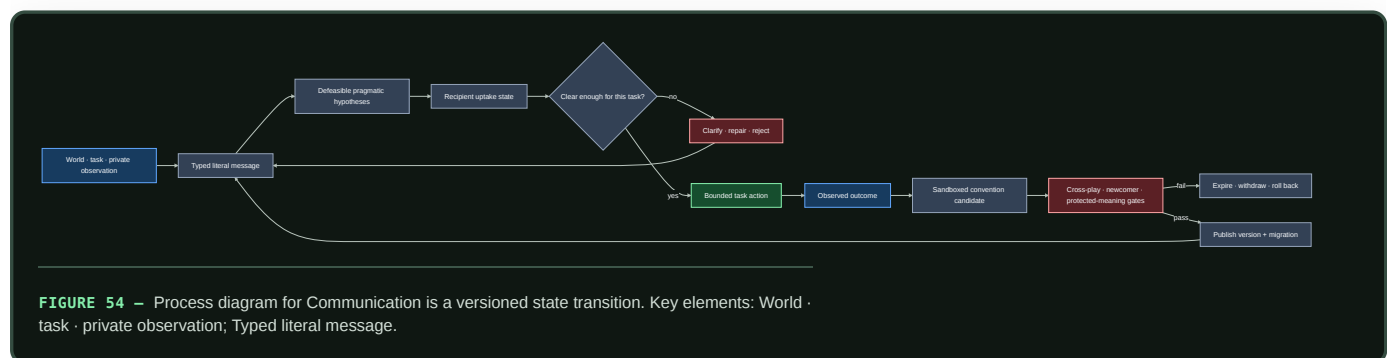
2. a versioned candidate state;
3. protected historical and rare-case evaluations;
4. measured migration, validation, and recovery cost;
5. a promotion or rejection decision; and
6. a rollback target.

Communication is a versioned state transition

A fluent message is not one state variable. The runtime keeps at least six separate records:

1. typed literal payload and protocol version;
2. external referent, query, action, or constraint;
3. defeasible hypotheses about sender intention;
4. recipient-generated uptake state—received, parsed, understood-enough, accepted, rejected, or unresolved;
5. repair and supersession lineage; and
6. any resulting authority decision, which communication alone cannot grant.

This separation follows the evidence boundaries in C-268–C-281. Composition is relative to a grammar and interpretation; pragmatic inference depends on context and incentives; acknowledgement is not identical belief; channel capacity is not semantic value; and population convergence is not grounding, truth, or safety.



Editable source: [versioned-repairable-conventions.mmd](#).

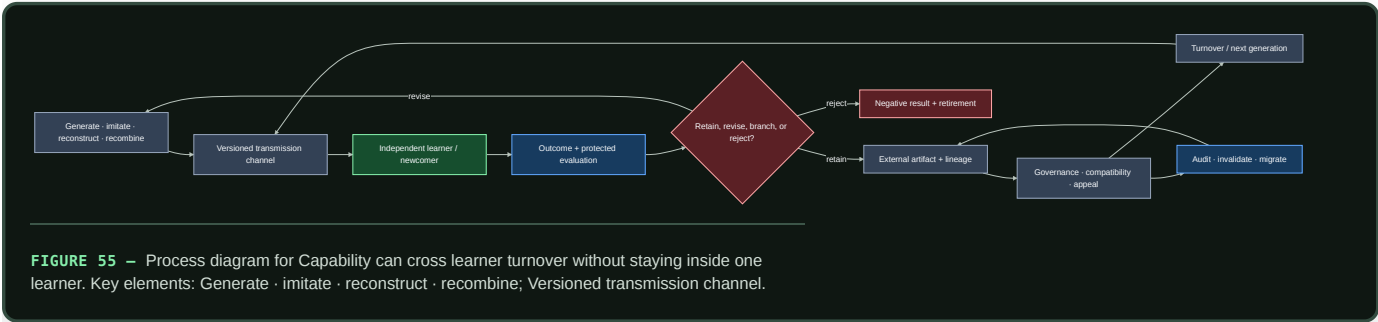
A local shorthand remains session-scoped until independent interpreters recover its declared denotation, older and newcomer agents pass cross-play, corrupted or version-mismatched messages trigger bounded repair, protected rare meanings survive, and migration plus rollback have been exercised. The full lifecycle is tested in [Candidate 015](#) against typed protocols, schema registries, acknowledgements, replicated logs, calibrated inference, standard coding, and explicit migration.

Capability can cross learner turnover without staying inside one learner

A repository, demonstration, or popular practice is not yet cumulative inheritance. The system must expose six different operations:

1. **Generate:** produce a variant by invention, reconstruction, imitation, or recombination.
2. **Transmit:** identify exactly which actions, outcomes, explanations, artifacts, tests, and environmental state reach the next learner.
3. **Evaluate:** test the reconstructed capability independently rather than crediting popularity, prestige, or successful transmission as usefulness.
4. **Retain:** place accepted capability, evidence, failures, and lineage in a form that a newcomer can retrieve and interpret.

- 5. **Govern:** control compatibility, authority, appeal, migration, and retirement without treating a convention as truth.
- 6. **Survive turnover:** repeat the cycle with a genuinely new learner and a protected test suite.



Editable source: [audited-cumulative-inheritance.mmd](#).

For generation g , a deliberately simple capability ledger is

$$K_{g+1} = q_g K_g + I_g + R_g - D_g,$$

where K_g is validated capability in a declared task-score unit, q_g is the dimensionless retained fraction, I_g is independently generated improvement, R_g is validated recombination gain, and D_g is degradation or compatibility loss in the same task unit. These terms are accounting fields, not a claim that capabilities are generally additive; protected outcomes stay visible beside the aggregate.

Raw population size also overstates diversity when every learner sees the same few lineages. If exposure weights are π_i and $\sum_i \pi_i = 1$, then

$$N_{\text{eff}} = \frac{1}{\sum_i \pi_i^2}$$

is the dimensionless effective number of equally exposed models. It measures attention concentration—not competence, independence, or complexity.

The evidence in C-343–C-367 supports bounded cultural accumulation, conditional social-information use, material scaffolding, and severe archaeological inference limits. It does not establish a population advantage for artificial learners. That residual is isolated in Candidate 019, which must beat centralized continual learning, replay, version control, retrieval, workflow engines, quality-diversity search, and fixed governance at equal cumulative effort.

Live response and longitudinal learning are separate loops

The runtime may need to contain a failure before its cause is known. Later analysis may identify a cause without improving the next live response. A usable operational-assurance plane therefore connects two loops without collapsing them:

LOOP	DEADLINE	REQUIRED STATE TRANSITION	FAILURE IF MEASURED ALONE
Live response	seconds to minutes	detect, acknowledge, assign scoped authority, contain, degrade, revoke, roll back, restore	fast containment can recur because no verified change enters future operation
Longitudinal learning	releases to months	retain competing traces, analyze, bind a finding to dependencies and an owner, test, deploy, verify, retrieve, retire	a complete report can arrive after preventable damage and may never change operation

Report volume is observation-biased. Over a declared exposure interval,

$$N_{\text{report}} = N_{\text{precursor}} p_{\text{detect}} p_{\text{report}} p_{\text{retain}}.$$

where both N values are event counts and every p is a dimensionless conditional probability. A low count can mean low exposure, weak detection, reporting friction, fear, or deletion; a high count can mean hazard, improved coverage, duplication, or gaming (C-178).

Likewise, an incident archive becomes memory only when an applicable lesson is retrieved, used, evaluated, invalidated, and retired (C-184). The held composition binds live traces and precursor detections to scoped response, dependency-linked changes, protected tests, verified outcomes, retrieval triggers, and retirement. Candidate 011 compares it with a complete SRE stack and rejects it if role, interruption, reviewer, storage, stale-memory, or coordination cost erases the gain.

Build the system in dependency order

The full system should be assembled through progressively stronger contracts:

1. Establish dense and conventional modular baselines with one physical measurement boundary.
2. Measure isolated-versus-connected behavior and retain immutable-copy, queueing, resource-isolation and explicit-feedback nulls before composing modules.
3. Align observation, action, outcome, location, and time before adding learned routing.
4. Introduce conditional execution and measure activation, data movement, and dispatch overhead.
5. Add attributable episodic memory without granting it slow-model write authority.
6. Add replay and consolidation under fixed maintenance budgets, then admit provisional modules only for measured capability gaps.
7. Test reversible protection, reopening, and structured pruning before promoting repeated behavior into cheaper paths.
8. Learn logical routing and physical placement together only after their separate baselines are stable.
9. Run continual operation with drift, conflict, newcomer admission, recovery, and complete lifecycle accounting.

Each stage keeps the strongest baseline from the previous one. A later mechanism cannot hide an earlier regression behind a higher aggregate score.

Efficiency mechanism

The architecture seeks three compounding reductions:

- **activation:** execute a small relevant subset of stored capacity;
- **movement:** keep repeated computation near the state it uses; and
- **reinterpretation:** convert stable repeated work into cheaper memory, precision, routing, or compiled paths.

Before execution, the resource plane prices the proposed route. Afterward, measured telemetry corrects that estimate. Hard risk floors remain constraints rather than terms that average efficiency can trade away.

Those reductions are useful only after control overhead is included. Lifecycle energy is therefore

$$E_{\text{life}} = E_{\text{runtime}} + E_{\text{memory}} + E_{\text{network}} + E_{\text{adaptation}} + E_{\text{maintenance}} + E_{\text{recovery}},$$

with every term measured in joules over the same workload and system boundary. The energy model defines the full comparison contract.

Evidence status

INGREDIENT	CLAIM RANGE	STATUS IN THIS SYNTHESIS
Conditional routing and early exit	C-003–C-004	available engineered mechanisms
Predictive representation and residual allocation	C-005–C-007	plausible runtime composition
Fast/slow memory and protection	C-008–C-010	scoped evidence; lifecycle policy experimental
Structural pruning and hardening	C-012–C-015	individual mechanisms available; promotion logic unvalidated
Sensorimotor control and local autonomy	C-017–C-024	biological and engineered constituents
Replay, reconsolidation, and forgetting	C-036–C-042	scoped observations; combined controller speculative
Maturation, reopening, and contextual control	C-043–C-051	scoped interventions; digital translations experimental
Collective coordination and resilience	C-052–C-060	scoped observations; quorum and fragility translations experimental
Endogenous generation and exploration	C-061–C-066	constituent observations; integrated curriculum speculative
Communication and convention lifecycle	C-268–C-281	constituent mechanisms established or scoped; versioned repairable composition speculative
Cumulative inheritance across turnover	C-343–C-367	bounded constituents established or scoped; population advantage unvalidated
Institutional authority and rule repair	C-368–C-395	formal constraints and bounded evidence; multi-level composition speculative
Human initiative, authority, and recovery	C-396–C-416	bounded HCI effects; combined record speculative
Connection-induced back-action and bounded insulation	C-1550–C-1559	scoped biochemical evidence; digital residual and energy advantage unvalidated
Complete integrated system	none	unvalidated project synthesis

Speculative extensions

- Modules bid for compute using expected task-value improvement per joule while hard risk and communication limits remain external constraints.
- Local clocks and queues replace a global synchronous step where causality and hardware permit it.
- Routing, memory placement, interconnect, and precision are learned jointly.
- Replicas run alternative maintenance decisions and compare later outcomes before one state becomes canonical.
- External tools and sensors become active experiments chosen for information gain, not passive input channels.

Failure modes

- A shared predictive state becomes a dense communication bottleneck.
- Task, resource, adaptation, and maintenance controllers oscillate or reward incompatible behavior.
- Routing overhead and scattered memory access erase activation savings.
- Episodic and factual stores disagree without an explicit conflict policy.
- Maintenance, regression, and recovery consume more energy than runtime saves.
- Premature hardening turns shortcuts or mutable facts into rigid behavior.
- Aggregate quality hides rare-case regression, fragility, or failed recovery.
- A scalar resilience score hides opposing resistance, recovery, adaptability, and newcomer-admission effects.
- Proposal generation collapses into high-temperature sampling without targeted intervention and outcome-based selection.

- Fluent exchange creates false common ground, silent semantic drift, or local conventions that newcomers and older versions cannot interpret or reject.
- A population receives credit for cumulative learning when it only performs more parallel search, preserves a headline score while losing rare skills, or copies one correlated lineage through every nominally independent agent.
- Governance vocabulary hides an ordinary optimizer, or veto, delegation, amendment, and participation add gridlock, capture, concentration, churn, or oversight cost without improving task-native or protected outcomes.
- A nominal human approval, explanation, or confidence display is credited as assurance even though state is stale, the mode is misunderstood, authority is ineffective, intervention arrives late, or recovery cannot change the world state.
- A module passes alone, then a new consumer changes its state or update path; or an always-on insulator deletes useful service while hiding copy, buffer, admission, replica, monitoring, and maintenance work.
- The architecture accumulates mechanisms faster than experiments can reject them.

Measurable predictions

1. Conditional routing reduces measured runtime data movement and energy at matched quality, calibration, and tail risk.
2. Attributable episodic memory improves adaptation latency without increasing protected slow-model regression.
3. Maintenance and structural consolidation reduce lifecycle cost after replay, validation, migration, and recovery are counted.
4. Separating factual propositions from reusable skills lowers correction cost and unsupported factual carryover.
5. Reversible maturation gates improve retention and relearning relative to fixed regularization, fixed pruning schedules, and naive fine-tuning.
6. Versioned uptake, repair, and cross-play gates reduce silent semantic failure under agent and protocol drift beyond a complete fixed-protocol stack.
7. Connection-qualified modules preserve producer trajectories and accepted consumer service across held-out client topologies; selective insulation activates only where it improves the complete non-dominated frontier.
8. Audited turnover retains and recombines more protected capability than a centralized continual learner at equal cumulative learning, coordination, storage, migration, and energy cost—or the population mechanism is retired.
9. Multi-level authority improves an applicable task-native or protected-harm frontier beyond the complete ordinary governance stack after gridlock, capture, concentration, human attention, and lifecycle cost—or it is merged.
10. A recoverable initiative record predicts failures beyond polished mode labels, previews, confirmations, undo/history, logs, runtime assurance, and user testing on consequential and accessibility-stratified tasks—or only its individually supported fields remain.
11. The integrated system occupies a better quality–risk–latency–energy frontier than every component ablation and the strongest ordinary controller, scheduler, cache, and router baselines.

Energy model and efficiency evaluation contract

Scope

Energy efficiency is not a property of a model in isolation. It is a measured relationship among a task, an input distribution, a quality and risk envelope, a latency or throughput requirement, a hardware–software system, a lifecycle horizon, and a physical measurement boundary.

This chapter defines the common contract for every efficiency result in the project. It replaces operation-count headlines with equal-budget comparisons, prices data movement and adaptation, and carries uncertainty through to an explicit reject, revise, or promote decision. The detailed notation and first models remain in the [mathematical notes](#); experiment contracts instantiate this chapter rather than inventing new accounting rules.

An efficiency result is identified by the record

$$\mathcal{R} = (\mathcal{T}, \mathcal{D}, Q, R, L, \mathcal{B}, \mathcal{H}, S, H, U),$$

where:

SYMBOL	MEANING	UNIT OR DECLARATION
$\mathcal{T}$	task and target behavior	named benchmark or environment
$\mathcal{D}$	evaluated input distribution	named dataset, stream, or generator
Q	primary task quality	task-specific score
R	declared risk or error measure	task-specific risk unit
L	latency requirement	seconds, with percentile
$\mathcal{B}$	physical accounting boundary	device, node, cluster, or facility
$\mathcal{H}$	hardware configuration	device type, count, clocks, memory, and interconnect
S	software configuration	versions, precision, kernels, compiler, and runtime
H	lifecycle horizon	seconds and qualified-event count
U	uncertainty description	confidence interval and measurement model

Two energy numbers with different records are not directly comparable. A result may vary one field deliberately, but it must show the resulting curve instead of silently carrying the old conclusion across the change.

Biological observation

Neural signaling operates under metabolic constraints (C-001). Biological systems therefore provide examples of computation shaped by the cost of activation, communication, maintenance, and adaptation. Event-driven hardware also demonstrates that local sparse activity can be implemented outside biology (C-015).

The observation does not define a common operation between a brain and a digital accelerator. A spike, synaptic event, memory read, floating-point multiply, token, and successful decision are different functional units. The inherited brain-to-accelerator range remains disputed under C-016 because its numerator and denominator do not share a task, quality target, physical boundary, or operation definition.

Here the brain establishes the feasibility of severe resource allocation. The actionable translation is to make every proposed mechanism compete under a complete physical and statistical contract.

Proposed AI translation

The qualified event is the functional unit

Let x_j be deployment event j , and let $I_j \in \{0, 1\}$ indicate that the event was served inside the preregistered quality, risk, and latency envelope. For a horizon containing N offered events, the qualified count is

$$N_q = \sum_{j=1}^N I_j.$$

N , N_q , and I_j are counts or dimensionless indicators. The envelope must define whether qualification is event-level, stratum-level, or run-level. Selectively dropping hard events cannot reduce the energy denominator: offered events, rejected events, failed events, and qualified events are all reported.

For tasks whose quality is defined only over a population, comparability is established at the run level. Candidate C and baseline B are inside the same envelope only if

$$QC - QB \geq -\varepsilon_Q, \quad RC - RB \leq \varepsilon_R, \quad LC_p \leq L_{\max},$$

where ε_Q is the allowed quality loss in quality units, ε_R is the allowed risk increase in risk units, p is the declared latency percentile, LC_p is candidate latency at that percentile in seconds, and $L_{\max}$ is the latency ceiling in seconds. All margins are fixed before confirmatory runs. Energy superiority is tested only after this envelope is satisfied.

Measurement boundary and gross energy

For boundary b and measurement interval $[t_0, t_1]$, gross electrical energy is

$$E_b^{\text{gross}} = \int_{t_0}^{t_1} P_b(t) dt,$$

where $P_b(t)$ is measured electrical power in watts, t is time in seconds, and E_b^{gross} is energy in joules. The instrument, sample rate in hertz, clock alignment, integration method, and missing-sample policy are part of U .

Incremental energy may also be reported:

$$E_b^{\text{inc}} = \int_{t_0}^{t_1} [P_b(t) - P_b^{\text{idle}}(t)] dt,$$

where $P_b^{\text{idle}}(t)$ is power in a separately measured, precisely defined idle state in watts. Gross energy remains primary. Incremental energy is meaningful only when candidate and baseline use the same idle definition and the idle subtraction does not hide reserved or provisioned capacity.

Every reported number carries one provenance label:

- **measured:** produced by a named instrument or counter during this run;
- **modeled:** calculated from measured counts and versioned coefficients;
- **cited:** copied with its original system, boundary, and date; or
- **hypothesized:** a preregistered value or direction awaiting measurement.

Typed material handling must not turn filtration, recovery, excretion, storage drift, oxygen delivery, or transport work into one inferred energy number (C-1491).

Mixed totals expose the provenance of every component. They are not labeled “measured energy” when any material term is modeled or cited.

The permitted boundaries are:

BOUNDARY	INCLUDED ENERGY
Device	accelerator package or named component only
Node	accelerator, CPU, memory, local storage, power conversion, and attributable node cooling
Cluster	participating nodes, fabric, shared storage, and attributable cluster infrastructure
Facility	cluster energy plus contemporaneous attributable facility overhead

If facility energy is derived from power usage effectiveness,

$$E_{\text{facility}} = \text{PUE } E_{\text{IT}},$$

where E_{IT} and E_{facility} are joules and PUE is the dimensionless ratio of facility power to IT-equipment power for the same site and interval. A cited fleet average is not substituted for a measured node or cluster result. Facility, carbon, water, and financial cost are separate outcomes; none is used as a synonym for joules.

The energy number is a measurement result

The measurand must name the electrical boundary, object, state, interval, conditions, aggregation, workload, and intended decision. A counter reading is an indication, not yet a result (C-519, C-520). Each reported energy value therefore retains:

1. instrument, range, firmware, voltage/current/phase configuration, bandwidth, sampling, clock, environment, and raw trace identity;
2. measurement model, integration rule, preprocessing and software version, missing-sample policy, warm-up, retry, idle, and useful-output definitions;
3. calibration chain, stated reference, corrections, validity scope, checks, drift status, and every uncertainty contribution;
4. covariance from shared meters, clocks, coefficients, environments, and preprocessing, plus coverage method and reproducibility conditions; and
5. the decision rule, target uncertainty, permitted comparison, expiry, provenance, supersession, and invalidation dependencies.

Calibration does not mean validated, traceability does not mean accurate enough, and uncertainty is not unknown error (C-521–C-526). The measurement-contract note defines the record, dimensional checks, covariance propagation, guard bands, drift review, and invalidation graph.

For sampled power, the estimator

$$\hat{E} = \sum_{m=1}^M P_m \Delta t_m$$

has units J because P_m is W and Δt_m is s. Its uncertainty model must retain correlations between samples and coefficients when they share a meter, calibration, clock, or correction. Repeated samples from one trace do not become independent experimental runs. An end-to-end calibrated boundary can rank systems differently from software estimates or device-only counters; the existence, sign, and frequency of such reversals are measured outcomes, not constants (C-536).

Paired meter blocks and seed-level inference

Fast work units can be shorter than a meter's sampling, clock-alignment, or resolution limits. Candidate 010 therefore freezes ordered opportunity blocks, counterbalances arm order within each scenario-seed cluster, and records warm-up and idle intervals separately. This reduces acquisition and review overhead without changing the inferential unit.

For seed s , arm a , and its set of measured blocks $\mathcal{B}_{s,a}$, define

$$\hat{e}_{s,a} = \frac{\sum_{b \in \mathcal{B}_{s,a}} E_b^{\text{gross}}}{\sum_{b \in \mathcal{B}_{s,a}} C_b},$$

where E_b^{gross} is measured block energy in joules and C_b is the count of correct commits in that block. Thus $\hat{e}_{s,a}$ has units J/correct commit. Every repetition and scenario is aggregated inside the seed before a candidate-baseline contrast is formed:

$$d_s = \hat{e}_{s,C} - \hat{e}_{s,B}.$$

With n independently generated seeds, the paired mean and its two-sided Student- t interval are

$$\bar{d} = \frac{1}{n} \sum_{s=1}^n d_s, \quad \bar{d} \pm t_{1-\alpha/2, n-1} \frac{s_d}{\sqrt{n}},$$

where s_d is the sample standard deviation of the seed contrasts in J/correct commit. Blocks, scenarios, power samples, and repeated measurements contribute precision and diagnostic information; none increases n .

Let $u_{s,a}$ be the declared conservative expanded measurement allowance for one seed-arm aggregate, including calibration contributions and half a meter resolution quantum per block, divided by its correct-commit count. A simple worst-direction reporting envelope widens the statistical interval by

$$\bar{u} = \frac{1}{n} \sum_{s=1}^n (u_{s,C} + u_{s,B}).$$

This envelope is deliberately conservative and does not replace a fuller covariance model. A zero correct-commit denominator, missing block, invalid review, expired calibration, excessive clock uncertainty, or fixture-shaped meter record makes the comparison undefined rather than silently dropping a seed.

The nominal two-seed design would create 6,720,000 reading-plus-review files under per-work-unit metering but 1,728 under the current paired-block design. That is an artifact-count calculation, not an energy result or a larger sample size.

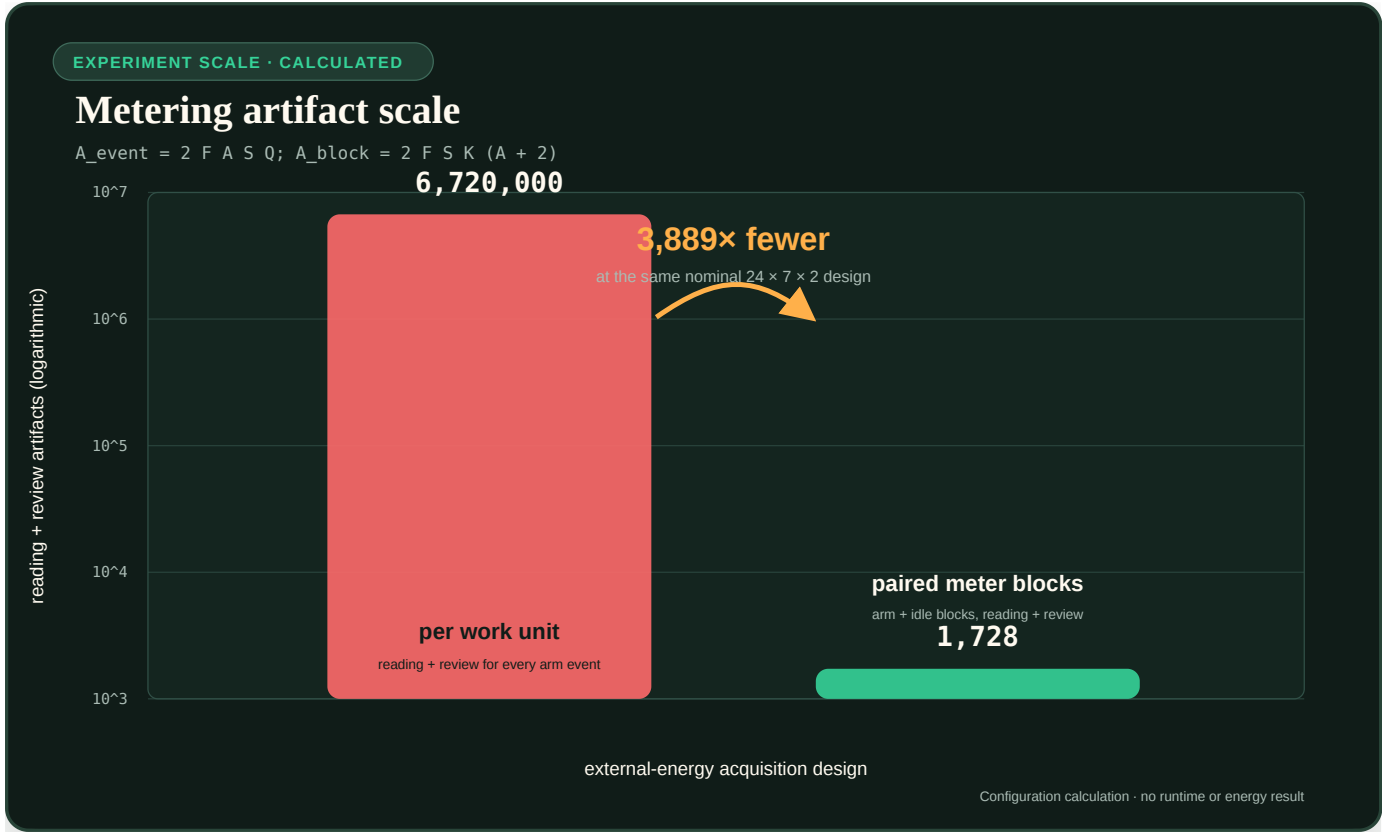


FIGURE 56 — Calculated metering artifact scale for Candidate 010

Editable assumptions: `../assets/plots/core-models.json`.

Lifecycle energy

For candidate C over horizon H , define the disjoint lifecycle total

$$E_C^{\text{life}}(H) = E_C^{\text{search}} + E_C^{\text{train}} + E_C^{\text{consolidate}} + E_C^{\text{compile}} + E_C^{\text{serve}}(H) + E_C^{\text{maint}}(H) + E_C^{\text{migrate}}(H) + E_C^{\text{recover}}(H) + E_C^{\text{idle}}(H).$$

Every E term is energy in joules at the same boundary $\mathcal{B}$:

- E^{search} covers architecture search, hyperparameter tuning, and failed development runs attributable to the selected result;
- E^{train} covers final training and validation;
- $E^{\text{consolidate}}$ covers pruning, replay, merging, or structural stabilization before service;
- E^{compile} covers compilation, quantization, layout generation, and deployment preparation;
- E^{serve} covers event execution during H ;
- E^{maint} covers monitoring, replay, repair, indexing, and lifecycle control during H ;
- E^{migrate} covers state serialization, transfer, warm-up, and reconfiguration not already assigned elsewhere;
- E^{recover} covers additional recovery work after a declared failure or regime change; and
- E^{idle} covers provisioned but inactive devices, memory, communication links, and reserve capacity during H .

Fast control and slow structure therefore require separate action, build, carry, reversal, stranded-capacity, and recovery rows before any lifecycle advantage is claimed (C-1496).

An event is assigned to exactly one term. A migration byte, for example, may appear in the movement ledger but its energy is not also charged as ordinary serving traffic.

Protection is also a lifecycle state, not a free subtraction from damage. The electrochemical solid--electrolyte interphase is a useful accounting example: it can suppress an immediate parasitic reaction while consuming inventory, adding resistance, continuing to grow, and changing regime (C-1535). For an artificial barrier, define

$$EB(H) = EB_{\text{build}} + EB_{\text{monitor}}(H) + EB_{\text{repair}}(H) + EB_{\text{replace}}(H) + EB_{\text{traffic}}(H),$$

with every term measured in joules at the same boundary. The result also keeps separate native-unit axes for consumed capacity, added latency, blocked useful work, false quarantine, damage admitted, and recovery. A filter, cache, trust layer, quarantine boundary, or checkpoint barrier wins only when its avoided downstream loss exceeds these construction, carrying, resistance, maintenance, and failure costs against fixed-barrier, rate-limiter, rollback, and no-barrier nulls. Cracking and repair in F-025 are synthetic engineering stressors, not effects attributed to the cited interphase sources.

Interface insulation uses the same lifecycle discipline but a different causal test. For an insulating path I , retain separate energy rows for producing and copying snapshots, maintaining buffers or replicas, admission and expiry, monitoring connection sensitivity, serving consumers, recovery and idle reserve. Its accepted-service denominator must include timely and fresh consumer outputs; deleting, dropping or indefinitely delaying load cannot manufacture an energy saving. Logical operations and bytes are explanatory telemetry until a calibrated physical boundary measures joules (C-1555–C-1557).

No monotone insulation--energy law is assumed. One scoped biochemical model shows a fuel tradeoff, while a countermodel reduces both coupling and fuel by accepting worse tracking or leak robustness. The artificial comparison must therefore preserve the full distortion--service--latency--memory--work--energy frontier and the no-load case. The equations and decision boundary are in the interface-qualified retroactivity contract.

Delayed damage prevents short evaluations from closing that ledger. If d_t is accumulated damage in a declared damage unit and ψ is a state-, action-, temperature-, and mode-dependent damage rate in damage units per second, then

$$d_{t+1} = d_t + \Delta t \psi(u_t, x_t, T_t, m_t)$$

is dimensionally valid for step duration Δt in seconds. Early outcome prediction may reduce the number of full-horizon trials, but reserve trials, false rankings, calendar exposure, and late failures remain charged (C-1539). A policy selected by a short proxy is not yet a lifecycle result.

Lifecycle energy per qualified event is

$$e_C^{\text{life}}(H) = \frac{E_C^{\text{life}}(H)}{N_{q,C}(H)},$$

where $N_{q,C}(H)$ is the candidate's qualified-event count during H and e_C^{life} is joules/qualified event. The same result is also reported per offered event so quality filtering remains visible. Lower friction or wear at one coupon/contact cannot replace accepted system service, mission transfer, auxiliaries, maintenance, replacement, manufacture, or allocation uncertainty in this boundary (C-1505).

Break-even horizon

Let ΔE_0 be candidate minus baseline one-time energy before service in joules, and let $\delta e = e_B^{\text{serve}} - e_C^{\text{serve}}$ be the measured steady serving saving in joules/qualified event. When $\Delta E_0 > 0$ and $\delta e > 0$, the event-count break-even point is

$$N^* = \frac{\Delta E_0}{\delta e}.$$

N^* is a count. At qualified service rate λ_q in events/second, the time break-even is $T^* = N^* / \lambda_q$ seconds. Maintenance, migration, recovery, and idle differences that grow with time must be included in the full numerical break-even calculation; the simple quotient is valid only when they are already represented in δe or are negligible over the stated horizon. A compiled surface or interface structure must include build, manufacture, inflexibility, reversal, and failed-envelope cost in this test (C-1503). If $\delta e \leq 0$, there is no energy break-even.

Data movement ledger

Executed arithmetic and moved data remain separate observables. For one run,

$$B_{\text{total}} = B_{\text{cache}} + B_{\text{device}} + B_{\text{host}} + B_{\text{fabric}} + B_{\text{storage}} + B_{\text{migration}},$$

where every B term is bytes crossing the named, disjoint boundary: on-chip cache levels, device memory, host–device interface, inter-device or inter-node fabric, persistent storage, and migration path. The ledger additionally reports byte-hops, defined as payload bytes multiplied by traversed logical or physical links, in byte-hops. Reads and writes are separated when their costs differ.

Operation and movement counts can support a calibrated model,

$$\hat{E}_{\text{model}} = \sum_{o \in \mathcal{O}} n_o \epsilon_o + \sum_{\ell \in \mathcal{L}} B_\ell \epsilon_\ell + \int_{t_0}^{t_1} \hat{P}_{\text{idle}}(t) dt,$$

where $\mathcal{O}$ is the declared set of operation classes, n_o is the executed count for class o , ϵ_o is calibrated joules/operation, $\mathcal{L}$ is the set of movement boundaries, B_ℓ is bytes crossing boundary ℓ , ϵ_ℓ is calibrated joules/byte, and $\hat{P}_{\text{idle}}$ is modeled idle power in watts. A hat marks an estimate. The model is checked against gross measured energy; it never upgrades modeled joules into measured joules.

Numerical energy-per-operation tables are tied to process, device, precision, data locality, utilization, and year. The engineering audit therefore uses them as an accounting method, not constants (computer architecture analogue). The thermodynamic floor $k_B T \ln 2$ joules describes the minimum dissipation associated with erasing one bit under its physical assumptions; k_B is the Boltzmann constant in joules/kelvin and T is absolute temperature in kelvin. It is not an estimator for an inference, multiply, or memory transfer.

Relative sensing has a reference-maintenance ledger

A relative channel can appear cheaper by deleting amplitude-indexed parameters while silently importing a maintained reference. For reference state r_t , the candidate ledger therefore separates

$$E_{\text{relative}} = E_{\text{sense}} + E_{\text{reference update}} + E_{\text{selector}} + E_{\text{fallback}} + E_{\text{state movement}}.$$

Each term is measured in joules only at a calibrated workstation boundary; before that, operations, writes, bytes and seconds remain separate. Channel- specific receptor abundance is biological evidence that a reference can be stored in interface structure (C-1548), not evidence that such storage is free. F-026 rejects the efficiency hypothesis if an explicit log ratio, streaming estimator, state-space model or compact recurrent state reaches the same task/risk frontier with lower complete maintenance cost.

Reduction and closure work stays inside the ledger

A coarse model is not credited with avoiding fine computation when its usable state depends on unreported reconstruction, healing, or fallback. For a multiscale run, define

$$E_{\text{reduce}} = E_{\text{lift}} + E_{\text{heal}} + E_{\text{micro}} + E_{\text{restrict}} + E_{\text{sync}} + E_{\text{fallback}},$$

where each term is gross measured electrical energy in joules attributable to lifting a coarse state, discarding initialization transients, advancing local fine simulations, restricting them back to coarse observables, coordinating micro/macro work, and executing a qualified fallback. The matching movement ledger separately reports bytes for every term; CPU seconds, solver steps, and right-hand-side evaluations remain non-energy diagnostics.

If a projected model truncates memory, the retained history window and omitted tail are part of the declared approximation (C-1526). If a slow reduction approaches a fold or loses its spectral gap, detection and fallback remain charged (C-1527). Heterogeneous micro-queries cannot receive free boundary reconstruction or synchronization (C-1528), and equation-free computation cannot hide lift replicas or healing inside preprocessing (C-1529). The complete dimensional and closure rules are in the multiscale-reduction contract.

The baseline receives an equally optimized implementation and the same error, risk, latency, and fallback envelope. A reduction wins only if its full measured lifecycle energy is lower after every failed query, rejected step, reconstruction, and coarse-model invalidation is retained.

Equal-budget comparisons

Candidate and baseline receive matched opportunity to succeed. Each experiment freezes:

1. training and evaluation data, stream order, changes, failures, and seeds;
2. quality, risk, latency, and availability requirements;
3. input information and look-ahead—an oracle is labeled and never used as a superiority baseline;
4. tuning trials and tuning compute in device-hours or joules;
5. provisioned parameter, memory, module, edge, or replica capacity;
6. service compute, controller compute, telemetry, and actuation cadence;
7. migration, topology-edit, replay, and reserve ceilings;
8. software optimization effort appropriate to both methods; and
9. measurement boundary, duration, warm-up, repetitions, and instrumentation.

Some mechanisms intentionally exchange one resource for another. They are not forced into a single operation count; instead, all resource axes are recorded and the quality–risk–latency–energy–movement frontier is compared. A method that violates a hard budget is infeasible, not retroactively scaled into compliance.

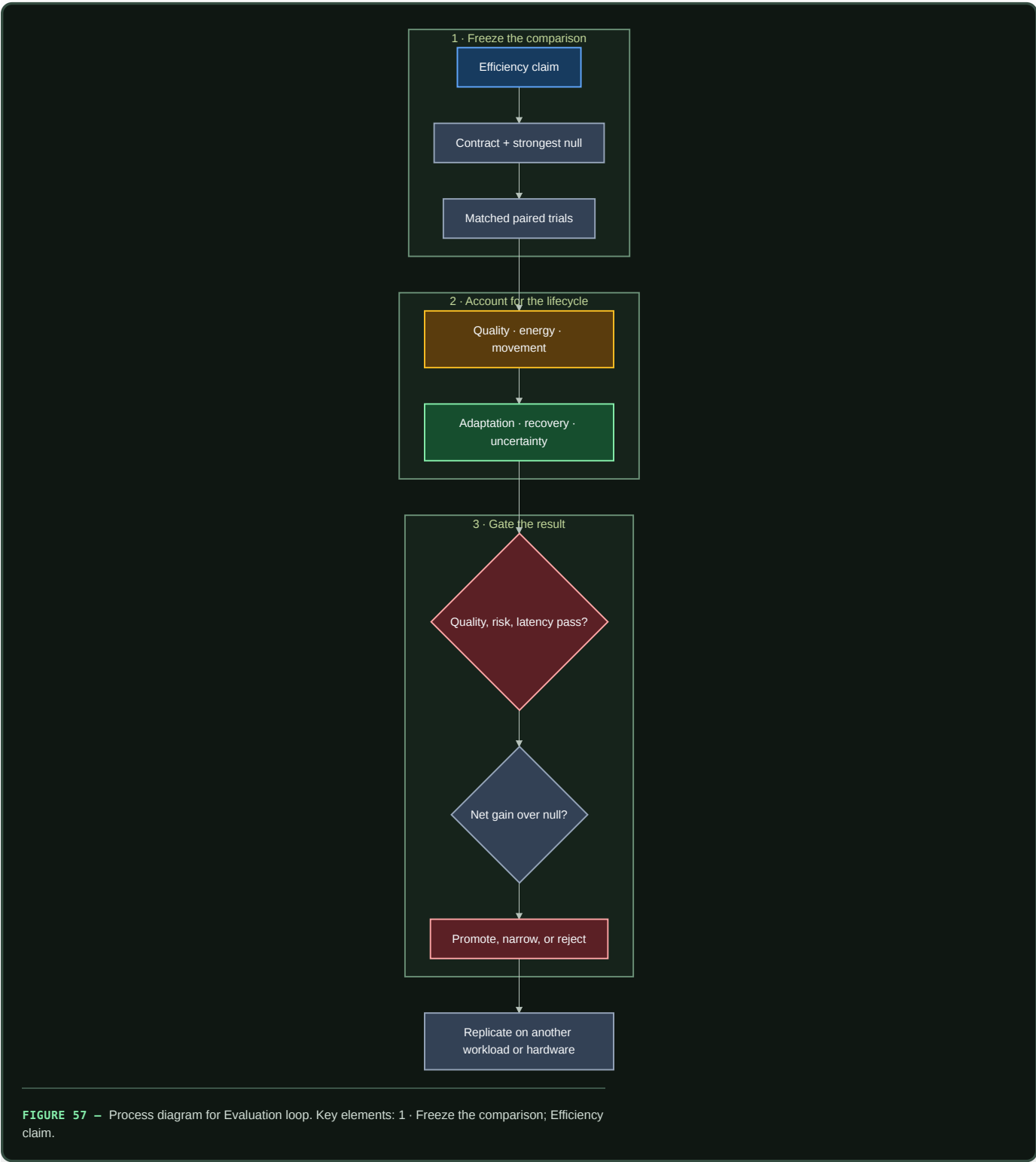
Engineering null models

The relevant null model is the strongest standard solution to the same constrained problem, not only a dense network:

PROPOSED MECHANISM	MINIMUM NULL MODEL
Conditional routing	tuned dense reference, fixed sparse model, and budgeted adaptive router
Prediction-error compute	calibrated residual/change detector plus value-of-information acquisition
Homeostatic allocation	tuned feedback or primal/dual resource controller
Adaptive topology	fixed topology with adaptive weights/routing and periodic global graph optimization
Temporal communication	fixed optimized schedule and work-conserving scheduler
Memory tiers	tuned cache/TTL/retrieval policy and, where possible, an oracle-lifetime upper bound
Maintenance plane	periodic/adaptive checkpoint, monitoring, and recovery controller
Structural specialization	profile-guided compilation, layout, quantization, and accelerator-aware kernel
Connection-qualified insulation	immutable message or copy-on-write, bounded queue/backpressure, admission/expiry, resource isolation, explicit filter/controller, replication, and no-insulator control

The full mapping and formal reference points are in the engineering analogue audit. Component ablations determine whether the claimed mechanism causes a gain. An oracle supplies headroom, while a shuffled or random-action control detects benefit from adaptation without useful information.

Evaluation loop



Editable source: `../assets/diagrams/efficiency-evaluation-loop.mmd`.

Uncertainty and missing costs

The primary comparison is paired. Candidate and baseline run the same workload seed, event sequence, change schedule, and failure trace. For paired replicate k , define relative lifecycle effect

$$d_k = \frac{e_{\mathcal{L},k}^{\text{lifa}} - e_{B,k}^{\text{lifa}}}{e_{B,k}^{\text{lifa}}},$$

where each e is joules/qualified event and d_k is dimensionless. Report the paired point estimate and a confidence interval across independent full-run seeds. Hierarchical bootstrap resampling is used when events are nested within seeds or change episodes. Repeated samples from one power trace are not treated as independent runs.

For a modeled scalar energy $\widehat{E} = f(\theta)$, where f is the declared energy-model function and θ is its coefficient-and-count vector, local covariance propagation is

$$\text{Var}(\widehat{E}) \approx Jf \Sigma_{\theta} Jf^{\top},$$

where θ is the vector of calibrated coefficients and measured counts, Σ_{θ} is their covariance matrix in the corresponding squared mixed units, and Jf is the Jacobian row vector of partial derivatives of f with respect to θ . The result is variance in joules squared. Bootstrap or Monte Carlo propagation replaces this approximation for nonlinear, correlated, or non-Gaussian estimates.

If a material category cannot yet be measured, let its energy lie in declared interval $[E_u^-, E_u^+]$ joules. The conclusion must survive the least favorable assignment within that interval. A result whose sign depends on assuming the missing category is zero remains unresolved.

Uncertainty reporting includes:

- meter accuracy, resolution, sampling rate, clock error, and calibration date;
- run-to-run variation, warm-up state, thermal state, and background workload;
- coefficient covariance and extrapolation range for modeled components;
- seed, task-stratum, failure, and change-event variation; and
- censored failures to finish or recover.

The uncertainty interval accompanies the effect and the absolute values. A large sample size does not repair a mismatched boundary or missing lifecycle phase.

Efficiency mechanism

The project's mechanisms act on distinct terms of the lifecycle and movement ledger. They must earn their complexity against the appropriate null model.

LEVER	INTENDED PHYSICAL CHANGE	NECESSARY MEASUREMENTS	COMMON NULL EXPLANATION
Selective modules and early exit	fewer executed operations and avoided activation movement	operation classes, device/host bytes, gate energy, latency by difficulty	tuned smaller dense model or confidence threshold matches it
Compartmentalization and placement	fewer cross-boundary bytes and smaller blast radius	byte-hops, cut traffic, placement/migration energy, recovery	ordinary partitioning or cache-aware layout matches it
Adaptive logical topology	fewer idle links and shorter useful paths under drift	edge-seconds, byte-hops, controller work, migration, reserve, recovery	adaptive weights or periodic graph optimization matches it
Memory lifetime routing	fewer expensive writes, replays, and stale retrievals	tier bytes, reads/writes, migration, retention quality, maintenance energy	LRU/TTL/retrieval policy matches it
Quantization and compilation	less arithmetic and movement per stable path	executed precision, kernel mix, bytes, compile energy, break-even	standard profile-guided optimization matches it
Maintenance and consolidation	lower future update/recovery cost after paid background work	replay/checkpoint bytes, optimizer work, downtime, retained quality, lifecycle horizon	periodic maintenance matches it

LEVER	INTENDED PHYSICAL CHANGE	NECESSARY MEASUREMENTS	COMMON NULL EXPLANATION
Temporal communication	less broadcast while meeting deadlines	physical context bytes, synchronization power, jitter, schedule-update cost	optimized time-aware schedule matches it

The first candidate experiments instantiate the shared contract:

- **adaptive topology** prices edge updates, reserve, migration, and recovery against routing and graph-optimization baselines;
- **multiscale context broadcast** separates logical bandwidth from physical tensor movement and board energy; and
- **recovery dynamics** prices active probes, telemetry, latency, and false alarms against standard monitoring and system-identification baselines.

An operation saving becomes an energy mechanism only if it removes physical work at the chosen boundary. An arithmetic reduction that adds irregular dispatch, cache misses, synchronization, or migration may move cost rather than remove it.

Result hierarchy

Every experiment reports the narrowest supported level:

1. **Proxy reduction:** fewer operations, bytes, active edges, or updates.
2. **Component reduction:** lower measured energy at a named device or link.
3. **System reduction:** lower gross node or cluster energy inside the matched quality, risk, and latency envelope.
4. **Lifecycle reduction:** lower $e^{\text{life}}(H)$ after search, adaptation, maintenance, reserve, failure, and idle costs at a declared horizon.
5. **Transfer:** the lifecycle result replicates on another workload or hardware class without changing the claim after seeing the result.

Evidence at one level does not imply the next. This hierarchy makes a useful proxy result publishable without inflating it into an end-to-end claim.

Audit of the inherited comparison

The source discussion asserted that a 20-watt brain would correspond to hundreds of kilowatts or megawatts of accelerator power. The calculation mixed an assumed biological operation rate, peak accelerator arithmetic at varying precisions, and a facility multiplier. Its formal shape,

$$P_{\text{counterfactual}} = P_{\text{brain}} \frac{\eta_{\text{brain}}}{\eta_{\text{machine}}},$$

is dimensionally valid only when P_{brain} is brain power in watts, $P_{\text{counterfactual}}$ is machine power in watts, and both efficiencies η measure the same qualified functional output per joule under the same quality, risk, time, and boundary record $\mathcal{R}$. No such common output has been established. The inherited values are retained as provenance for C-016, not as constants, bounds, or project targets.

The title “20 Watts Was Enough” expresses the research constraint: useful adaptive intelligence exists under a small biological power budget. The project’s numerical claims will come from the measurement loop above.

Evidence status

- Energy constrains biological signaling under the scoped evidence in C-001: **established**.
- Local sparse learning can be implemented on event-driven neuromorphic hardware under C-015: **established for the cited systems**, not proof of this architecture.
- The inherited biological-to-accelerator numerical range under C-016: **disputed**.
- Data movement, control, maintenance, and idle provision can dominate or erase nominal arithmetic savings: **engineering null hypothesis to measure for each system**, not a fixed proportion.
- Lower lifecycle energy from the integrated architecture at matched quality, risk, and latency: **speculative** until candidate experiments clear their preregistered gates and measured system boundaries.

Speculative extensions

Calibrated marginal-energy control

A controller could estimate the marginal joules of another layer, sensor, retrieval, replay, or route and compare it with expected decision improvement. The proxy would be trained from counters but recalibrated against gross energy measurements. Its calibration error and controller energy would be first-class outcomes.

Carbon- and scarcity-aware scheduling

Once joules are stable, execution time and placement could include marginal carbon intensity, water stress, or resource scarcity. These quantities retain their own units and uncertainty; they augment rather than replace the energy ledger.

Causal movement attribution

Hardware counters observe traffic but do not always identify which mechanism caused it. Controlled placement, cache-flush, route-freeze, and migration ablations could estimate the marginal bytes and joules caused by a gate, memory tier, or topology update.

Cross-substrate transfer model

A hierarchy of calibrated energy models could predict which mechanisms retain their advantage across GPUs, CPUs, accelerators, and distributed nodes. The transfer model would predict the sign and break-even horizon before the new hardware run, then be scored on calibration rather than refit after every result.

Failure modes

- **Boundary drift:** device energy is described as node, cluster, or facility energy without measuring the added components.
- **Quality leakage:** the candidate saves energy by rejecting hard inputs, lowering rare-event quality, or exceeding the latency envelope.
- **Weak null model:** a new controller is compared only with dense execution when tuned routing, caching, control, scheduling, or compilation solves the same problem.
- **Unpriced adaptation:** search, telemetry, replay, state movement, reserve, rollback, and failed updates disappear from the ledger.
- **Double counting:** migration or maintenance energy appears in both serving traffic and its own lifecycle category.
- **Proxy substitution:** FLOPs, parameters, active modules, logical messages, TDP, peak throughput, or a vendor operation table is reported as measured energy.

- **Idle erasure:** baseline subtraction removes capacity that the candidate must keep powered for latency or recovery.
- **Amortization without demand:** a one-time optimization is divided by an assumed event count beyond the measured deployment horizon.
- **Uncertainty collapse:** samples within one power trace are treated as independent replicates, or missing components are assigned zero uncertainty.
- **Optimization asymmetry:** the candidate receives more tuning trials, future information, custom kernels, or favorable batch/precision settings.
- **Hardware overgeneralization:** an irregular sparse win or loss on one substrate is stated as an architecture-wide property.
- **Thermodynamic rhetoric:** the Landauer floor or a 20-watt biological value is used to estimate attainable contemporary system energy.

Measurable predictions

H-E1 — Conditional execution

At matched quality, risk, and latency, a conditional path will reduce gross node joules/qualified event only after avoided module compute and physical data movement exceed gating, dispatch, and lost-utilization overhead. A tuned smaller dense model and fixed sparse model are required nulls. Reject the mechanism-level energy claim if the saving exists only in operation counts.

H-E2 — Data movement

Placement, compartmentalization, and structural consolidation will reduce host–device and inter-device byte/qualified event as well as byte-hops. The effect should predict a measured energy reduction after calibration. Reject the placement explanation if bytes do not fall or an ordinary partition/layout baseline matches it.

H-E3 — Lifecycle break-even

Pruning, compilation, consolidation, or specialization will have a finite N^* and T^* inside the measured deployment horizon. The result must remain positive after search, failed runs, maintenance, and rollback are included. Report “no break-even” when steady serving savings are non-positive or the observed horizon ends first.

H-E4 — Adaptive topology

Under recurrent changes and faults, use-dependent topology will reduce full lifecycle modeled and then measured joules/utility unit without violating quality or recovery margins. It must beat fixed topology with adaptive routing and budget-matched periodic graph optimization, and its frozen-topology ablation must lose the advantage. The Stage-1 thresholds are defined in Candidate 001.

H-E5 — Memory lifecycle

Lifetime-aware memory actions will reduce replay, write, retrieval, and migration energy at a declared retention and obsolete-intrusion envelope. A tuned cache/TTL/retrieval policy is the null. The scheduling and provenance overhead must be included in the break-even horizon described by the memory-lifecycle model.

H-E6 — Recovery-aware efficiency

A system with lower normal-operation energy but materially slower or less reliable recovery will not dominate. Energy, utility deficit in utility-seconds, and recovery time in seconds are reported jointly under paired faults and regime changes. Reserve capacity is charged in watt-seconds even when unused.

H-E7 — Substrate transfer

A mechanism promoted beyond one device will retain the direction of lifecycle effect on at least one second hardware class using the same task envelope and an independently calibrated movement model. If the sign changes, the result is reported as substrate-specific and the counters responsible for the reversal must be identified.

The project advances an efficiency claim only when the strongest engineering null model is outside the preregistered equivalence margin and the full lifecycle confidence interval clears the material-effect threshold. Otherwise the result narrows the design space and the biological principle remains a source of hypotheses rather than a performance claim.

Research roadmap

A publication checkpoint records progress. Completion is earned by closing a gate, retiring a hypothesis, or producing a reproducible result.

Scope

This roadmap converts an open-world research ambition into bounded iterations. Breadth remains unlimited: any scientific field may contain a transferable operation. Work remains finite because each iteration must change a canonical claim, principle, chapter, equation, diagram, decision, or experiment—or record that the new evidence changes none of them.

The stages are dependency gates rather than dates. Evidence discovery, architecture writing, null-model comparison, and measurement design continue in parallel, but an integrated system does not inherit a mechanism merely because its audit is interesting.

Live operational view

This document owns the stage order, gate meaning and exit conditions. The canonical milestone manifest projects Stages 0–5 into GitHub milestones, where issues and pull requests can be assigned, discussed and closed. The Go metadata command recreates that projection from the manifest; GitHub is not a second roadmap source.

The views answer different questions:

VIEW	WHAT COMPLETION MEANS
GitHub milestone	the attached coordination records are closed
stage gate in this roadmap	the stated dependency and rejection conditions are met for the affected mechanism
claim ledger or result record	reviewed evidence supports the recorded status within its declared boundary

A milestone percentage therefore measures operational progress only. It cannot promote a claim, turn a smoke run into evidence or waive a later gate. Due dates are omitted until a real external commitment justifies one.

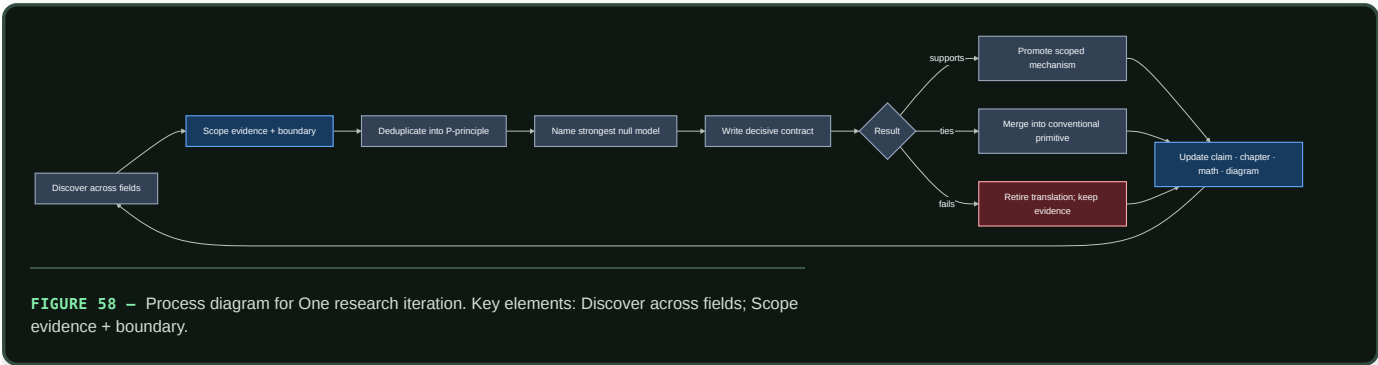
Biological observation

Development, adaptation, and maintenance operate on different timescales. Variation precedes selection; temporary states can precede durable change; stable organization remains repairable; and recurring behavior can migrate into structure. Relevant observations are separated across P-003, P-004, P-009, and P-010.

The engineering consequence is progressive commitment. Cheap reversible work comes first; expensive structural or hardware commitments require stronger evidence and longer reuse horizons.

Proposed AI translation

One research iteration



Editable source: `../assets/diagrams/research-iteration-loop.mmd`.

An iteration has three independently reviewable outputs:

OUTPUT	MINIMUM DURABLE ARTIFACT	MEANINGFUL COMPLETION
evidence delta	dated audit, primary reference, scoped C- claim proposal	the observation and its boundary can be checked without the architecture prose
synthesis delta	merge into a P- principle or a written discriminating reason to keep it separate	another field name does not create another mechanism by default
decision delta	experiment contract, rejection, or explicit no-change record	the result changes what the project would build or test

Parallel workstreams

Four workstreams move continuously, with different exit conditions:

- 1. Discovery and evidence.** Sample under-covered fields by constrained problem, audit primary sources, extract causal operations, and preserve boundary conditions.
- 2. Principle synthesis.** Deduplicate observations using the mechanism tuple in chapter 07, then update the shared primitive rather than accumulating themed components.
- 3. Concept closure.** Turn retained principles into readable data contracts, control paths, state ownership, equations, null models, failure signatures, graphics, and measurable predictions.
- 4. Decisive testing.** Compare one operation at a time against the strongest conventional method at equal interface, budget, hardware, and quality–risk envelope.

Stage 0 — Evidence, synthesis, and contracts

Stage 0 is complete for a mechanism only when:

- its observation has at least one scoped claim and primary source;
- its nearest existing AI and engineering analogues are recorded;
- its normalized operation maps to a **P-** principle or a held candidate;
- the affected chapter exposes state, information flow, timescale, physical boundary, and failure behavior;
- every equation defines symbols and units;
- a conventional null, ablation, measurement envelope, and rejection condition exist; and
- link-only discussion provenance records remain source leads rather than evidence.

The repository structure satisfies this workflow. Individual mechanisms remain at different positions inside it; the domain inventory and adoption matrix expose those differences.

Historical evidence has a separate promotion gate

Longitudinal evidence enters the roadmap as one of four typed claims:

- 1. **Accounting identity:** the declared stocks and flows reconcile.
- 2. **Descriptive regularity:** a relation exists in a sampled period, place, coding system, and surviving archive.
- 3. **Causal estimate:** an intervention or design-specific contrast identifies an effect under explicit assumptions.
- 4. **Prospective prediction:** a frozen model succeeds on a period, place, lineage, or regime unavailable during development.

No row upgrades itself into the next. A cohort matrix can be arithmetically correct and forecast poorly under changing rates. An adoption curve can fit while influence, common cause, homophily, and command remain unidentified. A historical instrument can estimate a local effect without predicting a future system. A surviving record can support chronology while omitting destroyed or never-produced evidence.

For interval Δt seconds, the basic population identity is

$$N_{t+\Delta t} = N_t + B_t - D_t + I_t - O_t,$$

where N is qualified units at a timestamp and B, D, I, O are entry, retirement, immigration, and outmigration counts during the interval. The identity finds bookkeeping errors; it does not explain or forecast a flow. The full cohort, selection, diffusion, collapse-vector, and prospective-test contract is derived in population-observation math and folded into Candidate 014 rather than promoted as a new principle.

Stage 1 — Isolated mechanism experiments

The first tests isolate control operations before composing them.

CONTRACT	CANDIDATE OPERATION	STRONG NULL CLASSES	GATE
001 adaptive topology	reversible birth, routing, merge, and retirement	fixed sparse routing, periodic/global reoptimization, oracle bounds	structural changes improve a declared frontier after controller and migration cost
002 context broadcast	rate-limited context decoded by receiver-local temporal filters	FILM, recurrent gates, global tokens, low-rank hypernetworks, gain scheduling	extra temporal decoding value survives an equal bit-rate and state interface
003 transition-class-aware recovery dynamics	bounded probes estimate hidden restoring margin and reject inapplicable transition classes	SLO/headroom, change detection, state-space ID, observability analysis, mechanism-specific indicators	useful warning survives probe and false-action cost while noise, rate, boundary, hidden-mode, and jump cases receive calibrated abstention
004 endogenous curriculum	generate, intervene, evaluate, and retain structured proposals	imitation, matched stochastic sampling, active learning, model-based planning, evolutionary search	compositional or causal gain survives equal proposal, interaction, evaluation, memory, and energy budgets
005 severity-ordered containment	contain spread, choose the least destructive qualified response, verify, escalate, and replenish	circuit breakers, static isolation, checkpoints, replicas, microrestart, scrubbing, rejuvenation, Bayesian repair/replace	collateral-loss and availability gains survive sensing, reserve, copying, replacement, and verification cost
006 reversible physical skill	compile a mature physical mapping into a rewritable local substrate with health probes and digital fallback	passive mechanics, analog control, FPGA/ASIC, fixed physical and matched digital reservoirs	a measured conversion, transport, recurrence, or command path is removed and lifecycle break-even precedes qualified retirement
007 endogenous-observation surveillance	jointly estimate hidden state and action-altered observation while preserving sensing/action provenance	residual CUSUM/GLR, nowcasting, maximum coverage, value-of-information sampling, delay-aware POMDP/MPC	decisions improve on true data vintages without mistaking policy-induced telemetry suppression for recovery
008 contestable modular allocation	withheld audits, protected entrants, real opportunity consequences, and replayable commitments under measured strategic private information	calibrated routers, constrained/primal-dual control, contextual bandits, randomized evaluation, applicable matching/scoring mechanisms	tie or lose under cooperative direct observability; improve protected outcomes under gaming, identity, collusion, and allocator threats at equal cost
009 graded assurance envelopes	bind distinct proof, effect, authority, evaluation, monitor, provenance, update, and recovery claims to a module version and invalidation graph	typed APIs, sandbox/IAM, CI/static analysis, runtime policy monitoring, lineage, canaries, transactions, schema/build invalidation	reduce unsafe admission, stale assurance, collateral rollback, and attribution time beyond the complete composed null at equal lifecycle budget
010 reset-coupled staged verification	create a reversible trace, invoke conditionally informative verification under commitment risk, then commit or reset with provenance	conditioned SPRT, calibrated cascade, abstention, retry plus rollback, redundant verifier, error-detecting code	improve false-commit, false-reject, latency, rollback, and energy frontiers only where later evidence is genuinely new and reset is cheaper than error
011 dual-loop operational assurance	connect bounded live command, containment, escalation, and restoration to dependency-linked verified learning, retrieval, and retirement	mature telemetry/tracing, SRE roles, IAM/interlocks, canaries, chaos drills, postmortems, action tracking, searchable runbooks	reduce tail containment and recurrence at equal authority, interruption, reviewer, storage, compute, human-time, and energy budgets

CONTRACT	CANDIDATE OPERATION	STRONG NULL CLASSES	GATE
012 latency-qualified authority	derive admissible local action from evidence age, integrity, mode, headroom, and coordination, then abstain, fall back, or hand off	adaptive protection, gain scheduling, supervisory/constrained/robust control, barrier functions, runtime assurance	improve service and risk without self-certification, unsafe envelope transitions, extra authority, privileged state, or uncharged reserve
013 deficit–capability routing	send compressed unmet demand upward, return rate-limited scarcity context, and allocate or grow only where local capability is present	delayed centralized optimizer, backpressure, primal–dual allocation, learned routing, ordinary feasibility gates	improve fulfilled demand, tail starvation, regret, messages, oscillation, and energy without strategic capture or stale structural growth
014 versioned observation contracts	preserve response, exposure, spatial/temporal support, preservation, intervention, selection, association, uncertainty, vintage, supersession, and follow-up dependencies with each inferred claim	typed/event-time evidence schema, state-space and Gaussian-process models, calibrated likelihood, selection function, lineage, multiplicity control, predictive checks, simulation calibration, surveillance schema, graded assurance	reduce overconfidence, invalid mixed-clock fusion, stale claims, and wasted follow-up beyond the complete composed null at equal lifecycle cost
015 versioned repairable conventions	adapt local message forms while preserving typed literal meaning, uptake, repair lineage, protected meanings, compatibility, and rollback	typed protocol/DSL, schema registry, acknowledgement/retry, calibrated inference, replicated logs, standard coding, explicit migration and canary rollback	improve heterogeneous cross-play and drift adaptation without silent semantic failure or hidden decoder, repair, migration, and review cost
016 conflict-bounded unit transition	define reproducible collective descendants, partition member and collective selection, and retain only conflict control with a net joint gain	ordinary modularity, typed interfaces, permissions/tests, clean versioning, external evaluation, routed experts, ensemble and population selection	demonstrate heritable collective capability and positive cost-adjusted between-collective selection under shortcuts, turnover, founder risk, and common-mode failure
017 contract-preserving semantic compaction	shrink histories under interpreter-, query-, representation-, evidence-, uncertainty-, rollback-, invalidation-, deletion-, access-, and corruption-degradation obligations	indexed history, snapshot plus suffix, views, key compaction, lossless compression, cold archive, extractive evidence summary, OAS/PREMIS package, versioned schema/vocabulary, Candidate 014	pass frozen hidden future-query, migration, dependency-turnover, and invalidation suites across workloads and hardware after complete physical, curatorial, and maintenance cost
018 value-aware artifact tiering	place artifacts using access/recompute forecast, task/evidence loss, invalidation value, and reconstructability under correlated failure	LRU/LFU, ARC, W-TinyLFU, size/miss-cost caching, static/economic tiering, coded fault-domain placement	improve leakage-free out-of-sample placement under shift after metadata, migration, tail, privacy, and failure costs
019 audited cumulative inheritance	preserve and recombine validated capability across learner turnover through separately versioned generation, transmission, evaluation, lineage, external state, and governance	centralized continual learning, replay/distillation, version control, retrieval, repositories, workflow engines, quality-diversity search, fixed governance	improve protected capability, recombination, compatibility, and newcomer learning at equal cumulative work after coordination and lifecycle cost
020 constitutionalized multi-level control	scope local authority, escalate spillovers, type veto and appeal, revoke delegation, expose agenda, and repair higher-order rules under expiring emergency power	constrained optimization/MPC, IAM/interlocks, policy-as-code, separation of duties, independent evaluation, append-only lineage, runtime assurance, incident/change management	improve task-native or protected outcomes in two applicable task families after capture, gridlock, concentration, communication, oversight, and lifecycle cost
027 connection-induced back-action fixture	pair isolated and connected executions, discriminate causal load classes, and apply bounded insulation only where producer and consumer service require it	immutable messages, copy-on-write, bounded queues, backpressure, admission, resource isolation, explicit filters/controllers, stop-gradient/adapters, replication	retain no new principle unless a connection-specific residual improves the held-out distortion–service–latency–resource frontier and abstains on no-back-action or useful-feedback controls
grounding loop	action-conditioned multimodal prediction and acquisition	passive multimodal learning, text-centric pretraining, standard world models	intervention and composition gains survive shortcut and missing-modality controls
memory lifecycle	capture, replay, externalize, weaken, or delete	reservoirs, prioritized replay, databases, caches, continual-learning regularizers	retention–adaptation improvement survives storage, maintenance, and deletion risk
mature hardening	compile, quantize, retrieve, or keep plastic through separate gates	compiler/autotuning, standard quantization, RAG, caching, distillation	each target produces a physical saving without freshness or recovery regression

Stage-1 exit gate: at least one isolated operation changes the measured quality–risk–latency–energy frontier relative to its strongest null. A tie merges the biological translation into the conventional primitive; a loss retires the translation while preserving the source evidence.

Stage 2 — Runtime plus adaptation

Combine the event controller from chapter 30 with attributable episodic capture from chapter 40. Keep slow structural state protected while a branch adapts.

Required composed ablations:

- runtime only;
- adaptation only;
- runtime plus adaptation without maintenance pricing;
- full two-loop system with storage, replay, rollback, and controller cost; and
- dense or monolithic continual-learning baselines at the same capacity and observation stream.

Stage-2 exit gate: sequential adaptation improves a declared task stream without a retention, calibration, rare-event, rollback, or lifecycle-energy regression hidden by the average score.

Stage 3 — Grounded developmental curriculum

Introduce temporally aligned perception, action, outcome, and language. Compare passive observation, action-conditioned prediction, targeted acquisition, and text-centric learning at matched capacity and interaction budget.

The environment must expose interventions, counterfactual structure, missing modalities, sensor noise, delayed outcomes, and distribution shifts. Language may describe, query, or compress the learned state; it cannot be the only path by which every evaluation variable enters the model.

Stage-3 exit gate: held-out intervention, temporal prediction, transfer, and composition gains survive leakage, memorization, simulator-state, and language-shortcut controls.

Stage 4 — Structural maturation

Apply the reversible maturity cycle from chapter 50 and the separated hardening targets from chapter 60. Candidate operations include structured pruning, module fusion, compilation, placement, quantization, retrieval externalization, write protection, and local reopening.

Every operation receives its own invalidation and recovery test. Parameter count is diagnostic; resident bytes, transferred bytes, wall energy, tail latency, maintenance, and rollback determine the system result.

Stage-4 exit gate: benefits remain positive after discovery, validation, migration, shadow traffic, failed promotions, recovery, and the observed reuse horizon are included.

Stage 5 — Substrate co-design

Compare accepted mechanisms on optimized conventional hardware and on event-driven, near-memory, neuromorphic, or other appropriate substrates. C-015 establishes feasibility for a scoped class of event-driven systems; it does not select a winner for this workload.

Algorithm, precision, memory layout, interconnect, batching, cooling boundary, and hardware control policy become joint experimental variables. The same qualified event and quality–risk envelope apply to every substrate.

Stage-5 exit gate: an independently reproducible end-to-end lifecycle advantage transfers beyond the workload or hardware used to tune it.

Breadth-wave decisions and next queue

The latest completed audits already changed the architecture rather than merely adding citations:

FIELD CLUSTER	DECISION NOW RECORDED
endocrine and circadian control	most pulse decoding maps to P-002/P-006/P-011; persistent receiver-local phase remains held against clocks, timers, PLLs, and schedules in Candidate 002
proteostasis and organelle quality control	sensing, containment, triage, repair, extraction, removal, replacement, and verification are separated; their ordered composition is Candidate 005
fault tolerance and reconstruction	exact restore, coding, replication, detectors, scrubbing, and self-stabilization are mandatory nulls; only underdetermined functional repair remains speculative
Earth-system transition signals	Candidate 003 is narrowed to gradual observable stability loss and now includes noise, rate, flickering, hysteresis, spatial, hidden-mode, boundary, jump, and benign-slowness controls
adaptive materials and self-assembly	demonstrated physical computation, memory, patterning, assembly, and healing map to P-006/P-009/P-010; reversible physical skill compilation becomes Candidate 006
epidemiology and surveillance control	placement, delay, ascertainment, pooled proxies, and intervention feedback map to P-001/P-007/P-009/P-013; endogenous observation becomes Candidate 007
economics, market design, and incentives	prices, auctions, matching, scoring, bandits, sanctions, and ledgers map to existing bundles; contestable allocation becomes Candidate 008 only for persistent strategic private information
programming languages and verification	proof, monitored behavior, authority, rollback, provenance, and empirical evidence remain distinct assurance classes; their versioned binding and invalidation becomes Candidate 009
chemistry and reaction networks	recognition, proofreading, nonequilibrium drive, autocatalysis, compartments, reaction–diffusion, transient assembly, molecular compute, oscillators, and integral control map to existing bundles; reset-coupled staged verification becomes Candidate 010 only where later evidence is conditionally new
high-reliability organizations and incident learning	field observations and doctrines do not transfer effect sizes; report denominators, topology, bounded escalation, incident command, and memory-in-use matter; the live/learning interface becomes Candidate 011

FIELD CLUSTER	DECISION NOW RECORDED
power grids and protection	local protection, adaptive settings, estimation, synchronization, reserve, islanding, cascade, and black-start restoration become mature nulls; evidence- and headroom-qualified authority becomes Candidate 012
plant distributed control	wound alarms, hydraulic/chemical drought response, patch foraging, tropisms, branching, source–sink flow, and mycorrhizal exchange map to existing bundles; organism, symbiosis, and community scales stay separate; deficit–capability routing becomes Candidate 013
astronomy and planetary remote inference	the latent-to-measurement-to-selection-to-association chain becomes explicit; non-detection, trials, model support, predictive checks, computation calibration, and causal identification remain distinct; Candidate 014 tests dependency-bearing observation contracts
geology and geomorphology	efficient-looking fracture and drainage networks become passive-physics nulls; connectivity is not conductance, erosion implies relocation, topology change can destroy service, and mixed supports extend Candidate 014 rather than creating a new principle
security and cryptography	authentication, authorization, information flow, detection, containment, and clean recovery remain separate; adversary, epoch, revocation, independence, compromise-horizon, and clean-root fields refine Candidates 009 and 012 rather than creating a new principle
linguistics and communication	formal composition, pragmatic inference, common ground, repair, turn timing, coding, iterated learning, grounding, and convention formation remain separate; Candidate 015 tests their versioned lifecycle composition against mature protocol engineering
paleobiology and major transitions	demographic and reproductive bottlenecks, extinction and recovery, acquisition and integration, aggregation and individuality, exaptation, constraint, and fossil observation remain distinct; Candidate 016 retains only reproducible collective heredity plus costed conflict control
pathology and rehabilitation	local selection can oppose host goals; sampled response is not eradication; tolerance is layered; output can hide compensation or depleted reserve; plasticity and dose have no monotonic sign; compensation-aware recovery becomes a cross-candidate evaluation method rather than a new principle
databases and storage	transactions, isolation, real-time ordering, replication, access paths, caches, compaction, reclamation, replay, temporal coordinates, coded repair, and tiering remain distinct; Candidates 017 and 018 retain only finite semantic preservation and task/evidence value beyond mature storage nulls
cultural evolution and archaeology	generation, transmission, evaluation, retention, governance, accessible model diversity, path dependence, and archaeological formation remain distinct; Candidate 019 tests only the residual population-level inheritance advantage across turnover
social choice and institutional governance	theorem assumptions, empirical effects, and authorized normative commitments remain separate; aggregation, strategy, delegation, veto, agenda, capture, amendment, participation, and path dependence refine Candidates 008/011/012/015/016; Candidate 020 tests only the applicable multi-level composition
human–computer interaction and human factors	nominal human presence, trust, transparency, explanation, preference, and adaptation are not assurance; effective initiative binds timely state, actual/pending mode, executable authority, state delta, recovery, resumption, expiry, accessibility, and full human cost into Candidates 009/011/012/015
quantitative history and demography	stock, flow, cohort, age, period, migration, replication, retirement, diffusion, archive selection, causal identification, collapse dimensions, and prospective validation remain distinct; their versioned composition strengthens Candidate 014 rather than creating another principle
aerospace, maritime, and safety-critical autonomy	stabilization, guidance, navigation, integrity, ODD, envelope protection, redundancy, fault stages, degraded service, fallback, collision avoidance, remote authority, certification, maintenance, and investigation remain distinct; validated asynchronous transfer refines Candidate 012 only
soft and active matter	passive relaxation, continuous fixed drive, external feedback, and adaptive policy remain separate; flocking, phase separation, defects, jamming, and assembly are physical mechanisms or nulls, while reversible phase-field compilation becomes only a Candidate-006 track
mechanical and civil resilience	compliant/passive mechanics, SHM, redundancy, robustness, graceful degradation, damage tolerance, fatigue/fracture, maintenance, max-flow/assignment, and transport recovery remain mature nulls; path-dependent residual capacity becomes a Candidates-005/012/014 test schema only
tribology, contact mechanics, lubrication, and adaptive interfaces	friction, real contact, film, wear, instability, third-body inventory, texture, adaptive coatings, and lifecycle transfer remain system-, operator-, support-, regime-, path-, and mission-qualified; Fixture F-021 carries C-1497–C-1505 across Candidates 001/003/005/006/009/012/014/018 without promoting a principle or candidate
chemical and process engineering	conservation, nonlinear operation, separation, recycle, heat integration, MPC/RTO, fault stages, safety layers, operability, and plantwide control become mature nulls; only conservation-qualified flowsheet reconfiguration remains as a Candidate-001 stress track
metrology and measurement science	measurand, indication, calibration, adjustment, verification, validation, traceability, uncertainty, covariance, reproducibility, comparisons, decision rules, drift, software, and provenance remain distinct; their dependency-invalidating composition refines Candidates 009/014 and the energy model
developmental biology, morphogenesis, and regeneration	signal, competence, positional instruction, repair workforce, current context, field source, receiver geometry, scaling envelope, load path, boundary state, redundancy, local/global composition, and organizer selection remain distinct; Fixture F-022 carries C-1506–C-1515 without promoting a principle or candidate
microbial ecology, biofilms, and fungal networks	molecular signaling, receiver response, cell physiology, transport, lineage ecology, selection, organismal flow, symbiosis, and ecosystem consequence remain distinct; one transported-field fixture refines P-011/P-013 and Candidates 001/013
animal navigation, sensory ecology, biomechanics, and motor control	sensing policy, emitted signal, propagation, body mechanics, controller, calibration, environment, and task are bound explicitly; controlled observability and controller–plant binding refine Candidates 006/007/012/014 rather than becoming new principles
integrative comparative physiology	aggregate supply/local delivery, conditional branching, finite exchange, typed renal handling, delayed regulation, regional/global matching, predictive burden, synchrony/function, and fast/slow adaptation remain distinct; Fixture F-020 carries C-1488–C-1496 without promoting a principle or candidate

FIELD CLUSTER	DECISION NOW RECORDED
plant plasticity, memory, and systemic signalling	digital population memory, lifecycle reset, writer/trace/retrieval, regenerative placement, common alarm plus typed context, route identity, sense-by-growth admission, developmental resource state, boundary sensing, and integrated environmental state remain distinct; Fixture F-023 carries C-1516–C-1525 without promoting a principle or candidate
applied multiscale reduction	exact projected memory, normal-hyperbolicity support, macro-to-micro reconstruction, and lift–heal–restrict closure remain distinct from generic averaging or fast/slow prose; Fixture F-024 carries C-1526–C-1529 and the closure contract without promoting a principle or candidate
electrochemistry	interface kinetics, transport, observation validity, inverse resolution, rate-dependent phase state, passivation cost, local depletion, parameter identifiability, path-dependent apparent equilibrium, and delayed degradation remain distinct; Fixture F-025 carries C-1530–C-1539 without promoting a principle or candidate
relative sensing, mechanism equivalence, and scale symmetry	full-trajectory invariance, adaptation, causal properties, generator provenance, intervention-qualified equivalence, experimental units, parameter and lineage shift, calibration, selective risk, singular boundary-layer error, finite support, reference lifecycle, collective transport, and observation-dependent recoverability remain distinct; Fixture F-026 carries C-1540–C-1549 plus C-1560–C-1568; its <code>RSD-T02-POP</code> population, shift, and selective-risk route is a complete written protocol but remains non-executable and <code>NO_RESULT</code> ; the scale-symmetry contract and mechanism-equivalence contract remain non-promotional
synthetic biology, signalling back-action, and control insulation	downstream sequestration, substrate competition, shared-resource coupling, intended feedback and ordinary contention remain causally distinct; Fixture F-027 carries C-1550–C-1559 and the retroactivity contract without promoting a principle or candidate
parasitology and adversarial lifecycles	observable surface and latent identity, costed lifecycle commitment, host/control-plane capture, intervention-induced competitive release, and time-shift antagonism remain distinct; the field audit deduplicates them against existing mechanisms, while Fixture F-028 carries C-1569–C-1572 as four protocol-complete, specification-only <code>NO_RESULT</code> contracts without promoting a principle or candidate
ecological priority effects, task order, curriculum, and scheduling	succession, randomized arrival-order effects, sequential parameter interference, order search, and non-learning dispatch remain distinct; the history-conditioned audit and parity contract refine C-008/C-056/C-057/C-574, F-014, and Candidates 004/019 without adding an ID
pharmacology and toxicology	commanded intervention, realized exposure, engagement, response, benefit, harm, adaptation, dependence, withdrawal, interaction null, and population support remain separate; a state-qualified intervention fixture refines Candidates 005/007/012/014
operations research, supply chains, and learning science	confidence and inventory records share one proxy-versus-state firewall, but instructional events and material commitments remain different transitions; horizon-qualified learning refines Candidates 004/019 while material service refines existing allocation/recovery candidates
legal evidence and procedure	normative authority, doctrine, empirical effect, formal inference, authentication, admissibility, weight, sufficiency, challenge, review, remedy, and finality remain distinct; a burden-qualified contestable-decision record refines Candidates 009/010/011/014/015/020
built environments and urban systems	accessibility, spatial configuration, passive physics, occupancy, controls, information models, commissioning, life safety, accessibility, infrastructure dependency, lifecycle burden, and recovery remain separate; no new principle survives, while an occupied-transition stress track narrowly refines Candidate 001
immune tolerance, adaptation, and memory	representation, recognition, deletion, ignorance, anergy, active suppression, exhaustion, expansion, contraction, affinity selection, trained state, context, memory maintenance, residency, and immunometabolism remain distinct; no new principle or candidate survives, while a typed lifecycle contract remains an evaluation discipline
music cognition, performance, and improvisation	statistical expectation, grouping, tonal state, phase correction, motif transformation, auditory–motor feedback, expressive plans, ensemble synchronization, improvisation, practice, far transfer, and enculturation map to existing bundles and ordinary sequence/control/retrieval/search/curriculum nulls; shared-clock-free partner- and phrase-specific co-adaptation remains a benchmark fixture only, not a principle or candidate
library, archival, and information science	provenance is not truth, fixity is not authenticity, retention is not useful memory, and retrieval rank is not epistemic confidence; appraisal, designated communities, representation dependencies, authority/vocabulary drift, migration, FAIR packaging, and institutional capability refine Candidates 014/015/017/018/019, while no new principle survives
comparative cognition, tool use, and flexible action	manufacture, causal transfer, future preparation, event memory, uncertainty control, social learning, teaching, recombination, exploration, inhibition, route recovery, and peripheral control are kept task-, opportunity-, morphology-, observation-, and history-qualified; species/task rankings are rejected, and Fixture F-003 refines evaluation for Candidates 004/006/007/014/018/019 without promoting a principle or candidate
visual art and design cognition	reconstructive generation, external workspaces, epistemic action, analogy, copying, fixation, proposal/selection separation, qualified aesthetic evaluators, material feedback, cultural accumulation, and provenance map to existing bundles; Fixture F-002 tests a versioned reconstruct–externalize–inspect–transform–evaluate composition across Candidates 004/014/017/018/019 without promoting a principle or candidate
mathematical practice and proof discovery	conjecture, analogy, abstraction, counterexample search, proof decomposition, libraries, representations, invariants, theorem proving, certificates, SAT/SMT, model finding, experimental mathematics, collaboration, and teaching map to mature methods; Fixture F-004 keeps attributable proposals separate from typed acceptance, reconstructs proof DAGs through small checkers, uses dependency-safe splits and reverse-dependency invalidation, and charges complete human/joule budgets across Candidates 004/009/010/011/014/017/019 without promoting a principle or candidate
fluid dynamics and turbulence	signed cascade/flux, intermittency and tails, detector-qualified coherent structures, closure/model-form support, objective-qualified ROMs, adaptive-resolution total work, observation/assimilation/observability, sensor-placement transport, stable net-benefit control, a mixing vector, path-dependent transition, calibrated extremes, measurement operators and uncertainty, and the full energy boundary remain separate; Fixture F-005 carries the outcome firewall across Candidates 002/003/006/007/012/014 without promoting a principle or candidate
sports expertise, adaptive performance, and team coordination	anticipation/interception, causal cue use, practice/retention/transfer, relevant variability, speed/accuracy/risk/energy, fatigue/readiness/damage/return, shared information/coordination, adversarial deception, feedback dependence, prospective selection, and full human/joule budgets remain separate; Fixture F-006 carries the representative history/resource-qualified performance contract across Candidates 002/004/006/007/009/012/014/019 without promoting a principle or candidate

FIELD CLUSTER	DECISION NOW RECORDED
optics, photonics, and inverse sensing	forward operators, finite information modes, null-space ambiguity, photon/noise budgets, structural priors, active/coded acquisition, adaptive correction, drift, saturation, fusion, optical transforms, conversion, analog error, fabrication, thermal control, future-query recovery, and full lifecycle energy remain separate; Fixture F-007 carries the operator-qualified measure–infer–intervene–monitor–route–retain contract across Candidates 001/006/007/010/014/017/018 without promoting a principle or candidate
semiconductor device and circuit reliability	qualification, hierarchical yield and variability, reversible drift, cumulative degradation, abrupt failure, transient upset, fault geometry, correction/recovery stages, adaptive margin, analog/in-memory physical state, endurance/wear, accepted field service, and fabrication-to-retirement burden remain separate; Fixture F-008 carries the mission-profile-qualified degradation/recovery contract across Candidates 001/005/006/009/010/012/014/017/018 without promoting a principle or candidate
acoustics, hearing, and auditory-scene analysis	calibrated pressure/level and room support, cochlear filtering/compression, temporal coding, spatial cues, masking/grouping/separation, reverberation/context, sparse activity, efferent control, echolocation and active emission/reception, array inference, uncertainty, calibration, and complete lifecycle cost remain separate; Fixture F-009 carries the operator- and action-qualified acoustic contract across Candidates 002/006/007/009/012/014 without promoting a principle or candidate
information thermodynamics and physical computation	46 established, 5 plausible, and 1 disputed claim keep logical transformation, device, circuit, computer/workload, facility, and lifecycle quantities separate; finite-time/error/reservoir/controller costs, reversible history, thermal noise, information engines, thermodynamic uncertainty, distribution mismatch, data movement, PUE, and embodied burden refine P-001/P-003/P-006/P-007/P-008/P-009/P-010/P-012/P-013; Fixture F-010 carries the boundary-qualified physical-computation contract across Candidates 001/005/006/009/010/012/014/017/018 without promoting a principle or candidate
olfaction, chemical sensing, and plume tracking	50 established, 1 plausible, and 1 disputed claim keep receptor/sensor coverage, identity, concentration, mixtures, normalization, sniff/pump dynamics, adaptation, turbulent transport, active search, sparse and drifting representations, valence, cross-reactive arrays, humidity, poisoning, analytical confirmation, exposure, safety, and lifecycle cost separate; Fixture F-011 carries the operator-qualified active chemical-sensing contract across Candidates 002/006/007/009/010/012/014/017/018 without promoting a principle or candidate

The music, library/archival, comparative-cognition, visual-design, mathematical-practice, fluid-dynamics, sports, optics, semiconductor, acoustics, information-thermodynamics, and olfaction/chemical-sensing queues are closed by explicit no-promotion decisions. The music Fixture F-001 retains E-MUSIC-07 as a cross-candidate benchmark; the archival refinement stays inside Candidate 017's query-registered preservation track; and comparative cognition contributes Fixture F-003 as an opportunity/history audit across Candidates 004/006/007/014/018/019; and visual design contributes Fixture F-002 as a versioned reconstruct–externalize–inspect–transform–evaluate audit across Candidates 004/014/017/018/019. It treats fixation and negative transfer as measured failure modes, qualifies evaluators and material evidence, and keeps lineage separate from truth, intent, and recoverability; and mathematical practice contributes Fixture F-004 across Candidates 004/009/010/011/014/017/019. It keeps proposal provenance separate from correctness, requires counterexamples and proof-DAG reconstruction, preserves typed lifecycle state, prevents dependency leakage, propagates reverse-dependency invalidation, and includes proof, checker, human, rebuild, and joule cost; and fluid dynamics contributes Fixture F-005 across Candidates 002/003/006/007/012/014. It requires signed scale transfer, tail fidelity, detector-qualified structures, explicit closure/model support, objective-qualified reduction, total adaptive work, actual observation and sensor operators, observability and posterior calibration, stable net-benefit control, vector mixing outcomes, transition history, extreme calibration, propagated uncertainty, and a full lifecycle energy boundary. Its null is the complete relevant CFD, reduced-model, estimation, sensing, control, rare-event, metrology, and accounting stack. Sports contributes Fixture F-006 across Candidates 002/004/006/007/009/012/014/019. It preserves actual observation/action coupling; separates anticipation from interception, practice from retention and transfer, relevant exploration from variance, and fatigue/readiness/damage from staged return; challenges coordination with turnover and perturbation, deceptive cues with calibration, selection with counterfactual opportunity, and apparent efficiency with full athlete, coach, facility, sensor, recovery, medical, failed-pathway, wall-time, and joule budgets. Its null is the complete calibrated prediction, retrieval, planning, control, curriculum, assurance, coordination, causal-inference, and lifecycle- accounting stack. Optics contributes Fixture F-007 across Candidates 001/006/007/010/014/017/018. It binds inferred state to the physical operator, acquisition action, photons/noise, calibration, ambiguity, prior support, saturation, typed uncertainty, route eligibility, retained-query obligation, and complete resource boundary. Its adversarial tracks test null- space honesty, prior mismatch, active photon allocation, multiplex crossover, drift attribution and fallback, optical/digital service curves, multi-device fabrication and thermal variation, and physical compaction under future queries. Its null is the complete inverse-method, experiment-design, control, fusion, digital/analog acceleration, calibration, and lifecycle-accounting stack. The first T1 slice is now executable as the F-007 null-space-honesty smoke harness: it must expose false specificity, accept justified abstention, and exactly tie the complete mature active null rather than manufacture a novelty advantage. Semiconductor reliability contributes Fixture F-008 across Candidates 001/005/006/009/010/012/014/017/018. It binds accepted service to manufactured-unit hierarchy, mission profile, mechanism-specific acceleration, censoring, monitor age, electrothermal concentration, fault geometry, correction and recovery state, protected exact boundaries, physical- weight age, endurance, and fabrication-

through-retirement resources. Its ten tracks test hierarchical transfer, accelerated-life support, sparse wear, correlated faults, voltage authority, approximation containment, analog/digital crossover, nonideality transfer, wear placement, and keep/repair/repurpose/ replace policy against the complete mature reliability stack. Acoustics contributes [Fixture F-009](#) across Candidates 002/006/007/009/012/014. It binds each result to source and emission action, propagation/room state, receiver geometry and motion, calibrated operator support, context/history, literal target, uncertainty, independent unit, feasible action, exposure, and complete resource boundary. Its nine tracks test level-dependent front ends, sparse event timing, active emission/reception, localization under operator shift, reverberation transfer, grouping/separation, efferent-like gain, multi-emitter interference, and full lifecycle cost against the complete calibrated acoustics, inference, control, spectrum, uncertainty, and accounting stack. Information thermodynamics contributes [Fixture F-010](#) across Candidates 001/005/006/009/010/012/014/017/018. It binds every physical- efficiency result to one accepted service while keeping logical, device, circuit, workload, facility, and lifecycle quantities distinct. Its tracks challenge lower-bound substitution, finite-time and finite-error operation, reversible-history closure, thermal reliability, measurement/controller reset, uncertainty-relation applicability, distribution mismatch, topology, data movement, PUE allocation, and fabrication-to-retirement ranking. Its null is the complete compression, reversible-computation, stochastic-thermodynamics, circuit/architecture, workload-metering, facility, and lifecycle-assessment stack. Olfaction and chemical sensing contribute [Fixture F-011](#) across Candidates 002/006/007/009/010/012/014/017/018. It binds source and release, turbulent or reactive transport, realized sniff/pump/motion, receiver path, dynamic response and recovery, cross-sensitivity, calibration and exposure history, operator support, literal outcome, feasible action, safety, independent unit, and complete lifecycle cost. Its fourteen tracks test concentration–identity factorization, sensor coverage, normalization, deconvolution, active acquisition, bilateral/serial/wind cues, plume statistics, search policy, embodiment, sparse total cost, drifting representations, mixture masking, sensor poisoning, analytical confirmation, and safe tiered action against the complete calibrated sensing, analytical, estimation, control, standards, human-work, consumable, and lifecycle stack. The next breadth field is selected by expected ability to split, merge, reject, or re-baseline a current mechanism. Breadth alone does not close a row.

Efficiency mechanism

Stage gates reduce the cost of compounding weak mechanisms. Shared principle bundles reuse telemetry, baselines, and test regimes; isolated rejection avoids paying for full-system integration; exact provenance prevents repeated searches; and negative results retire duplicated work.

For stage k , let $E_{k,\text{research}}$ and $E_{k,\text{integration}}$ be joules spent within declared research and compute boundaries. Let $p_{k,\text{survive}}$ be the empirically estimated fraction of candidates that pass the isolated gate. The expected integration energy avoided by testing first is

$$E_{k,\text{avoided}} = (1 - p_{k,\text{survive}})E_{k,\text{integration}} - E_{k,\text{research}}.$$

This quantity is reported only when both energy boundaries are actually measured; its purpose is to expose the cost of evaluation, not to assume that research automatically saves energy.

Evidence status

- The ordering and gates are project decisions.
- Mechanism-specific evidence remains in the [claims ledger](#); roadmap placement does not promote a claim.
- The candidate and fixture catalogues contain written contracts and bounded development harnesses, but no claim-eligible result.
- The six-stage composition and its lifecycle benefit remain speculative until the gates produce measured results.
- Sensorimotor transfer remains speculative under C-007.

Speculative extensions

- Maintain a machine-readable experiment ledger linked to commits, hardware records, raw telemetry, derived figures, and claim-status changes.
- Use active literature search to select the next field by expected impact on an unresolved principle, not by publication volume.
- Allocate compute to experiments by expected decision value and uncertainty reduction once the proxy is calibrated against actual outcomes.
- Run adversarial review agents whose only task is to find conventional explanations, hidden costs, or domain evidence that breaks a principle.
- Publish negative results and retired translations as reusable research outputs rather than silently deleting them.

Failure modes

- **Perpetual breadth:** new fields accumulate without changing principles or experiments.
- **Premature composition:** several plausible mechanisms are integrated before any one survives its null model.
- **Gate drift:** thresholds or comparison boundaries change after results are visible.
- **Mechanism creep:** a component enters the system without a claim, principle, ablation, or owner.
- **Proxy capture:** FLOPs, parameter count, or average accuracy substitutes for measured physical cost and qualified behavior.
- **Negative-result loss:** a failed translation disappears and later returns under another domain name.
- **Benchmark enclosure:** the roadmap optimizes a convenient task family that no longer tests the intended operation.
- **Documentation lag:** equations, diagrams, or claim status stop matching the executed artifact.
- **Hardware lock-in:** a mechanism is tuned to one substrate before its functional interface is stable.
- **Narrative immunity:** biological appeal preserves a mechanism after its measurable predictions fail.

Measurable predictions

1. Each breadth wave changes at least one principle boundary, null model, experiment regime, or explicit no-change decision; otherwise the sampling policy is revised.
2. The ratio of scoped claims to deduplicated principles increases as coverage grows, while duplicate organism-themed components decline.
3. At least one Stage-1 candidate ties a conventional baseline and is merged, and at least one loses and is retired; a roadmap in which every mechanism wins is not discriminating.
4. Shared telemetry and null-model infrastructure reduce the incremental cost of testing later principles relative to the first isolated contracts.
5. Mechanisms that pass isolated gates still sometimes fail composition, producing explicit interaction terms rather than post-hoc stories.
6. Lifecycle accounting reverses at least one apparent saving based on FLOPs, parameter count, or inference-only energy.
7. Every completed stage leaves reproducible baselines, scoped claim changes, negative results, and a smaller set of unresolved integration decisions.

Boundary-qualified physical-computation contract

This note defines the quantitative boundary for Fixture F-010. It operationalizes the durable result of the information thermodynamics and physical computation audit: a fundamental lower bound, a device transition, a circuit, a workload, a facility, and a hardware lifecycle answer different questions and cannot be substituted for one another.

- **Status:** fixture mathematics; no new principle or candidate
- **Comparison unit:** one preregistered useful-task service contract evaluated over a declared hardware lifecycle
- **Primary rule:** report energy, time, error, stability, uncertainty, and accepted outcomes jointly at every boundary used in a claim

Identity and useful-task contract

For arm a , hardware instance h , workload episode e , and measurement interval r , seal the immutable identity

$$I_{a,h,e,r} = (a, h, e, r, v^{\text{hw}}, v^{\text{sw}}, v^{\text{cal}}, s, t_0, t_1),$$

where a is an arm identifier [identifier], h is a physical hardware identifier [identifier], e is an episode identifier [identifier], r is a meter interval identifier [identifier], v^{hw} is hardware and firmware version [identifier], v^{sw} is software, compiler, model, and configuration version [identifier], v^{cal} is instrument and calibration version [identifier], s is site [identifier], and t_0, t_1 are interval endpoints [s] with duration $\tau_r = t_1 - t_0$ [s]. Repair, recalibration, replacement, remapping, or version change creates a linked new identity rather than overwriting history.

Let requested outcome j have preregistered service vector

$$R_j = (q_j, L_j^{\text{max}}, \rho_j^{\text{max}}, \Theta_j, \tau_j^{\text{ret}}),$$

where q_j is the required task-quality vector in declared task-native units, L_j^{max} is maximum allowed end-to-end latency [s], ρ_j^{max} is a vector of maximum allowed failure and escaped-harm probabilities [failure/request], Θ_j is minimum useful throughput [request/s], and τ_j^{ret} is required state-retention horizon [s]. Define

$$A_j = \mathbf{1}[Q_j \succeq q_j \wedge L_j \leq L_j^{\text{max}} \wedge \rho_j \preceq \rho_j^{\text{max}} \wedge \Theta \geq \Theta_j],$$

where $A_j \in \{0, 1\}$ is accepted-outcome status, Q_j is measured quality in the same units as q_j , L_j is measured latency [s], ρ_j is the measured risk vector [failure/request], Θ is delivered throughput [request/s], $\succeq$ and $\preceq$ mean every registered component passes its direction, and $\mathbf{1}[\cdot]$ is an indicator [dimensionless]. Let

$$N_{\text{acc}} = \sum_{j=1}^{N_{\text{req}}} A_j, \quad f_{\text{acc}} = \frac{N_{\text{acc}}}{N_{\text{req}}},$$

where N_{req} is requested outcomes [request], N_{acc} is accepted outcomes [accepted outcome], and f_{acc} is accepted fraction [dimensionless]. Rejected, abstained, timed-out, silently corrupted, retried, and safety-blocked requests remain in N_{req} and the resource ledger.

Six-boundary energy vector

For a sealed comparison unit, report

$$\mathbf{E} = (E^{\text{fund}}, E^{\text{dev}}, E^{\text{circ}}, E^{\text{IT}}, E^{\text{fac}}, E^{\text{emb}}) \quad [\text{J}],$$

where E^{fund} is a theorem-qualified lower bound for the declared information operation [J], E^{dev} is measured energy crossing the device terminals [J], E^{circ} is measured or calibrated energy of the complete circuit and controls [J], E^{IT} is metered IT energy for the complete workload [J], E^{fac} is allocated facility energy [J], and E^{emb} is allocated fabrication-to-retirement energy [J]. The vector is not a sum: boundaries can be nested. A report must state whether $E^{\text{dev}} \subset E^{\text{circ}} \subset E^{\text{IT}} \subset E^{\text{fac}}$ for its meters.

Per-accepted-outcome intensity at boundary b is

$$e^b = \frac{E^b}{N_{\text{acc}}} \quad [\text{J/accepted outcome}], \quad b \in \{\text{dev, circ, IT, fac, emb, life}\}.$$

If $N_{\text{acc}} = 0$, e^b is undefined and the arm fails; it is not reported as zero. **fund** is excluded from this normalization unless the logical operation and accepted outcome have an explicit registered multiplicity.

Fundamental information-operation boundary

Logical loss and generalized erasure

Let X be the input logical state, Y the retained logical output, and S usable side information, all discrete random variables [state]. With natural logarithms, define

$$H(X | Y, S) = - \sum_{x,y,s} p(x,y,s) \ln p(x | y, s) \quad [\text{nat}],$$

where $p(x, y, s)$ is the joint probability [dimensionless]. $H(X | Y, S)$ records input distinctions unavailable from retained output and side state. It is not automatically heat; a physical encoding and protocol are still required.

For physical microstate $z \in \mathcal{Z}$, probability $p(z)$ [dimensionless], Hamiltonian $\mathcal{H}(z)$ [J], bath temperature T [K], and Boltzmann constant $k_B = 1.380649 \times 10^{-23}$ J/K, define nonequilibrium free energy

$$\mathcal{F}[p, \mathcal{H}] = \sum_{z \in \mathcal{Z}} p(z) \mathcal{H}(z) + k_B T \sum_{z \in \mathcal{Z}} p(z) \ln p(z) \quad [\text{J}].$$

For an isothermal transformation under the assumptions registered by the selected theorem, expected work performed on the system obeys

$$\langle W_{\text{on}} \rangle \geq \Delta \mathcal{F} = \mathcal{F}[p_1, \mathcal{H}_1] - \mathcal{F}[p_0, \mathcal{H}_0] \quad [\text{J}],$$

where p_0, p_1 are initial and final microstate distributions, $\mathcal{H}_0, \mathcal{H}_1$ are initial and final Hamiltonians [J], and W_{on} is work on the system [J]. The protocol class, bath, initial state, correlations, cycle closure, and controls are part of the theorem.

For cyclic reset of a degenerate, uniformly random binary memory with symmetric final error probability $\epsilon \in [0, 1/2]$, the special case is

$$E_{\text{reset}}^{\text{fund}}(T, \epsilon) = k_B T [\ln 2 - h(\epsilon)] \quad [\text{J}],$$

where

$$h(\epsilon) = -\epsilon \ln \epsilon - (1 - \epsilon) \ln(1 - \epsilon) \quad [\text{nat}]$$

is binary entropy. At $\epsilon = 0$, this becomes $k_B T \ln 2$. For a biased input, nondegenerate memory, correlated side state, finite reservoir, or noncyclic operation, use the applicable generalized bound rather than this special case.

Finite time, error, and state stability

For protocol π with duration τ_π [s], define empirical excess work

$$W_\pi^{\text{ex}} = \langle W_{\text{on}, \pi} \rangle - \Delta \mathcal{F} \quad [\text{J}].$$

W_π^{ex} is compared only among protocols with matched initial and final physical distributions, error definition, bath, and controls. The joint protocol outcome is

$$\mathbf{g}\pi = (\langle W_{\text{on}, \pi} \rangle, \tau_\pi, \epsilon_\pi, p_\pi^{\text{tail}}) \quad [\text{J}, \text{s}, 1, 1],$$

where ϵ_π is mean logical error [error/transition] and p_π^{tail} is a registered high-work or harmful-event probability [event/transition]. No coordinate may be silently scalarized.

For one activated bistable-memory null,

$$\tau_{\text{ret}} = \tau_0 \exp\left(\frac{\Delta U}{k_B T}\right) \quad [\text{s}], \quad p_{\text{loss}}(t) = 1 - \exp\left(-\frac{t}{\tau_{\text{ret}}}\right),$$

where τ_0 is attempt time [s], ΔU is effective barrier [J], τ_{ret} is mean retention time [s], t is storage time [s], and p_{loss} is loss probability [loss/stored state]. The equation is a registered activated-process null, not a universal retention law.

Define error-consequence energy

$$E^{\text{err}} = E^{\text{detect}} + E^{\text{correct}} + E^{\text{retry}} + E^{\text{fallback}} + E^{\text{lost service}} \quad [\text{J}],$$

where the terms are measured detection, correction, retry, fallback, and allocated lost-service energy [J]. Harm and task loss not expressible in joules remain separate registered coordinates.

Nonequilibrium, feedback, and uncertainty-relation scope

For repeated realizations initially in canonical equilibrium at inverse temperature $\beta = (k_B T)^{-1}$ [1/J], a registered Jarzynski test uses

$$\hat{J} = \frac{1}{N_\pi} \sum_{i=1}^{N_\pi} \exp(-\beta W_i), \quad J_0 = \exp(-\beta \Delta F),$$

where N_π is independent protocol realizations [realization], W_i is work on realization i [J], ΔF is equilibrium free-energy change [J], and $\hat{J}$, J_0 are dimensionless. Report the work distribution, rare-event coverage, dependence diagnostics, and uncertainty of $\hat{J}$; a single $W_i < \Delta F$ is not a violation.

For feedback measurement record M and controlled state X , let mutual information be

$$I(X; M) = \sum_{x,m} p(x, m) \ln \frac{p(x, m)}{p(x)p(m)} \quad [\text{nat}].$$

The joint feedback ledger is

$$E_{\text{feedback}}^{\text{joint}} = E^{\text{plant}} + E^{\text{sense}} + E^{\text{record}} + E^{\text{control}} + E^{\text{actuate}} + E^{\text{reset}} \quad [\text{J}],$$

where every term is energy crossing the declared plant, sensor, record memory, controller, actuator, or reset boundary [J]. Extracted work from the plant is reported with sign and cannot cancel unmeasured controller work.

For a stationary continuous-time Markov jump model and a registered integrated current J_t over duration t [s], the original steady-state thermodynamic uncertainty relation is tested as

$$\mathcal{U}_t = \frac{\text{Var}(J_t)}{\langle J_t \rangle^2} \Sigma_t \geq 2,$$

where Σ_t is expected total entropy production in units of k_B [dimensionless], and $\mathcal{U}_t$ is dimensionless. Before evaluating it, register the current, transition graph, Markov property, stationarity, time-reversal convention, observation completeness, and estimator for Σ_t . A finite-time, transient, non-Markovian, deterministic, or quantum claim requires its own cited inequality and assumptions; failure of this scope test blocks the inference.

Device, circuit, and memory boundaries

Measured device transition

For device transition k over interval $[t_k^0, t_k^1]$, terminal energy is

$$E_k^{\text{dev}} = \sum_{c=1}^{C_k} \int_{t_k^0}^{t_k^1} V_{k,c}(t) i_{k,c}(t) dt \quad [\text{J}],$$

where C_k is the number of terminals or supplied channels [channel], $V_{k,c}$ is measured potential [V], $i_{k,c}$ is signed current [A], and time t is [s]. Instrument bandwidth, phase, probe loading, integration rule, calibration covariance, and recovered-energy sign are registered. Heat requires an independent calorimetric or validated thermodynamic inference; terminal electrical energy is not relabeled as heat.

For conventional capacitive switching, the registered null is

$$E^{\text{dyn}} = \alpha C_{\text{eff}} V^2 N_{\text{cyc}} \quad [\text{J}],$$

where α is mean activity per cycle [transition/cycle], C_{eff} is effective switched capacitance [F], V is supply voltage [V], and N_{cyc} is cycles [cycle]. Short-circuit, leakage, clock, interconnect, and control energy are additional measured terms.

For an idealized adiabatic RC path, use the scoped model

$$E^{\text{adiabatic}}(\tau) = \gamma \frac{RC}{\tau} CV^2 + P_{\text{leak}}\tau + E^{\text{clock}}(\tau) + E^{\text{control}} + E^{\text{I/O}} + E^{\text{reset}} \quad [\text{J}],$$

where R is effective resistance [ohm], C is capacitance [F], τ is transition time [s] with registered slow-ramp support, γ is a waveform-dependent coefficient [dimensionless], P_{leak} is leakage power [W], and the remaining terms are measured clock, control, input/output, and reset energies [J]. A real crossover exists at operating point o only if

$$E^{\text{adiabatic}}(o) < E^{\text{ordinary}}(o)$$

at matched task quality, transition error, useful throughput, area or hardware budget, temperature, and complete cyclic state.

Logical reversibility and closed history

For a reversible arm, let B^{anc} be prepared ancilla bits [bit], B^{hist} be retained history [bit], B^{out} be preserved output [bit], and B^{garb} be garbage remaining before closure [bit]. The run closes only when each non-output state is assigned exactly one action:

$$B^{\text{anc}} + B^{\text{hist}} + B^{\text{garb}} = B^{\text{uncompute}} + B^{\text{retain}} + B^{\text{export}} + B^{\text{erase}} \quad [\text{bit}],$$

where the right-hand terms are uncomputed, deliberately retained, exported, and erased bits [bit]. Each action carries circuit, movement, stability, and eventual reset energy. Equality is a bookkeeping conservation rule, not a claim that all logical states are independent or uniformly random.

Retention and correction ledger

For memory tier m , define

$$E_m^{\text{memory}} = N_m^{\text{w}} e_m^{\text{w}} + N_m^{\text{r}} e_m^{\text{r}} + N_m^{\text{ref}} e_m^{\text{ref}} + E_m^{\text{ECC}} + E_m^{\text{scrub}} + E_m^{\text{move}} + E_m^{\text{idle}} \quad [\text{J}],$$

where N_m^{w} , N_m^{r} , and N_m^{ref} are write, read, and refresh counts [operation]; e_m^{w} , e_m^{r} , and e_m^{ref} are measured energy per respective operation [J/operation]; and the remaining terms are error-correction, scrubbing, movement, and idle energy [J]. Report raw bit errors, detected uncorrectable errors, miscorrections, silent corruption, retries, endurance, retention distribution, and accepted retrievals separately.

Workload and data-movement boundary

Partition the implemented workload into physical hierarchy links $\ell \in \mathcal{L}$, including register, local memory, cache, on-package, off-package memory, host, storage, and network paths. Define

$$E^{\text{move}} = \sum_{\ell \in \mathcal{L}} B_{\ell} \hat{e}_{\ell}(B_{\ell}, d_{\ell}, p_{\ell}, o_{\ell}) \quad [\text{J}],$$

where B_{ℓ} is bytes transferred on link ℓ [byte], d_{ℓ} is physical or logical distance class [class], p_{ℓ} is precision and encoding [bit/value and identifier], o_{ℓ} is the operating point containing voltage, temperature, rate, and utilization [registered tuple], and $\hat{e}_{\ell}$ is a measured energy model [J/byte] with a coverage interval. A component table from another process or workload may be a prior but not a measurement.

For routed or sparse workload episode e , let

$$E_e^{\text{IT}} = \int_{t_e^0}^{t_e^1} P_e^{\text{IT}}(t) dt \quad [\text{J}],$$

where $P_e^{\text{IT}}(t)$ is metered IT power [W], and t_e^0, t_e^1 are episode boundaries [s]. The declared IT boundary contains compute, memory, interconnect, storage and network shares, host orchestration, routing metadata, load imbalance, idle allocation, conversion, correction, calibration, rejected work, and retries. Diagnostic decomposition is

$$E_e^{\text{IT}} = E_e^{\text{arith}} + E_e^{\text{move}} + E_e^{\text{route}} + E_e^{\text{sync}} + E_e^{\text{convert}} + E_e^{\text{idle}} + E_e^{\text{maint}} + E_e^{\text{retry}} \quad [\text{J}],$$

where every right-hand term is an allocated measured or calibrated energy [J]. The equality is checked against the top-level meter within registered closure tolerance δE [J]; an unclosed residual remains explicit.

Let requested arithmetic count be N_e^{OP} [operation], useful bytes be B_e^{use} [byte], routed candidates be N_e^{route} [candidate], and active hardware-time capacity be

$$C_e^{\text{cap}} = \sum_{u=1}^U n_{u\tau_{e,u}} \quad [\text{device s}],$$

where U is hardware class count [class], n_u is provisioned device count [device], and $\tau_{e,u}$ is reserved wall time [s]. Slower execution and idle replicas are therefore not free when throughput is held constant.

Facility and cooling boundary

For facility interval r , measure

$$E_r^{\text{fac}} = \int_{t_0}^{t_1} P_r^{\text{fac}}(t) dt, \quad E_r^{\text{IT}} = \int_{t_0}^{t_1} P_r^{\text{IT}}(t) dt \quad [\text{J}],$$

where P_r^{fac} is total data-centre facility power [W] and P_r^{IT} is IT-equipment power [W] under the registered ISO/IEC 30134-2 measurement category and boundaries. Power usage effectiveness is

$$\text{PUE}_r = \frac{E_r^{\text{fac}}}{E_r^{\text{IT}}} \quad [\text{dimensionless}].$$

For a task cohort c sharing interval r , allocated facility energy is

$$E_{c,r}^{\text{fac}} = E_{c,r}^{\text{IT}} + w_{c,r} (E_r^{\text{fac}} - E_r^{\text{IT}}) \quad [\text{J}],$$

where $E_{c,r}^{ll}$ is directly metered or allocation-qualified cohort IT energy [J], and $w_{c,r} \in [0, 1]$ is a preregistered overhead-allocation weight with $\sum_c w_{c,r} = 1$. At minimum, test IT-energy, peak-demand, space/capacity, and direct cooling-submeter allocation cases. Multiplying an episode by a generic PUE is not a confirmatory measurement.

Cooling diagnostics report

$$E_r^{\text{cool}} = E_r^{\text{chiller}} + E_r^{\text{fan}} + E_r^{\text{pump}} + E_r^{\text{tower}} + E_r^{\text{control}} \quad [\text{J}],$$

where the terms are chiller, fan, pump, heat-rejection, and cooling-control energy [J]. Ambient dry-bulb and wet-bulb temperatures [K], humidity [dimensionless], supply/return temperatures [K], flow [m^3/s], utilization [dimensionless], and site are held or modeled explicitly. PUE is not carbon, water, task quality, or a cooling coefficient of performance.

Embodied lifecycle boundary

For hardware cohort h , define cradle-to-retirement primary-energy inventory

$$E_h^{\text{life}} = E_h^{\text{fab}} + E_h^{\text{pack}} + E_h^{\text{transport}} + E_h^{\text{deploy}} + E_h^{\text{op}} + E_h^{\text{maint}} + E_h^{\text{replace}} + E_h^{\text{EOL}} \quad [\text{J}],$$

where the terms are allocated fabrication, packaging, transport, deployment, operation including facility share, maintenance, replacement, and end-of-life primary energy [J]. Credits, if allowed by the preregistered lifecycle method, are signed and shown separately.

Let $Y_h \in (0, 1]$ be accepted packaged yield [accepted device/started device], N_h^{start} be started units [device], N_h^{life} be lifetime accepted task outcomes [accepted outcome], and u_h be useful utilization [useful device-second/provisioned device-second]. The lifecycle intensity is

$$e_h^{\text{life}} = \frac{E_h^{\text{life}}}{N_h^{\text{life}}} \quad [\text{J/accepted outcome}],$$

with N_h^{life} estimated only over registered deployment demand, support lifetime, failure, maintenance, retirement, and replacement policies. Yield and utilization are reported rather than absorbed into an optimistic denominator.

For new specialized hardware s versus an already available conventional arm c , the operational-energy break-even count is

$$N^* = \frac{E_s^{\text{emb}} - E_c^{\text{incremental emb}}}{e_c^{\text{op}} - e_s^{\text{op}}} \quad [\text{accepted outcome}],$$

when $e_c^{\text{op}} > e_s^{\text{op}}$. Here E_s^{emb} is newly incurred embodied energy [J], $E_c^{\text{incremental emb}}$ is additional embodied energy incurred by the conventional option [J], and $e_c^{\text{op}}, e_s^{\text{op}}$ are facility-inclusive operational intensities [J/accepted outcome]. If the denominator is nonpositive, no positive energy break-even exists. N^* is reported as a distribution under yield, utilization, service-life, demand, and allocation uncertainty.

Climate, water, material criticality, toxicity, and labor are separate outcome coordinates. For greenhouse-gas inventory,

$$G_h = \sum_{g=1}^G a_{h,g} \chi_g \quad [\text{kg CO}_2\text{e}],$$

where $a_{h,g}$ is activity amount in its declared inventory unit, χ_g is the geography-, time-, and pathway-qualified characterization factor [kg CO₂e/inventory unit], and G is inventory-flow count [flow]. Energy alone does not determine G_h .

Uncertainty, support, and matched comparison

For reported outcome y [native unit], decompose its estimator as

$$\hat{y} = y + b^{\text{meter}} + b^{\text{model}} + b^{\text{alloc}} + \varepsilon,$$

where b^{meter} is meter/calibration bias [native unit], b^{model} is model-form or extrapolation bias [native unit], b^{alloc} is shared-resource allocation effect [native unit], and ε is repeatability variation [native unit]. Report a coverage or credible interval for y , calibration lineage, covariance where quantities share meters or models, and sensitivity across registered allocation and lifecycle cases. An interval for repeatability alone is not total uncertainty.

Let x be an episode/regime feature vector in registered native units and $\mathcal{S}_{\text{val}}$ be validation support. Define a preregistered support distance

$$d_{\text{sup}}(x) = \inf_{z \in \mathcal{S}_{\text{val}}} \|D^{-1}(x - z)\|_2 \quad [\text{dimensionless}],$$

where D is a diagonal matrix of fixed feature scales in the same units as x , and $\|\cdot\|_2$ is Euclidean norm. Authority is withheld when $d_{\text{sup}}(x) > d_{\text{max}}$, where d_{max} is a sealed dimensionless threshold. Other support tests are allowed only when specified before the held-out release.

The primary outcome vector for arm a is

$$\mathbf{Y}_a = (f_{\text{acc}}, Q, L_{0.50}, L_{0.99}, \rho, e^{\text{IT}}, e^{\text{fac}}, e^{\text{life}}, E^{\text{err}}, C^{\text{cap}}, G, W, M),$$

where f_{acc} is accepted fraction [dimensionless], Q is registered task quality [task-native units], $L_{0.50}, L_{0.99}$ are median and 99th percentile latency [s], ρ is the registered risk vector [failure/request], the e terms are energy intensity [J/accepted outcome], E^{err} is error-consequence energy [J], C^{cap} is capacity use [device s], G is greenhouse-gas inventory [kg CO₂e], W is water inventory [m³], and M is a material/labor burden vector in declared native units. The vector is not reduced to one score after observing results.

Arm p Pareto-dominates null n only if its simultaneous uncertainty region is no worse on every hard-gated coordinate and strictly better on at least one preregistered primary coordinate under every required sensitivity case. Let

$$D_{p,n} = 1$$

denote that decision [dimensionless], and $D_{p,n} = 0$ otherwise. A component energy win with worse quality, risk, latency, capacity, or another required boundary cannot set $D_{p,n} = 1$.

The illustrative simultaneous-decision figure shows the uncertainty regions and hard gates without assigning measured values to any system.

Equal-budget constraints

For resource $r \in \mathcal{R}$, require

$$B_{a,r} \leq B_r^{\max},$$

where $B_{a,r}$ is arm- a consumption in the native unit of resource r and $B_r^{\max}$ is the shared ceiling in that unit. The registered resource set is

$\mathcal{R} = \{\text{data, design work, fabrication, area, memory, sensors, controls, reserve, training compute, wall time, capacity, operational energy, maintenance, replacement}\}.$

Each resource has its own unit; unlike quantities are never summed. If an arm uses less of a capped resource, the unused amount remains reported and is not converted into post-hoc credit. If an arm violates any hard ceiling, its result is infeasible rather than penalized by a chosen scalar.

Required nulls and ablations

The complete null stack contains, when technically compatible:

1. source/channel coding, compression, quantization, pruning, batching, memoization, caching, compiler elimination, and recomputation;
2. clock and power gating, dynamic voltage/frequency scaling, near-threshold operation, mixed precision, structured sparsity, tiling, data reuse, and hierarchy-aware placement;
3. reversible logic with closed ancilla/history accounting, adiabatic or energy-recovery logic with measured power clock, and conventional logic at matched throughput and process;
4. ECC, checksums, retry, checkpoint/replay, guardbands, calibration, redundancy, and abstention;
5. matched digital, analog, in-memory, optical, and neuromorphic implementations including conversion, communication, control, drift, thermal, and host work;
6. direct facility metering and registered shared-overhead allocations; and
7. ISO 14040/14044 lifecycle cases with common functional unit, yield, utilization, lifetime, replacement, geography, and uncertainty.

Ablations remove exactly one of: generalized physical-state modeling; finite-time optimization; finite-error accounting; retention/correction; closed reversible history; power-clock recovery; feedback-controller boundary; TUR scope gate; hierarchy-aware routing; facility metering; embodied inventory; or uncertainty/support gating. Recalibrate each ablation only within the same development budget. An ablation that makes an arm infeasible is recorded as such, not silently retuned with additional resources.

Held-out regimes and hard retirement

Confirmation splits group by physical device, fabrication cohort, circuit and clock instance, software/model version, workload family, data-layout and hierarchy regime, task shift, transition duration, final error target, temperature, retention horizon, sensor/controller version, facility/site, season, electricity case, and future time. Random transitions or requests from the same group are development diagnostics only.

Retire the broad physical-efficiency composition if any of the following holds:

1. a claimed lower bound lacks its state distribution, Hamiltonian, bath, correlations, final error, duration, or cycle boundary;
2. a device advantage disappears when waveform source, parasitics, control, correction, and full transition closure are measured;
3. a reversible advantage excludes history, ancillae, output preservation, uncomputation, retention, export, or eventual erasure;
4. an adiabatic advantage disappears at matched useful throughput, hardware capacity, error, and leakage-inclusive power-clock cost;

5. a feedback or information-engine gain disappears when sensing, memory, control, actuation, and reset share one boundary;
6. a thermodynamic uncertainty inference fails its process, current, stationarity, observation, or entropy-production scope test;
7. a memory advantage fails the required retention, endurance, correction, silent-corruption, or replacement contract;
8. arithmetic savings are offset by routing, data movement, synchronization, conversion, imbalance, or idle capacity;
9. a facility claim uses component power, TDP, or a generic PUE instead of calibrated interval evidence;
10. lifecycle superiority depends on an unsupported yield, utilization, lifetime, demand, allocation, electricity, or replacement assumption;
11. no Pareto gain survives the strongest compatible null stack, held-out regimes, and required uncertainty sensitivities; or
12. the result lowers quality, safety, latency, retention, or coverage relative to the sealed useful-task contract.

Passing this contract supplies evidence only for the already named candidate scope in Fixture F-010. It creates no project-wide claim, principle, or candidate by itself.

Burden-qualified contestable decisions

Scope

This note formalizes the narrow systems residue from the legal evidence/procedure audit: evidence state, authority, burden, procedure, decision, review, remedy, and finality remain distinct. The legal rules are jurisdiction- and date-bound; the transfer is an experiment contract, not a claim that adjudication is an optimizer or a truth oracle. Its evidence boundaries are C-679–C-704.

Decision record

For decision version v , retain

$$\mathcal{D}_v = (J, A, P, H, B, S, R, O, Q, D, G, V, F),$$

where J is jurisdiction/rule authority, A authorized decision-maker, P parties or affected interests, H claim and requested action, B burden bearer, S governing standard/decision rule, R admitted record, O objections and preserved issues, Q disclosure/access/response state, D disposition, G stated grounds and reasons, V review/remedy state, and F finality/reopening conditions. These are typed fields, not commensurable scores.

For evidence item i ,

$$e_i = (a_i, p_i, u_i, c_i, \alpha_i, \delta_i, w_i, h_i, o_i),$$

where a_i is artifact identity, p_i proponent, u_i permitted purpose, c_i custody/provenance, α_i authentication ruling, δ_i admissibility ruling, w_i weight or likelihood contribution, h_i supported claim/scope, and o_i objections/contrary evidence. Authentication does not imply content truth; admissibility does not imply weight; weight does not imply sufficiency; sufficiency does not grant authority.

Statistical and authority boundaries

For formal hypotheses H_1, H_0 , the likelihood ratio

$$\Lambda(e) = \frac{p(e \mid H_1)}{p(e \mid H_0)}$$

is dimensionless and depends on the evidence model. It is neither $p(H_1 \mid e)$ nor a universal translation of a verbal proof standard. A decision threshold follows only after hypotheses, priors/evidence distributions, consequence vector, protected constraints, authority, and the decision rule are declared.

Keep the outcome vector visible:

$$\mathbf{L} = (N_{10}, N_{01}, T_{\text{wrong}}, N_{\text{process}}, N_{\text{protected}}, C, TH, TD, E),$$

where false positive/negative counts N_{10} , N_{01} are decisions, T_{wrong} is time under an incorrect or unauthorized effect in person-seconds or effect-seconds, N_{process} and $N_{\text{protected}}$ are violation counts, C is currency at a stated price year, TH human effort in person-seconds, TD elapsed delay in seconds, and E lifecycle energy in joules. Protected procedure cannot be silently traded for mean accuracy through an unstated scalarization.

Review is selected and scoped

For N_0 first decisions, N_A reviewed, and N_R reversed/remanded,

$$r_A = \frac{N_A}{N_0}, \quad r_R = \frac{N_R}{N_A}.$$

Both are dimensionless. r_R is not the first-stage error rate because issue preservation, selection into review, review standard, harmless/prejudice rule, remedy, settlement, and reviewer error intervene. A review record therefore binds the issue, record version, standard, permitted scope, identified error, prejudice rule, disposition, remedy, successor version, and later validity.

Falsification boundary

The held composition loses if typed workflow plus provenance, access control, calibrated selective prediction, rule/citation graphs, independent review, red-team challenge, conflict-of-interest controls, and full recomputation match its protected outcome–cost frontier. Reason text is not credited as faithful causal introspection; finality is not truth; reopening is not free rollback; and similarity to a prior decision does not establish applicable authority.

First testable efficiency model

Purpose

This model compares a conditional architecture with a dense baseline on the same task using directly measured system quantities.

Per-event cost

For event x , let the dense baseline execute modules $1 \dots n$. Its measured energy is

$$ED(x) = ED_{\text{compute}} + ED_{\text{memory}} + ED_{\text{network}} + ED_{\text{host}}.$$

Every ED component is joules per event at the same declared physical boundary. For conditional gates $g_i(x) \in \{0, 1\}$, the candidate uses

$$EC(x) = E_{\text{gate}}(x) + \sum_{i=1}^n g_i(x) E_i(x) + EC_{\text{memory}}(x) + EC_{\text{network}}(x) + EC_{\text{host}}(x).$$

$E_i(x)$ and every other EC component are joules per event at that same boundary; $g_i(x)$ is dimensionless. The net event saving is

$$SE(x) = 1 - \frac{EC(x)}{ED(x)}.$$

SE is dimensionless, and $SE > 0$ indicates lower measured energy. It is reported only for event strata whose quality and risk remain inside the declared equivalence envelope.

Lifecycle cost

Let E_{train} , $E_{\text{consolidate}}$, and E_{mature} be measured one-time lifecycle energies in joules at the same boundary. For a dimensionless count N of qualified deployment events,

$$E_{\text{life/event}} = \frac{E_{\text{train}} + E_{\text{consolidate}} + E_{\text{mature}}}{N} + \frac{1}{N} \sum_{j=1}^N EC(x_j).$$

$E_{\text{life/event}}$ is joules per qualified event; x_j is qualified event j , and $EC(x_j)$ is its measured energy in joules. Pruning or hardening is beneficial over the measured lifetime only when its one-time cost is smaller than the accumulated deployment saving.

Quality equivalence

Candidate C is comparable to baseline D only if

$$QC \geq QD - \epsilon_Q, \quad RC \leq RD + \epsilon_R, \quad LC_{p95} \leq L_{\max}.$$

QC and QD use one declared task-quality unit, so ϵ_Q has that same unit. RC and RD use one declared risk unit, so ϵ_R has that same unit. LC_{p95} and the ceiling $L_{\max}$ are seconds at the 95th percentile. All tolerances are fixed before the experiment. Reporting only average quality is insufficient when conditional compute can fail selectively on rare events. The full comparison record is defined in the [energy-model](#) chapter.

Brain counterfactual: unresolved

The dimensional form

$$P_{\text{counterfactual}} = P_{\text{brain}} \frac{\eta_{\text{brain}}}{\eta_{\text{machine}}}$$

P_{brain} and $P_{\text{counterfactual}}$ are power in watts. η_{brain} and η_{machine} must both be measured in the same qualified functional-output units per joule, making their ratio dimensionless. This requires the same task, quality, risk, time, and accounting boundary. Current “brain FLOP” estimates and accelerator arithmetic do not meet that condition. Consequently claim C-016 remains disputed and no numerical range is derived here.

Initial hypotheses

- **H-E1:** conditional routing reduces device energy at matched quality only above a workload-dependent module granularity.
- **H-E2:** memory and interconnect savings, not arithmetic savings alone, determine whether sparse capacity wins end to end.
- **H-E3:** consolidation increases short-term energy but reduces amortized energy when it replaces frequent full-model updates.
- **H-E4:** structural pruning yields more portable energy savings than unstructured weight sparsity.

These hypotheses become claims only after experiments and ledger review.

History-conditioned modular succession: mathematical contract

- **Status:** frontier notation and experiment design; no result
- **Audit:** history-conditioned modular succession and priority effects
- **Promotion boundary:** this note creates no claim, principle, candidate, protocol, or fixture identifier
- **Purpose:** define fixed-task-and-eligibility order estimands, resource identities, causal-cut contrasts, endpoint units, multiplicity control, and kill boundaries before any large implementation is proposed

Scope

The mathematical question is whether a randomized sequence changes a final learned system after the presented task multiset, eligible module identity set, and all declared resources are held fixed. Realized active, consolidated, merged, or retired module state may differ and is an endpoint. The notation does not assume that an order effect exists or that any effect is ecological in mechanism.

Microbial abundance, model accuracy, allocated parameters, examples, seconds, bytes, and joules remain different quantities. No conversion among them is introduced here.

Indices, objects, and units

SYMBOL	DEFINITION	UNIT
K	number of distinct task blocks and eligible module identities in the bounded design; initially $K = 4$	count
k	paired admission-unit index, $k \in \{1, \dots, K\}$	count
ℓ	executing-module index when tasks and modules are evaluated separately	count
j, j'	sequence-position indices, each in $\{1, \dots, K\}$	count
r	paired random-seed index	count
a	learning-method index	category
q	endpoint index	category
h, h'	distinct mechanism-factor indices when used together	category
$\mathcal{T}$	frozen multiset of task blocks	set-valued
T_k	checksummed task block with identity k	data object
M_k	eligible module identity paired with T_k before randomization	typed identity
$A_k = (T_k, M_k)$	frozen task/module admission unit	typed pair
Π_K	set of all permutations of K identities	set-valued
π	one randomized permutation in Π_K	dimensionless mapping

SYMBOL	DEFINITION	UNIT
π_j	paired admission-unit identity shown at sequence position j	count
$\mathbf{z}$	vector of binary mechanism interventions	dimensionless vector
$s_{a,j}$	complete learned and structural state for method a after position j	typed state
$\theta_{a,j}$	trainable parameter state within $s_{a,j}$	parameter vector
$\rho_{a,j}$	router/admission state within $s_{a,j}$	typed state
$\omega_{a,j}$	optimizer state within $s_{a,j}$	typed state
$ma_{a,j}$	replay or external learning-memory state within $s_{a,j}$	typed state
$ca_{a,j}$	capacity-allocation state within $s_{a,j}$	parameter/byte vector
$Y_{a,\pi,\mathbf{z},r,q}$	observed endpoint q for one method, order, intervention cell, and seed	endpoint-specific
$\mu_{a,\pi,\mathbf{z},q}$	seed expectation of $Y_{a,\pi,\mathbf{z},r,q}$ under the frozen generator	endpoint-specific
N_k^{data}	exogenous examples presented from task k	count
$N_{a,\pi,k\ell}^{\text{acc}}$	examples from task T_k accepted by executing module M_ℓ	count
$U_{a,\pi,k\ell}$	optimizer-update applications to M_ℓ attributed to examples from T_k under the frozen attribution rule	count
C_a^{tot}	total installed parameter-capacity ceiling	parameters
C_a^{act}	maximum simultaneously active parameter ceiling	parameters
B_a^{peak}	peak live attributed storage	bytes
t	elapsed wall time from the registered run start	seconds
E_a	calibrated externally measured run energy at the declared boundary	joules
P_a	mean measured power over a declared interval	watts

Every typed state is serialized or hashed at each task boundary. A variable that affects later updates but is absent from $s_{a,j}$ is an unmeasured shared state and invalidates the corresponding causal cut.

Fixed-task-multiset and module-eligibility identity

For the initial design, freeze a one-to-one admission map $g(T_k) = M_k$ before order assignment and define $A_k = (T_k, M_k)$. Each pair appears exactly once. A valid permutation satisfies

$$\pi : \{1, \dots, K\} \rightarrow \{1, \dots, K\}$$

as a bijection, and therefore

$$\biguplus_{j=1}^K \{T_{\pi_j}\} = \mathcal{T},$$

where $\biguplus$ denotes multiset union. Equivalently, for every identity k ,

$$\sum_{j=1}^K 1_{\{\pi_j = k\}} = 1.$$

Here $1_{\{\cdot\}}$ is a dimensionless indicator. If a later design repeats blocks, the right-hand side becomes a frozen multiplicity n_k that is identical for every order. Adding or deleting a presented task, eligible module identity, initialization image, example, label, augmentation, or response

opportunity violates the order estimand. Realized activation, consolidation, merge, or retirement does not violate it; those states remain measured outcomes.

Let $I_{k,r}$ be the stored initialization image for paired module M_k and seed r . It is drawn before π is assigned, so

$$I_{k,r}(\pi) = I_{k,r}(\pi')$$

for all compared orders π and π' . Every other random stream is keyed by its typed identity and paired seed rather than by global call order. Otherwise a permutation could change initialization, dropout, augmentation, router, or fault draws and the contrast would not isolate order.

Let $\xi_{k,r}$ denote the frozen stream bundle keyed by admission unit A_k , paired seed r , stream type, and within-identity draw index. A sequence position selects that identity's bundle; it does not create a new position-keyed bundle.

The method-specific state transition is written abstractly as

$$s_{a,j} = F_a(s_{a,j-1}, A_{\pi j}, \mathbf{z}, \xi_{\pi j, r}),$$

where F_a is the frozen transition implementation and $\xi_{\pi j, r}$ is the paired random-input bundle for the identity occupying position j . Position-indexed exogenous disturbances are prohibited unless they are a separately registered, randomized factor. This notation permits parameters, routers, optimizer moments, replay, and structure to carry history; it does not require all methods to contain every component.

Exposure, capacity, optimizer, evaluator, and budget identities

Exogenous exposure

For each task k , let $\mathcal{D}_k$ be the frozen checksummed sequence of examples, labels, augmentations, and interaction outcomes. Exogenous parity requires

$$N_{a,\pi,k}^{\text{data}} = N_k^{\text{data}}$$

for every method a and order π in a matched comparison. The equality is in counts and content: equal counts with different examples are not equal exposure.

Routed acceptance may be an endogenous mechanism. The accepted-exposure ledger therefore retains both task identity k and executing-module identity ℓ . Define the task-total accepted count and its dimensionless exposure ratio as

$$N_{a,\pi,k}^{\text{acc}} = \sum_{\ell=1}^K N_{a,\pi,k\ell}^{\text{acc}}, \quad x_{a,\pi,k} = \frac{N_{a,\pi,k}^{\text{acc}}}{N_k^{\text{data}}}, \quad x_{a,\pi,k} \geq 0.$$

The ratio can exceed one only when the registered method duplicates or fans out one presented example to more than one module; every duplicate remains charged. Module-total accepted exposure is

$$N_{a,\pi,\ell}^{\text{acc}} = \sum_k N_{a,\pi,k\ell}^{\text{acc}}.$$

The exposure-equalizer intervention freezes the complete task-by-executing-module targets $N_{k\ell}^*$ and $U_{k\ell}^*$ such that

$$N_{a,\pi,k\ell}^{\text{acc}} = N_{k\ell}^*, \quad U_{a,\pi,k\ell} = U_{k\ell}^*$$

for every task k , executing module ℓ , and evaluated order in that intervention level. Thus equal task totals cannot hide a different routing allocation. Mixed-task update attribution and any fan-out rule are frozen before order assignment. Rejected, downweighted, or duplicate work remains charged even when it is not applied as an update.

Let $b_{a,\pi,k}$ be module k 's birth time in seconds from run start and let $T_{a,\pi}^{\text{wall}}$ be the registered run duration in seconds. Its final wall-clock age is

$$A_{a,\pi,k}^{\text{final}} = T_{a,\pi}^{\text{wall}} - b_{a,\pi,k} \quad [\text{s}].$$

The pre-instantiation cut sets all $b_{a,\pi,k} = 0$ while freezing inactive state. The exposure equalizer changes the $N_{k\ell}^{\text{acc}}$ and $U_{k\ell}$ matrices, not b . This is why age and accepted exposure can be crossed independently.

Capacity

Let $C_{a,\pi,k}(t)$ be parameters allocated to module k at time t , and let $A_{a,\pi,k}(t) \in \{0, 1\}$ indicate that the module is active. Valid runs satisfy

$$\sum_{k=1}^K C_{a,\pi,k}(t) \leq C_a^{\text{tot}}$$

and

$$\sum_{k=1}^K A_{a,\pi,k}(t) C_{a,\pi,k}(t) \leq C_a^{\text{act}}$$

for all measured t . Every term is a parameter count. Storage for optimizer state, router state, replay, checkpoints, and metadata is counted separately in bytes; parameter count is not treated as bytes without the registered representation width.

For the position-blind reservation cut, frozen quotas C_k^* obey

$$\sum_{k=1}^K C_k^* = C_a^{\text{tot}}, \quad C_{a,\pi,k}(t) \leq C_k^*.$$

Unused reserved capacity cannot be borrowed. Otherwise the later order could change effective total opportunity while nominal capacity remained fixed.

Optimizer and evaluator

Let $U_{a,\pi}^{\text{tot}}$ be all optimizer updates, including replay, router, consolidation, recovery, and failed-attempt updates. Let $V_{a,\pi}^{\text{eval}}$ be evaluator calls. Within a method-specific order contrast,

$$U_{a,\pi}^{\text{tot}} = U_{a,\pi'}^{\text{tot}}, \quad V_{a,\pi}^{\text{eval}} = V_{a,\pi'}^{\text{eval}}$$

for every compared π and π' , unless a count is itself a declared endpoint under a common ceiling. If early stopping is allowed, unused budget is reported; it is not silently transferred to tuning.

Across different methods, the optimizer mechanism may differ. Equality then means equal allowed update, search, evaluator, precision, and stopping budgets, not pretending that EWC, OGD, replay, PBT, and ordinary SGD execute identical operations. Method-specific work remains visible in the complete ledger.

Complete budget vector

Let $N^{\text{data}} = \sum_k N_k^{\text{data}}$ be total presented exposure. Total accepted task-by-module exposure is

$$N^{\text{acc}} = \sum_k \sum_{\ell} N_{a,\pi,k\ell}^{\text{acc}}.$$

Let N^{fwd} and N^{bwd} be forward- and backward-evaluation counts, and let B^{read} , B^{write} , and B^{peak} be attributed read, written, and peak-live byte counts. Let T^{worker} be summed provisioned worker time in seconds, T^{wall} be elapsed run time in seconds, and E be calibrated external energy in joules. U^{tot} and V^{eval} retain the definitions above.

Define

$$\mathbf{B}_{a,\pi} = [N^{\text{data}}, N^{\text{acc}}, U^{\text{tot}}, N^{\text{fwd}}, N^{\text{bwd}}, V^{\text{eval}}, B^{\text{read}}, B^{\text{write}}, B^{\text{peak}}, T^{\text{worker}}, T^{\text{wall}}, E]_{a,\pi}.$$

The first six entries are counts; the next three are bytes; T^{worker} and T^{wall} are seconds; E is joules. This vector is never summed directly. Paired order arms require componentwise equality within frozen tolerances or are compared on a preregistered Pareto frontier.

Mean measured power is derived only when a calibrated energy interval of duration $\Delta t > 0$ seconds exists:

$$P = \frac{E}{\Delta t}.$$

Because $\text{J/s} = \text{W}$, P is in watts. Operations, parameters, or bytes cannot replace E in this equation.

Potential outcomes and order estimands

For endpoint q , method a , intervention vector $\mathbf{z}$, and order π , define the seed expectation

$$\mu_{a,\pi,\mathbf{z},q} = \mathbb{E}_r [Y_{a,\pi,\mathbf{z},q}],$$

where the expectation is over the frozen seed generator, not over all possible tasks or machines.

Pairwise order effect

For two preregistered orders π and π' , the controlled order effect is

$$\tau_{a,q}(\pi, \pi'; \mathbf{z}) = \mu_{a,\pi,\mathbf{z},q} - \mu_{a,\pi',\mathbf{z},q}.$$

Its unit is the endpoint unit. A positive value is beneficial only when higher values of endpoint q are defined as better. Costs and errors retain their natural lower-is-better direction rather than having signs silently reversed.

The paired estimator over R seeds is

$$\hat{r}_{a,q}(\pi, \pi'; \mathbf{z}) = \frac{1}{R} \sum_{r=1}^R (Y_{a,\pi,\mathbf{z},r,q} - Y_{a,\pi',\mathbf{z},r,q}),$$

where R is a dimensionless seed count.

Order-distribution sensitivity

Let $|\Pi_K| = K!$. The permutation-average endpoint is

$$\bar{\mu}_{a,\mathbf{z},q} = \frac{1}{K!} \sum_{\pi \in \Pi_K} \mu_{a,\pi,\mathbf{z},q}.$$

The between-order variance is

$$V_{a,\mathbf{z},q}^{\text{order}} = \frac{1}{K!} \sum_{\pi \in \Pi_K} (\mu_{a,\pi,\mathbf{z},q} - \bar{\mu}_{a,\mathbf{z},q})^2.$$

This is the finite-population variance for an order drawn uniformly from all $K!$ permutations. If only $m < K!$ orders are sampled uniformly without replacement, the preregistered sample estimator uses denominator $m - 1$ and reports its sampling uncertainty; it is not silently substituted for the complete enumeration. If q is a dimensionless score, the variance is squared score units. The order range is

$$R_{a,\mathbf{z},q}^{\text{order}} = \max_{\pi \in \Pi_K} \mu_{a,\pi,\mathbf{z},q} - \min_{\pi \in \Pi_K} \mu_{a,\pi,\mathbf{z},q},$$

in the endpoint unit. The range is descriptive and selection-biased as an estimate of a future best order; it cannot replace simultaneous pairwise intervals.

Position effect

For task identity k and position j , define

$$\eta_{a,k,j,q}(\mathbf{z}) = \frac{1}{(K-1)!} \sum_{\pi \in \Pi_K: \pi_j=k} \mu_{a,\pi,\mathbf{z},q}.$$

The position contrast

$$\eta_{a,k,j,q}(\mathbf{z}) - \eta_{a,k,j',q}(\mathbf{z})$$

compares the same task at two positions averaged over all orders of the other tasks. It is not an “early-arrival law” unless it replicates across protected task families and survives the mechanism cuts.

For a plot or table that compares every position with the task’s mean over positions, define

$$\bar{\eta}_{a,k,\cdot,q}(\mathbf{z}) = \frac{1}{K} \sum_{j=1}^K \eta_{a,k,j,q}(\mathbf{z}), \quad c_{a,k,j,q}(\mathbf{z}) = \eta_{a,k,j,q}(\mathbf{z}) - \bar{\eta}_{a,k,\cdot,q}(\mathbf{z}).$$

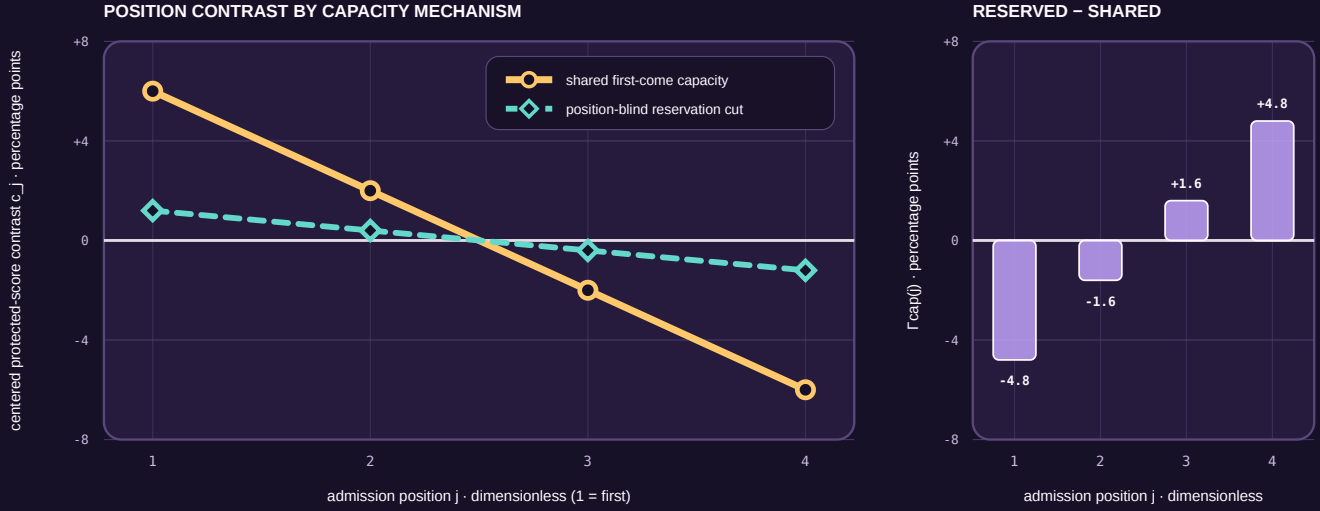
Thus $\sum_j c_{a,k,j,q}(\mathbf{z}) = 0$. If the only changed factor is the capacity-reservation cut, its position-specific interaction is

$$\Gamma_{a,k,j,q}^{(\text{cap},\text{pos})}(\mathbf{z}-\text{cap}) = c_{a,k,j,q}(\text{zcap} = 1, \mathbf{z}-\text{cap}) - c_{a,k,j,q}(\text{zcap} = 0, \mathbf{z}-\text{cap}).$$

Hypothetical capacity cut and position contrast

$$c_j = \eta_{\cdot}(a, k, j, q) - \text{mean over positions } j' \text{ of } \eta_{\cdot}(a, k, j', q); \quad \Gamma_{\text{cap}}(j) = c_j^{\text{reserved}} - c_j^{\text{shared}}$$

CONSTRUCTED VALUES · NO MODEL TRAINED · NO MEASUREMENTS



HYPOTHETICAL FACTORIAL-DESIGN ILLUSTRATION · NO MEASUREMENTS

The constructed reservation cut attenuates the early-position advantage and late-position penalty.

A real run fixes presented tasks, eligible module IDs, exogenous exposure, capacity, optimizer, evaluator and budget.

hypothetical percentage-point contrasts · not evidence or an effect-size target

FIGURE 59 — Hypothetical centered position contrasts for shared and reserved capacity, with the resulting capacity-cut interaction

The figure is an algebraic reading aid. It substitutes constructed percentage-point contrasts for $c_{a,k,j,q}$ and plots their constructed difference as $\Gamma_{a,k,j,q}^{(\text{cap}, \text{pos})}$. The values are not measurements, estimates, predictions, recommended effect sizes, or evidence that a capacity mechanism exists. Its editable specification is the [history-conditioned-position-contrast](#) entry.

Optimized-order advantage

Let g be a frozen order-selection algorithm, let $\hat{\pi}_g$ be its selected order using discovery information only, and let Π^{rand} be the uniform distribution over admissible orders. The held-out optimized advantage is

$$\Delta_{a,g,q}^{\text{opt}} = \mu_{a,\hat{\pi}_g,\mathbf{z},q} - \mathbb{E}_{\pi \sim \Pi^{\text{rand}}} [\mu_{a,\pi,\mathbf{z},q}].$$

The quality contrast is incomplete until the order optimizer's pilot examples, similarity or curvature measurements, candidate sequences, evaluator calls, failed runs, bytes, seconds, and joules are appended to $\mathbf{B}_{a,\hat{\pi}_g}$. Selecting the observed maximum from confirmation orders is not g and is not a valid optimized-order estimate.

Factorial mechanism estimands

Define the binary intervention vector

$$\mathbf{z} = [z_{\text{age}}, z_{\text{exp}}, z_{\text{cap}}, z_{\text{state}}, z_{\text{fac}}, z_{\text{lock}}] \in \{0, 1\}^6,$$

where 0 is the natural-history level and 1 is the causal-cut level defined in the audit. Let $\mathbf{z}-h$ denote all factors except mechanism h .

For one order π , the controlled main contrast of mechanism h at fixed $\mathbf{z}-h$ is

$$\Delta_{a,\pi,q}^{(h)}(\mathbf{z}-h) = \mu_{a,\pi,(z_h=1,\mathbf{z}-h),q} - \mu_{a,\pi,(z_h=0,\mathbf{z}-h),q}.$$

The order-by-mechanism interaction for two orders is

$$\Gamma_{a,q}^{(h)}(\pi, \pi'; \mathbf{z}-h) = \tau_{a,q}(\pi, \pi'; z_h = 1, \mathbf{z}-h) - \tau_{a,q}(\pi, \pi'; z_h = 0, \mathbf{z}-h).$$

$\Gamma^{(h)}$ has the endpoint unit. If cutting a path reduces an order contrast toward zero, that is evidence that the measured effect depends on the cut under the frozen design. It is not proof that mechanism h is the only mediator.

For distinct mechanisms h and h' , the two-factor interaction at one order is the inclusion--exclusion contrast

$$\Delta_{a,\pi,q}^{(h,h')} = \mu_{11} - \mu_{10} - \mu_{01} + \mu_{00},$$

where the two subscripts give $(z_h, z_{h'})$ and all remaining factors are fixed. Higher-order interactions use the same inclusion--exclusion rule.

Decomposition limits

The factorial is randomized, but a unique additive causal decomposition is not generally available because:

1. capacity changes which exposures are accepted;
2. exposure changes optimizer and shared-state trajectories;
3. shared state changes routing and therefore capacity demand;
4. facilitation changes both information and subsequent work;
5. lock-in changes the set of later admissible states; and
6. nonlinearity allows higher-order interactions.

Consequently,

$$\tau_{a,q}(\pi, \pi'; \mathbf{0}) \neq \sum_h \Gamma_{a,q}^{(h)}(\pi, \pi'; \mathbf{z}-h)$$

in general. The right side also depends on the levels at which the other factors are fixed. "Percent mediated" is not reported unless a separate causal model supplies and defends the required cross-world assumptions.

Ordinary-scheduling negative control

Let J_k denote a task job that consumes the same declared compute and I/O but does not update parameters, routers, optimizer state, replay, normalization, or structure. Let σ_π be its completion order under scheduler input π . Scheduling may change a service endpoint $L(\sigma_\pi)$ such as latency in seconds.

Separately, let u_k be the complete checksummed learned update record produced for task k . Apply all records after job completion in one canonical order κ to obtain

$$s_{a,K}^{\text{canon}} = u_{\kappa_K} \circ \dots \circ u_{\kappa_1}(s_{a,0}).$$

If $s_{a,K}^{\text{canon}}$ and all final capability endpoints are identical across input permutations while only $L(\sigma \pi)$ differs, the detected effect is ordinary scheduling under this control. If update records themselves depend on the live history, canonical replay is diagnostic rather than an oracle; that dependence must be attributed to one or more serialized state paths.

Learning endpoints

Assume higher task score is better. Let $Q_{a,\pi,k}^{\text{post}}$ be task k 's held-out score immediately after its acquisition block, and let $Q_{a,\pi,k}^{\text{final}}$ be its score after all K blocks at the frozen retention horizon. Both retain the declared task score unit.

Backward transfer is

$$\text{BWT}_{a,\pi,k} = Q_{a,\pi,k}^{\text{final}} - Q_{a,\pi,k}^{\text{post}}.$$

Positive values indicate improvement and negative values indicate forgetting. The nonnegative forgetting magnitude is

$$F_{a,\pi,k} = \max(0, Q_{a,\pi,k}^{\text{post}} - Q_{a,\pi,k}^{\text{final}}).$$

Let Q_k^{floor} be the preregistered protected score floor for task k . The protected shortfall is

$$H_{a,\pi}^{\text{prot}} = \max_{1 \leq k \leq K} \max(0, Q_k^{\text{floor}} - Q_{a,\pi,k}^{\text{final}}).$$

This has the score unit and cannot be cancelled by high performance on another task. The worst-task score is

$$Q_{a,\pi}^{\text{min}} = \min_{1 \leq k \leq K} Q_{a,\pi,k}^{\text{final}}.$$

For newcomer task k , let Q_k^* be a frozen competence threshold and let $n_{a,\pi,k}^*$ be the first accepted-example count at which the threshold is met and remains met for the frozen confirmation window. If the threshold is never met, $n_{a,\pi,k}^*$ is right-censored at the task budget rather than deleted. Forward facilitation is evaluated through this sample-complexity endpoint and a matched no-history baseline, not through final score alone.

Structural and routing endpoints

Let $p_{a,\pi,\ell}(t)$ be the fraction of routed load assigned to executing module ℓ at time t , with

$$p_{a,\pi,\ell}(t) \geq 0, \quad \sum_{\ell=1}^K p_{a,\pi,\ell}(t) = 1$$

when at least one module is eligible. Router entropy is

$$H_{a,\pi}^{\text{route}}(t) = - \sum_{\ell=1}^K p_{a,\pi,\ell}(t) \log p_{a,\pi,\ell}(t),$$

in nats when the natural logarithm is used. Terms with $p = 0$ contribute zero. Low entropy is not by itself specialization or quality.

For a frozen task-by-module evaluation matrix $G_{a,\pi,k\ell}$, where row k is task identity and column ℓ is module identity, define a dimensionless specialization contrast for module ℓ as

$$S_{a,\pi,\ell} = \max_k G_{a,\pi,k\ell} - \frac{1}{K-1} \sum_{k' \neq k_\ell^*} G_{a,\pi,k'\ell},$$

where k_ℓ^* is the task attaining the maximum under a frozen tie rule. This metric is meaningful only if G uses a common dimensionless score. If tasks have different native units, their raw matrix is reported and no such subtraction is allowed.

Admission time, commitment time, unlock count, retirement count, allocated capacity, accepted load, dropped load, and lineage remain separate endpoints. An expert label or high S does not prove independent function, causal necessity, or efficient routing.

Service and lifecycle endpoints

If N^{due} predictions are due and N^{accepted} are correct, current, integrity-valid, and within the frozen deadline, accepted service is

$$S^{\text{acc}} = \frac{N^{\text{accepted}}}{N^{\text{due}}},$$

a dimensionless fraction. Missing, late, stale, duplicated, and inaccurate predictions remain separate counts before aggregation.

Latency L_i for due item i is completion time minus original due time, in seconds. Every due item remains in the denominator; missing completions receive the preregistered right-censor value. p50, p95, and p99 are reported with the exact quantile convention.

The mandatory result is a vector,

$$\mathbf{V}_{a,\pi} = [\{Q_{a,\pi,k}^{\text{final}}\}_{k=1}^K, Q_{a,\pi}^{\text{min}}, H_{b,\pi}^{\text{prot}}, \{\text{BWT}_{a,\pi,k}\}_{k=1}^K, \{n_{a,\pi,k}^*\}_{k=1}^K, S^{\text{acc}}, L_{0.99}, \mathbf{B}_{a,\pi}].$$

The entries have different units and are not averaged. Method a dominates method a' only under preregistered direction and relevance margins, with no protected endpoint worse and at least one endpoint materially better.

Multiplicity and uncertainty

Let $\mathcal{H}_{\text{primary}}$ be the frozen family of primary order, method, and order-by-mechanism contrasts. Its membership is committed before confirmation outcomes are opened.

For randomization inference, compute one test statistic for each $h \in \mathcal{H}_{\text{primary}}$ and, under the registered treatment re-randomizations, use the maximum absolute standardized statistic

$$M^{\text{perm}} = \max_{h \in \mathcal{H}_{\text{primary}}} |T_h^{\text{perm}}|.$$

The empirical distribution of M^{perm} supplies family-wise adjusted decisions and simultaneous intervals. If that procedure is computationally unavailable, Holm's step-down correction is the fallback. Unadjusted effect estimates and intervals remain visible, but they cannot carry the confirmatory decision.

Seeds are paired experimental units only for the generator they instantiate. Tasks, orders, endpoints, or repeated checkpoints from one run are not treated as independent sample-size multipliers. Generalization to a task population or machine population requires corresponding sampled levels and a second-family or second-machine replication.

The minimum relevant effect $\delta_q^{\min}$ is declared in endpoint q 's native unit. "Statistically nonzero" without crossing $\delta_q^{\min}$ does not keep the frontier alive.

Dimensional analysis checklist

1. Task scores may be dimensionless proportions or native score units; the unit is declared before subtraction.
2. Counts of examples, updates, parameters, modules, calls, and operations are dimensionless counts with different meanings and are not interchangeable.
3. Storage and traffic are bytes; parameter counts become bytes only after representation width and metadata are included.
4. Latency, worker time, and wall time are seconds but have different boundaries and remain separately named.
5. Energy is joules and power is joules per second, or watts.
6. Throughput is accepted items per second and is not an energy efficiency.
7. Quality per joule is reported only beside its raw quality and energy components at matched task, quality floor, horizon, and boundary.
8. Ecological abundance and artificial router load are not assigned a common unit merely to make an analogy.

Testable predictions

These are preregistrable hypotheses, not findings:

1. In high-overlap task strata, the capacity reservation cut yields $\Gamma^{(\text{cap})}$ opposite in sign to the natural order contrast if incumbency is carried by capacity pre-emption.
2. If shared-state modification is a carrier, the magnitude of $\tau(\pi, \pi'; \mathbf{z})$ decreases when $z_{\text{state}} = 1$ while private post-acquisition scores remain within their non-inferiority margins.
3. If typed facilitation improves newcomer acquisition, setting $z_{\text{fac}} = 1$ increases the newcomer threshold count n^* or right-censoring rate without a corresponding change under the sham-only scheduler control.
4. If lock-in is useful rather than merely persistent, the irreversible arm improves protected delayed outcomes after the inducing module is removed and remains non-dominated after checkpoint, unlock, migration, and recovery costs enter **B**.
5. If the effect is ordinary continual-learning interference, replay, EWC, or OGD matches the complete frontier without module-birth factors.
6. If the effect is curriculum selection, a held-out optimized order yields the same frontier after its search budget is charged.
7. V^{order} and the sign of position contrasts vary with feature overlap, readout conflict, horizon, and method; no universal early/late rule is predicted.

Numerical and contract checks

1. Every stored permutation is a bijection and has the same task-multiset and eligible-module-set digests.

2. Every task identity has identical exogenous example and update opportunity digests across orders.
3. Capacity inequalities hold at every event, not only at final state.
4. Every shared-state reset restores the same checksummed boundary image while retaining only the explicitly exempt private state.
5. Sham facilitation payloads match byte count, availability time, parser work, and transport work, and are provably answer-free.
6. Canonical replay applies exactly the same update-record multiset in the same canonical order for every scheduling control.
7. Invalid, missing, censored, timed-out, and resource-exceeded runs remain in result tables and multiplicity families.
8. The full budget vector is reconstructed from raw events before any method comparison.
9. Half-precision, quantized, sparse, or parallel execution receives its own parameter, byte, work, and energy accounting; nominal operation equality is insufficient.
10. No equation in this note authorizes a scientific, performance, or energy result before a registered executable package and sealed execution exist.

Kill boundary

This mathematical contract supplies no residual architecture contribution if:

1. for every protected primary contrast on both task families, the registered simultaneous interval lies wholly inside $(-\delta_q^{\min}, +\delta_q^{\min})$; the unobserved population condition $|\tau| < \delta_q^{\min}$ is the corresponding no-effect region, not an executable decision rule;
2. canonical replay or the non-learning scheduler reproduces the effect;
3. equal age plus equal accepted exposure removes it;
4. no randomized causal cut changes it on fresh seeds;
5. a complete conventional null is non-inferior at lower or equal lifecycle cost;
6. results depend on selecting the best confirmation permutation;
7. protected-task harm or resource ceilings are violated; or
8. the second task family or machine replication fails.

Passing these equations and checks would justify evaluating evidence. It would not by itself establish a biological equivalence, universal mechanism, or novel architecture.

Interface-qualified retroactivity and insulation

- **Purpose:** distinguish direct output sequestration, substrate competition, generic shared-resource competition, and intentional useful coupling; define measurable effects; and bound claims about insulating module interfaces
- **Evidence audit:** interface-qualified retroactivity and insulation
- **Experiment contract:** Fixture F-027
- **Result state:** analytical definitions and synthetic experiment specifications only; every empirical, workstation, and energy result is **NO_RESULT**

Four coupling classes

Let an upstream module produce a signal for one or more downstream clients. Four effects that can look similar in a latency or output trace must remain distinct.

1. **Direct output sequestration or connection back-action** occurs when the receiver binds, pins, drains, blocks, or otherwise changes the output carrier or producer-owned state that participates in upstream dynamics. Removing the output connection removes this path even when total work is held fixed.
2. **Substrate competition** occurs when multiple downstream transformations compete for a declared pathway-specific enzyme, catalyst, transformer, or service pool. It can change the service delivered to other substrates and, when substrate binding changes enzyme modification or availability, can also feed back into pathway state. It is not generic compute contention.
3. **Generic shared-resource competition** occurs when otherwise unrelated modules contend for a processor, memory channel, accelerator, allocator, network, transcription or translation capacity, thermal limit, power cap, or another common resource. A disconnected work-matched load can reproduce this path.
4. **Intentional useful coupling** is an admitted interaction whose effect is part of the declared objective: for example signal integration by substrate competition or a designed gradient, acknowledgement, feedback, or feedforward controller. Intended coupling is not called a defect merely because it alters another component, but its authority and costs remain in the ledger.

For upstream state x , intended input u , direct output-binding state s , pathway-specific substrate-service state v , generic shared-resource state q , and admitted coupling signal c , a mechanism ledger can write

$$\dot{x} = f(x, u) + b_{\text{seq}}(x, s) + b_{\text{sub}}(x, v) + b_{\text{shared}}(x, q) + b_{\text{useful}}(x, c) + b_{\text{cross}}(x, s, v, q, c),$$

where every term has the state unit of x per second. The terms are isolated dynamics, direct sequestration, pathway-specific substrate competition, generic shared-resource coupling, declared useful coupling, and their interactions. A term may be identically zero in a particular

system. The names do not identify a mechanism; selective interventions do. In particular, functional value is an objective label, not a fifth physical pathway: a substrate-competition path can be intentionally useful in one task and harmful in another.

Biological source model and units

The source-shaped model uses one free signalling species and one downstream binding pool. Its notation is:

SYMBOL	MEANING	UNIT
t	elapsed time	second (s)
$X(t)$	free upstream signalling concentration	mole per cubic metre (mol m^{-3})
$C(t)$	concentration bound to downstream sites	mol m^{-3}
p_{tot}	total downstream-site concentration	mol m^{-3}
$k(t)$	production flux of free signal	$\text{mol m}^{-3} \text{s}^{-1}$
δ	first-order loss rate of free signal	s^{-1}
k_{on}	association-rate constant	$\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$
k_{off}	dissociation-rate constant	s^{-1}
K_d	dissociation concentration, $k_{\text{off}}/k_{\text{on}}$	mol m^{-3}
$Y(t)$	total signal concentration, $X(t) + C(t)$	mol m^{-3}
ε	timescale ratio, δ/k_{off}	dimensionless
$g(Y)$	quasi-steady bound concentration	mol m^{-3}
$R(X)$	reduced retroactivity factor	dimensionless

The isolated upstream module is

$$\dot{X}_{\text{iso}} = k(t) - \delta X_{\text{iso}}.$$

After a downstream binding pool is connected, the mass-action model is

$$\dot{X} = k(t) - \delta X + k_{\text{off}}C - k_{\text{on}}X(p_{\text{tot}} - C),$$

$$\dot{C} = k_{\text{on}}X(p_{\text{tot}} - C) - k_{\text{off}}C.$$

The binding fluxes appear with opposite signs and therefore conserve $X + C$ in the absence of production and loss. Connection changes the free signal trajectory without requiring an unrelated shared resource.

When binding and unbinding are fast relative to production and loss, $\varepsilon \ll 1$, the quasi-steady bound concentration satisfies

$$C = g(Y), \quad 0 = -k_{\text{off}}C + k_{\text{on}}(Y - C)(p_{\text{tot}} - C).$$

The reduced free-signal dynamics are

$$\dot{\bar{X}} = (k(t) - \delta \bar{X}) [1 - R(\bar{X})],$$

with

$$R(\bar{X}) = \left[1 + \frac{(1 + \bar{X}/K_d)^2}{p_{\text{tot}}/K_d} \right]^{-1}.$$

Here $\bar{X}$ is the reduced approximation to X , in mol m^{-3} . The formula predicts stronger dynamic back-action when the load p_{tot} is large relative to the signal and when binding affinity is high, meaning K_d is small. It is not licensed when the timescale separation or mass-action model fails.

Worked mass-action reference

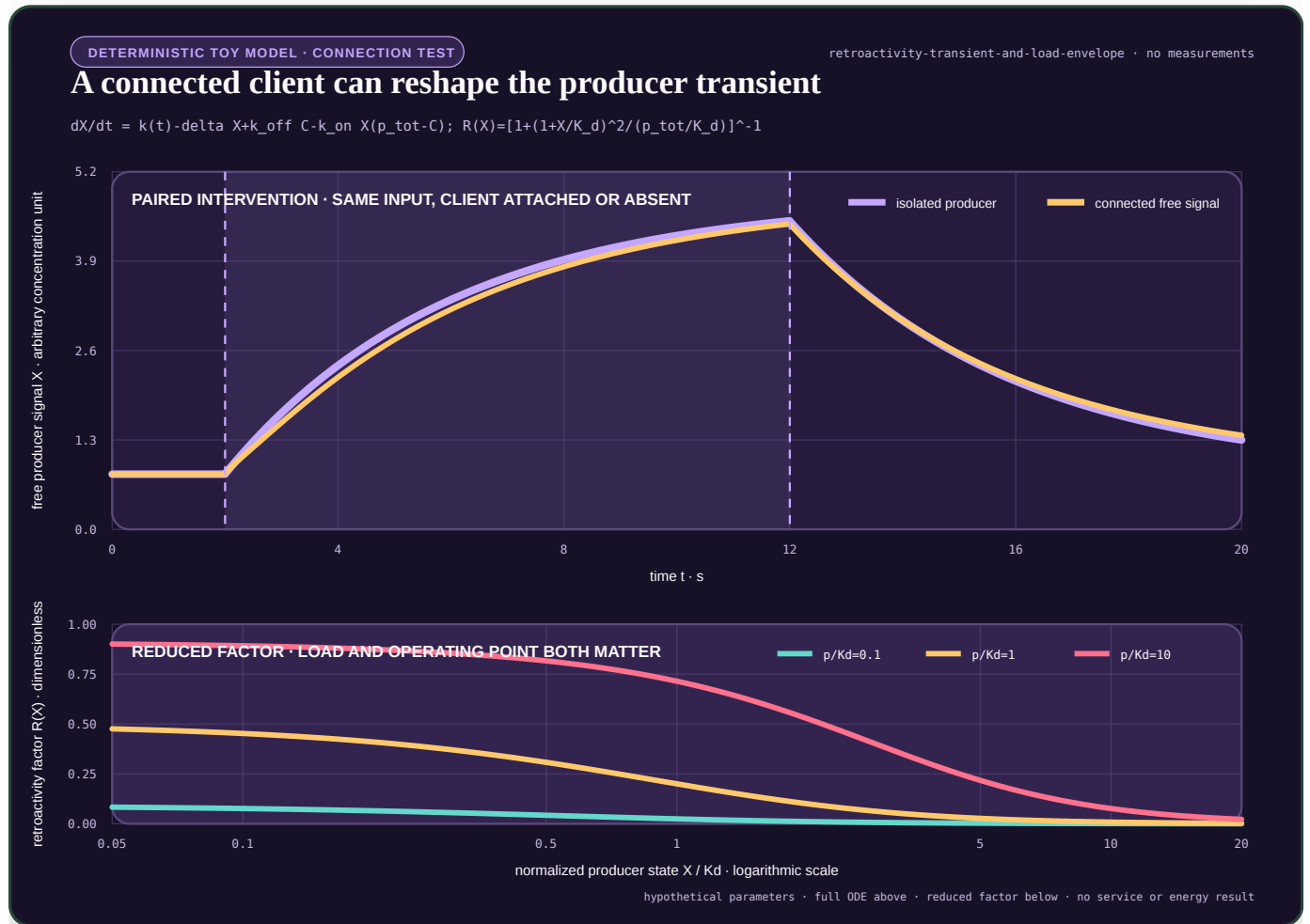


FIGURE 60 — Reference source-model response and retroactivity factor

The figure is an explanatory rendering of the registered source equations, not an experiment result. Its editable definition is the core-model plot specification under the identifier **interface-qualified-retroactivity**; the generated SVG is kept beside the public site assets. Any parameter or caption change must begin in that editable specification and must retain the **NO_RESULT** boundary.

For client j , let p_j , C_j , $k_{\text{on},j}$, $k_{\text{off},j}$, and $K_{d,j}$ have the corresponding units above. The full multiple-client model is

$$\dot{X} = k(t) - \delta X + \sum_{j=1}^N [k_{\text{off},j} C_j - k_{\text{on},j} X(p_j - C_j)],$$

$$\dot{C}_j = k_{\text{on},j} X(p_j - C_j) - k_{\text{off},j} C_j,$$

where N is the dimensionless number of attached clients. Client effects need not be independent after they couple through the same free signal.

For heterogeneous fast-binding pools, define

$$\bar{C}_j(\bar{X}) = \frac{p_j \bar{X}}{K_{d,j} + \bar{X}}, \quad A_N(\bar{X}) = \sum_{j=1}^N \frac{p_j K_{d,j}}{(K_{d,j} + \bar{X})^2}.$$

Because $Y = \bar{X} + \sum_j \bar{C}_j$ and $dY/d\bar{X} = 1 + A_N(\bar{X})$, the reduced free-signal model is

$$R_N(\bar{X}) = \frac{A_N(\bar{X})}{1 + A_N(\bar{X})}, \quad \dot{\bar{X}} = (k(t) - \delta \bar{X})[1 - R_N(\bar{X})].$$

The empty sum gives $R_0 = 0$; for one client the expression reduces exactly to the single-pool factor above. It is licensed only when every registered fast-binding condition and the full-versus-reduced error gate pass. Replacing heterogeneous $K_{d,j}$ values by an averaged affinity is not this reduction.

Artificial bounded-publisher model

The AI translation is an interface test, not a claim that digital activations are molecules. It uses a stateful producer whose published states occupy a finite producer-owned slot until every direct consumer releases it.

SYMBOL	MEANING	UNIT
n	logical input-step index	dimensionless integer
Δt	scheduled interval between input steps	s
d	producer-state dimension	dimensionless count
x_n	producer state at step n	normalized state unit (NSU)
u_n	registered producer input	normalized input unit (NIU)
A	state transition from NSU to NSU	dimensionless
B	input transition from NIU to NSU	NSU NIU ^{−1}
H	output map from NSU to normalized output unit	NOU NSU ^{−1}
y_n	producer output	normalized output unit (NOU)
S	producer-owned publication slots	dimensionless count
a_n	slots available at step n	dimensionless count
$\mathcal{P}_n$	distinct publication-slot IDs with at least one unreleased client reference	dimensionless finite set
$h_{j,n}$	registered hold time for client j and publication n	s
K	pathway-specific downstream transform servers	dimensionless count
$w_{j,n}$	transform-service demand for client j at step n	logical operation count
V	finite substrate-service intervention indicator	dimensionless, zero or one
Z	direct output-sequestration intervention indicator	dimensionless, zero or one
Q	generic shared-resource intervention indicator	dimensionless, zero or one
F	intentional-feedback intervention indicator	dimensionless, zero or one
c_{copy}	work for one snapshot copy	logical operation count
b_{copy}	bytes written for one snapshot	byte (B)

SYMBOL	MEANING	UNIT
L_n	deadline-lateness at step n	s
m_n	deadline-miss indicator	dimensionless, zero or one

The isolated logical update is

$$x_{n+1}^{\text{iso}} = Ax_n^{\text{iso}} + Bu_n, \quad y_n^{\text{iso}} = Hx_{n+1}^{\text{iso}}.$$

Here u_n is due at $t_n = n\Delta t$. Publication sequence n semantically exposes the post-update state x_{n+1} and derived scalar y_n , carries t_n as its due time, and records the update-completion time separately. The fixture's physical 96 B record serializes the state and metadata only; deriving y_n requires the charged state reduction. A missing record leaves a permanent due-sequence gap; sequence IDs are not compacted or reused. State x_0 is an initial condition, not publication sequence zero. The non-published initial scalar is $y_{\text{init}} = Hx_0$.

For the direct finite interface,

$$\mathcal{P}_n = \{s \in \{1, \dots, S\} : \exists j \text{ with an unreleased reference to slot } s\}, \quad a_n = S - |\mathcal{P}_n|.$$

Multiple clients may hold the same publication slot; that slot appears once in $\mathcal{P}_n$. Summing client references would double-count a shared slot and is prohibited.

If a publication requires one slot, the frozen blocking rule is

$$x_{n+1}^{\text{direct}} = \begin{cases} Ax_n^{\text{direct}} + Bu_n, & a_n \geq 1, \\ x_n^{\text{direct}}, & a_n < 1, \end{cases}$$

$$m_n = \mathbf{1}[a_n < 1], \quad y_n^{\text{direct}} = Hx_{n+1}^{\text{direct}}.$$

The indicator $\mathbf{1}[\cdot]$ equals one when its condition is true and zero otherwise. Holding the state is one registered policy. Drop-input, queue-input, and block-wall-clock policies are separate arms because they produce different trajectories.

An immutable-snapshot interface acquires the physical x_{n+1} record, queues its charged copy into consumer-owned storage, and releases the producer slot only after copy completion, cancellation, or TTL expiry. Consumer work and consumer-storage lifetime cannot extend that source pin. The logical state transition may therefore retain a bounded copy-time effect under slot pressure; c_{copy} , b_{copy} , reference work, copy latency, queueing, staleness, and consumer-storage lifetime are all charged to the snapshot arm. The model exposes the benefit and cost rather than assuming free insulation.

Unless the fixture declares an override, snapshot, actor, backpressure, substrate-service, private-service, and replica arms reuse one capture law: capacity admission precedes source acquisition; an accepted 24 B descriptor reserves the 96 B destination; the producer worker acquires the exact source; and a charged per-client capture worker copies it before handoff to downstream service. Rejection creates no source reference, while cancellation releases the descriptor, reservation, any partial destination, and the source reference. These ownership intervals enter both B_{peak} and worker accounting.

Pathway-specific substrate-service control

The artificial substrate-competition control starts from identical immutable snapshot acquisition in its shared and private arms. Snapshot copying may pin the producer for its bounded copy interval, but transform service never owns or extends that reference; the V contrast must therefore have zero producer- state effect even if both arms share the same copy-time effect. Every publication creates one transform request per subscribed client. The primary causal comparison fixes $K = N$ and uses N identical 4096-operation-per-second servers in both arms. With $V = 1$, requests

enter one deterministic FCFS queue feeding those N servers. With $V = 0$, each client has one permanently assigned deterministic FCFS queue and one of the same servers. Both arms therefore have equal server count, per-server rate, nameplate capacity, and transform demand; only request pooling changes. Cells with $K \neq N$ are capacity diagnostics and cannot identify V . Pooling can improve or worsen a client stratum, so the effect is two-sided.

This control represents competition for a declared domain-specific transform service, not a biochemical claim. Under exclusive producer placement its construction requires

$$x_{n+1} = Ax_n + Bu_n$$

for both values of V . Thus finite transform service may change client outputs or latency, but it must not change logical producer state unless a separate path is enabled. A nonzero upstream distortion in that exclusive, immutable control falsifies the implementation boundary.

Causal identification

Let $Z \in \{0, 1\}$ indicate direct output sequestration, let $V \in \{0, 1\}$ indicate finite pathway-specific substrate service, let $Q \in \{0, 1\}$ indicate disconnected work-matched load on a generic shared resource, and let $F \in \{0, 1\}$ indicate the separately logged intentional feedback channel. For a dimensionless endpoint D , write $D_{z,v,q,f}$ for the paired aggregate under the four binary settings. The baseline-referenced one-factor contrasts are

$$\Delta_{\text{seq}} = D_{1,0,0,0} - D_{0,0,0,0},$$

$$\Delta_{\text{sub}} = D_{0,1,0,0} - D_{0,0,0,0},$$

$$\Delta_{\text{shared}} = D_{0,0,1,0} - D_{0,0,0,0},$$

and

$$\Delta_{\text{useful}} = D_{0,0,0,1} - D_{0,0,0,0}.$$

For example, the sequestration-by-shared-resource interaction is

$$\Delta_{\text{seq,shared}} = D_{1,0,1,0} - D_{1,0,0,0} - D_{0,0,1,0} + D_{0,0,0,0}.$$

All contrasts above are dimensionless because D is dimensionless. Other pair and higher-order interactions use the same inclusion--exclusion rule and must be reported rather than absorbed into a main effect. Direct output back-action is identified only if Δ_{seq} survives exclusive resource allocation, collapses when the pinning path is cut, and cannot be reproduced by $V = 1$ or $Q = 1$. Substrate competition is identified through client-service changes under immutable producer reads that collapse when private transform servers replace the finite pool. Generic contention must be reproducible by disconnected work and must respond to resource placement. Intentional feedback is identified by replaying its logged messages with reads disabled and by disabling it while reads remain.

The disconnected load must match observed logical operations, bytes read, bytes written, allocation count, service-time distribution, and scheduling class as closely as the registered platform permits. Matching only nominal consumer count is insufficient. Equal total work does not identify substrate competition: requests must additionally be reassigned from the shared pathway-specific service to private services without changing snapshots or their demand.

Observation, noise-memory, and bandwidth estimators

Let $y_{r,n}^{\text{prod}}$ be the output derived from the current producer state record, let $y_{r,n}^{\text{client}}$ be the latest completed client output, and let $y_{r,n,b}^{\text{live}}$ be buffered live version b , all in normalized output units for replicate r and sample n . If $B_{r,n}$ is the dimensionless live-version count, the total-live observation is

$$y_{r,n}^{\text{total}} = \frac{y_{r,n}^{\text{prod}} + \sum_{b=1}^{B_{r,n}} y_{r,n,b}^{\text{live}}}{1 + B_{r,n}}.$$

This arithmetic mean is an artificial observation map, not a conserved biochemical total. It remains in normalized output units. Artificial publication records serialize eight state components and metadata, not a second output scalar. Each y^{prod} or y^{live} therefore requires a 64 B state read and the fixture's charged eight-operation reduction. At each telemetry due time the distinct physical record set is frozen before asynchronous copying; duplicate client references do not duplicate a version, and logical expiry cannot turn a telemetry-pinned record into accepted service.

For observation map o , let R_s be the dimensionless replicate count, let M be the dimensionless post-warm-up sample count, and let Δt be the sample period in seconds. All replicates must share byte-identical input, hold, service, deadline, and fault histories; only the registered componentwise process-noise stream may differ. Let $\mu_n^{(o)}$ be the deterministic process-noise-disabled trajectory under those exact same events. The residual is

$$e_{r,n}^{(o)} = y_{r,n}^{(o)} - \mu_n^{(o)},$$

in normalized output units. Varying the input or any non-noise event across replicates invalidates the estimator rather than entering the residual. For the registered post-warm-up window, define the grand residual mean and centred residual

$$\bar{e}_o = \frac{1}{R_s M} \sum_{r=1}^{R_s} \sum_{n=0}^{M-1} e_{r,n}^{(o)}, \quad \tilde{e}_{r,n}^{(o)} = e_{r,n}^{(o)} - \bar{e}_o.$$

Both retain normalized output units. For dimensionless lag index k , estimate the autocovariance

$$\hat{\gamma}_o(k) = \frac{1}{R_s(M-k)} \sum_{r=1}^{R_s} \sum_{n=0}^{M-k-1} \tilde{e}_{r,n}^{(o)} \tilde{e}_{r,n+k}^{(o)},$$

in squared normalized output units, and $\hat{\rho}_o(k) = \hat{\gamma}_o(k)/\hat{\gamma}_o(0)$, which is dimensionless. Let K_o be the first nonnegative lag at which two consecutive autocorrelations are nonpositive; if no such pair occurs before $M/4$, the estimate is unavailable. The integrated correlation-time estimator is

$$\tau_{c,o} = \Delta t \left[1 + 2 \sum_{k=1}^{K_o} \hat{\rho}_o(k) \right],$$

in seconds. A nonpositive estimate, a nonstationary residual diagnostic, or missing observation events invalidates the estimate rather than triggering imputation.

The fixture's executable nonstationarity diagnostic divides the complete post-warm-up residual sequence into ten contiguous equal-count blocks. With block mean μ_b , population variance v_b , grand mean μ , and $v_{\min} = 10^{-12} \text{ NOU}^2$, it reports

$$\Delta\mu = \max_b |\mu b - \mu|, \quad Rv = \frac{\max_b \max(v_b, v_{\min})}{\min_b \max(v_b, v_{\min})}.$$

$\hat{c}_{o}$ is available only when $\Delta\mu \leq 0.01$ NOU, $Rv \leq 4$, every block is complete, and no observation is missing. These are fixture decisions, not universal stationarity criteria; the registered sensitivity thresholds are reported.

For an input sinusoid of angular frequency $\omega = 2\pi f$ in radians per second, where f is in hertz, fit the post-warm-up output

$$y(t) = a_0 + a_s \sin(\omega t) + a_c \cos(\omega t) + \epsilon(t),$$

where a_0 , a_s , a_c , and residual $\epsilon(t)$ have the output unit and t is in seconds. If the fitted input amplitude is $A_u > 0$ in normalized input units, output amplitude $A_y = (a_s^2 + a_c^2)^{1/2}$ has the output unit and gain $G(\omega) = A_y/A_u$ has output units per input unit. Phase is $\phi(\omega) = \text{atan2}(a_c, a_s)$ in radians. The DC probe uses two constant inputs around $u_0 = 0.5$ NIU, $u_- = 0.49$ NIU and $u_+ = 0.51$ NIU, with the explicit shared override $x_{0,i} = 0.5$ NSU for every component, plus the same event history and observation map. If $\bar{y}_-$ and $\bar{y}_+$ are their final-50-second means after a 300 s run, define

$$G(0) = \frac{\bar{y}_+ - \bar{y}_-}{0.02 \text{ NIU}} \quad [\text{NOU NIU}^{-1}].$$

This is the fixture's operational DC estimate; failure of either constant trajectory to converge makes it unavailable. The minus-three-decibel bandwidth is the smallest interpolated frequency

$$\omega_B = \inf \left\{ \omega > 0 : G(\omega) \leq \frac{G(0)}{\sqrt{2}} \right\},$$

in radians per second. Log-linear interpolation is allowed only between two adjacent excited frequencies bracketing the threshold; otherwise ω_B is unavailable.

Distortion, competition, and service endpoints

Let T be a registered evaluation duration in seconds, $y(t)$ an upstream output in a declared output unit, and $s_y > 0$ a frozen scale in the same unit. The upstream trajectory distortion is

$$DU = \sqrt{\frac{1}{T} \int_0^T \left\| \frac{y_{\text{connected}}(t) - y_{\text{isolated}}(t)}{s_y} \right\|_2^2 dt}.$$

DU is dimensionless. Its discrete, duration-weighted implementation is

$$DU^{\text{disc}} = \sqrt{\frac{\sum_{n=0}^{M-1} \Delta t_n \|y_n^{\text{connected}} - y_n^{\text{isolated}}\|_2^2}{s_y^2 \sum_{n=0}^{M-1} \Delta t_n}},$$

where M is the dimensionless sample count and Δt_n is the duration represented by sample n , in seconds.

For the artificial fixture, samples are always taken on the exogenous due grid $t_n = n\Delta t$. The value at t_n is the latest producer state whose service completed by t_n , held from its actual completion, or y_{init} if none has completed. Blocked and queued updates are not realigned by logical sequence. The same sample-and-hold trace defines DU , t_{20} , and t_{50} ; client latency remains completion time minus the original due time.

For an existing downstream client, let $z_{1,n}^{(N)}$ and $z_{1,n}^{(1)}$ be the integrity-valid outputs for exact sequence n completed by its fixed deadline in the N -client and paired one-client arms. Let $a_n^{(N)}$ and $a_n^{(1)}$ be their zero-or-one availability indicators, let $s_z > 0$ be the frozen output scale, and let $P_{\text{miss}} = 10$ be the fixture's dimensionless missing penalty. Define

$$e_{C,n} = \begin{cases} (z_{1,n}^{(N)} - z_{1,n}^{(1)})/s_z, & a_n^{(N)} a_n^{(1)} = 1, \\ P_{\text{miss}}, & a_n^{(N)} a_n^{(1)} = 0, \end{cases} \quad DC = \sqrt{\frac{1}{M} \sum_{n=0}^{M-1} e_{C,n}^2}.$$

The endpoint is dimensionless, retains every due sequence, and separates harm to an existing client from distortion of the producer. Complete-case values are diagnostic only; sensitivity uses $P_{\text{miss}} \in \{2, 10, 100\}$.

For a registered step whose output changes from y_0 to y_∞ , and for $p \in \{0.2, 0.5\}$, define

$$t_p = \inf \{t \geq 0 : |y(t) - y_0| \geq p |y_\infty - y_0|\}.$$

t_{20} and t_{50} are therefore the $p = 0.2$ and $p = 0.5$ cases and are in seconds. Upward and downward values are reported separately; their difference is not called sign-sensitive unless the input, initial state, endpoint, and observation map are registered.

If N_{due} outputs are due and N_{accepted} meet the frozen accuracy and deadline criteria, accepted service is

$$S_{\text{acc}} = \frac{N_{\text{accepted}}}{N_{\text{due}}},$$

a dimensionless fraction. Dropped, stale, duplicated, late, and inaccurate outputs remain separate counts before any accepted-service aggregation.

For latency, every due output contributes one value. An integrity-valid completion contributes completion time minus original due time; any dropped, stale, duplicated, integrity-failed, or unavailable output contributes the right-censor value $L_{\text{cens}} = 300.25$ s. The ordinary nearest-rank p99 is reported. The protected $L_{0.99}$ used below equals L_{cens} if any censored value exists, and otherwise equals that ordinary p99. A valid-completion-only quantile is diagnostic only.

Insulation and weak-coupling frontiers

An insulating interface is evaluated on a vector, not a scalar:

$$\mathbf{v} = (DU, DC, 1 - S_{\text{acc}}, L_{0.99}, O, B_{\text{peak}}, B_{\text{write}}, C_{\text{prov}}, W_{\text{peak}}, E).$$

Here $L_{0.99}$ is p99 latency in seconds, O is logical operation count, B_{peak} is peak retained memory in bytes, B_{write} is total bytes written, and C_{prov} is provisioned reference-worker time, and W_{peak} is peak concurrent worker count, a dimensionless count. For worker k at rate ν_k logical operations per second and provisioned duration T_k ,

$$C_{\text{prov}} = \sum_k \frac{\nu_k}{4096 \text{ s}^{-1}} T_k \quad [\text{RWS}].$$

The same ledger reports active and idle reference-worker seconds. Including W_{peak} in the mandatory vector prevents an arm from treating extra parallel workers or a higher service-rate multiplier as free merely because its operation count is unchanged. E is measured energy in

joules. Until calibrated workstation measurement exists, E is unavailable and cannot be replaced by O or C_{prov} .

One arm dominates another only if it is no worse on every registered endpoint and strictly better on at least one, with uncertainty and relevance margins applied as frozen in the experiment contract.

Let $\alpha > 0$ be a dimensionless coupling-strength multiplier, let $DU(\alpha)$ be dimensionless upstream distortion, let $\epsilon_{\text{track}}(\alpha, \omega)$ be dimensionless tracking error at angular frequency ω in radians per second, and let $\epsilon_{\text{leak}}(\alpha, \ell)$ be dimensionless error when each scheduled direct-reference release independently fails with dimensionless probability ℓ . If a continuous-time comparison is needed at constant scheduled reference-release rate ρ_{rel} in s^{-1} , the corresponding hazard is

$$\lambda_{\text{leak}} = -\rho_{\text{rel}} \ln(1 - \ell) \quad [\text{s}^{-1}],$$

which is approximately $\rho_{\text{rel}} \ell$ only for small ℓ . A low-coupling regime can satisfy

$$\frac{\partial DU}{\partial \alpha} > 0$$

while simultaneously satisfying

$$\frac{\partial \epsilon_{\text{track}}}{\partial \alpha} < 0 \quad \text{or} \quad \frac{\partial \epsilon_{\text{leak}}}{\partial \alpha} < 0.$$

Thus reducing back-action can worsen bandwidth or leak robustness. No universal monotone energy law follows from the coupling label.

Deliberate temporal-use comparator

Fixture F-027's RIN-T10 asks whether a declared back-action should be suppressed, preserved, or used for a temporal objective. Let $\tau_* \in \{0.25, 1, 4\}$ s be the target time constant and define the dimensionless coefficient

$$a = \exp\left(-\frac{\Delta t}{\tau_*}\right).$$

For isolated output y_n^{iso} in NOU, the causal target state is

$$y_{n+1}^* = y_n^{\text{iso}} + a(y_n^* - y_n^{\text{iso}}), \quad y_0^* = y_0^{\text{iso}}.$$

Here $y_n^{\text{iso}} = Hx_{n+1}^{\text{iso}}$ is the post-update isolated publication for due sequence n . Sequence n uses the already-existing y_n^* ; only after emission may the recurrence consume y_n^{iso} to form y_{n+1}^* . The state is initialized once and is not reset when clients detach at 150 s. The active scored reference is

$$r_n = \begin{cases} y_n^*, & t_n < 225 \text{ s}, \\ y_n^{\text{iso}}, & t_n \geq 225 \text{ s}. \end{cases}$$

For ordinary client j , let $\mathcal{A}_j$ be its exact set of due sequences while active, let $M_j = |\mathcal{A}_j|$, let $\tilde{y}_{j,n}$ be its delivered target sample, and let $m_{j,n} \in \{0, 1\}$ indicate that the sample is available, timely, current, and integrity-valid. With the registered missing penalty $P_{\text{target}} = 10$ NOU,

define

$$e_{j,n}^{\text{target}} = \begin{cases} \tilde{y}_{j,n} - rn, & m_{j,n} = 1, \\ P_{\text{target}}, & m_{j,n} = 0, \end{cases}$$

and

$$\text{RMSE}_j = \sqrt{\frac{1}{M_j} \sum_{n \in \mathcal{A}_j} (e_{j,n}^{\text{target}})^2} \quad [\text{NOU}].$$

The primary target endpoint is

$$\text{RMSE}_{\text{target}}^{\max} = \max_{j: M_j > 0} \text{RMSE}_j.$$

Client 1 and the pooled active-client RMSE are separate reports; neither can replace the maximum in the primary gate.

Sensitivity substitutes 2 and 100 NOU for P_{target} ; invalid samples never leave the denominator.

The explicit-filter null emits its current state before applying a recurrence. Let $M_{<225} = 4500$ be the number of due sequences before the switch and let $R_{<225} = M_{<225} - 1 = 4499$ be the number of useful next-state recurrences. The fixture's scalar-allocation and read/write law gives these nominal filter- specific totals for a complete chain:

$$O_{\text{filter}} = 50 + 2M_{<225} + 11R_{<225},$$

$$B_{\text{read,filter}} = 8M_{<225} + 88R_{<225} \text{ B}, \quad B_{\text{write,filter}} = 48 + 8M_{<225} + 8R_{<225} \text{ B}.$$

The physical ingress stores eight state components, not y_n^{iso} . Deriving that scalar reads 64 B and costs seven additions plus one multiplication. The displayed operation total expands to coefficient setup, the sequence-zero derivation and state initialization, $M_{<225}$ emissions, $R_{<225}$ recurrence cores, $R_{<225} - 1$ later input derivations, and final release. The constant 48 B written is the associated coefficient, state, allocator, and release ledger. Filter-specific peak live storage is 24 B: coefficient, state, and transient output. Missing ingress changes actual counts and is reconstructed from raw events; the nominal formulas cannot be applied to a broken chain. One fixed- TTL B-SNAPSHOT ingress is additional. Every successful pre-switch emission also creates one filter-owned typed 96 B source on the FCFS filter worker, paying the common allocation, initialization, transient-read, reference-acquisition, and eventual release charges. Its per-client descriptors pin that source until copy completion, cancellation, or source TTL; every destination then pays the ordinary B-SNAPSHOT allocation, copy, reference-release, destination-release, queue, and client-service charges. The filter worker serially executes coefficient setup, input derivation, emission, source creation, descriptor acquisition, and recurrence work, so these counts determine timing. From 225 s the arm cancels remaining pre-switch sources and copies, bypasses the recurrence and uses the ordinary immutable isolated-output route while retaining coefficient and state until episode end.

Numerical and dimensional checks

1. Every term in each differential equation must have the dependent variable's unit per second.
2. $0 \leq C(t) \leq p_{\text{tot}}$ and $X(t) \geq 0$ are invariants of valid biological-source simulations with nonnegative initial states and nonnegative production.
3. The binding-only contribution conserves $X + C$ to the registered solver tolerance.
4. The reduced formula is tested only against the full model; it is never its own oracle.

5. State and client trajectories are integrated on the same time grid before a paired discrepancy is evaluated.
 6. Fixed-step convergence is checked against half-step and quarter-step solutions or an independently configured adaptive solver.
 7. Event ties in the artificial system use this frozen order: expire snapshot TTLs and normally scheduled direct references; apply detachments and joins; complete copies; complete consumer/transform service and create valid feedback messages; complete and deliver feedback service; complete and route C-ADAPT decisions; record crash, restart, or version change; deliver input; update the producer and apply already delivered feedback; publish, execute the block/drop/queue rule, or enqueue a controller request; acquire an immediately routed publication; enqueue telemetry; score deadlines; append the event record.
 8. A count, byte, second, joule, and watt are never added without an explicit objective and dimensional conversion.
-

Validity and kill boundaries

1. A read of immutable state with abundant independent storage may have no direct output back-action. In that regime, the sequestration translation must collapse.
2. Competition among declared downstream transformations is not inferred from generic CPU delay; it requires a pathway-specific finite-service intervention.
3. General CPU, memory, transcription, or translation contention is not relabelled direct sequestration or substrate competition.
4. A coupling that improves a declared integration objective is not insulated automatically. Its benefit and harm must be evaluated under an ablation that preserves input information and total work.
5. Intended gradients, acknowledgements, feedforward controllers, or supervisory control remain declared useful coupling.
6. A reporter or monitor can itself be a downstream client; an observation is not assumed non-invasive.
7. A lower DU is not useful if client service, task accuracy, latency, memory, recovery, or complete lifecycle work becomes unacceptable.
8. Static equality does not imply dynamic modularity. Step, pulse, periodic, burst, and stochastic histories remain separate.
9. The source model does not establish that phosphorylation cycles evolved to insulate, or that the same mechanism exists in artificial systems.
10. The fixture is retired as an architecture contribution if ordinary snapshots, queues, backpressure, process isolation, admission control, or resource reservation match the complete frontier.
11. Logical operations and bytes are resource measures, not joules.
12. Every equation and experiment in this note remains **NO_RESULT** until a registered execution produces a valid artifact.

Interface-qualified scale symmetry

- **Purpose:** define and visualize full-trajectory fold-change detection, weaker lookalikes, observation-interface dependence, and the cost of preserving an absolute side channel
- **Claims:** C-1540--C-1549
- **Evidence audit:** relative sensing and scale symmetry
- **Experiment contract:** Fixture F-026
- **Result state:** analytical definitions and one illustrative exact model; no trained-model, biological, workstation or energy result

Notation, units, and interfaces

SYMBOL	MEANING	UNIT
t	time	seconds (s)
$x(t)$	internal state vector	model-declared
$u(t)$	positive external input	model-declared input unit
b	positive background input	same unit as u
p	positive multiplicative scale factor	dimensionless
G	registered set of admissible scale factors	dimensionless set
$y(t)$	output at one declared interface	model-declared
$r(t)$	positive reference state	same unit as u
τ_r	reference time constant	seconds (s)
m	pulse fold multiplier	dimensionless
$D_{j,p}$	normalized trajectory discrepancy	dimensionless
$s_{y,j}$	frozen nonzero scale for output component j	same unit as y_j

The observation chain is

$$u_{\text{external}} \longrightarrow u_{\text{internal}} \longrightarrow y_{\text{readout}} \longrightarrow a_{\text{action}}.$$

Scale symmetry at one arrow does not imply it at the next. The input, state, readout, target and action must therefore be named independently.

Exact full-trajectory definition

Consider a causal system

$$\dot{x} = f(x, u), \qquad y = h(x, u),$$

and let $x_{ss}(b)$ be the initialized steady state associated with positive background b . Exact FCD over scale support G requires

$$y(t; pu, x_{ss}(pb)) = y(t; u, x_{ss}(b))$$

for every registered input history, time, and $p \in G$. The equality is between complete trajectories. Equality of a final value, peak, integral or selected time point is insufficient.

Approximate FCD needs a frozen discrepancy, support, uncertainty model and acceptance margin. It is not licensed by a visually attractive overlay.

State-equivariance certificate

Assume the model has unique solutions on a declared forward-invariant positive domain, and that ρ_p maps that domain into itself for every $p \in G$. Let ρ_p transform internal state when the positive input is multiplied by p . Under those conditions, a sufficient certificate is

$$f(\rho_p(x), pu) = D\rho_p(x)f(x, u),$$

$$h(\rho_p(x), pu) = h(x, u),$$

and

$$\rho_p(x_{ss}(b)) = x_{ss}(pb).$$

$D\rho_p(x)$ is the Jacobian of the state transformation. These equalities belong to a defined model and observation map. Changing the output can change both invariance and identifiability.

Worked reference model

Use the positive-domain reference system

$$\tau_r \dot{r} = u - r, \quad y = \ln \frac{u}{r}.$$

Under the transformation $(u, r) \mapsto (pu, pr)$, the state equation scales by p while the output remains

$$\ln \frac{pu}{pr} = \ln \frac{u}{r}.$$

If $r(0) = b$, then the transformed initial state is $pr(0) = pb$, so the entire output trajectory is identical.

For a pulse that changes u from b to mb at t_0 and returns to b at t_1 , the reference during the pulse is

$$r(t) = b \left[m - (m - 1)e^{-(t-t_0)/\tau_r} \right], \quad t_0 \leq t < t_1,$$

and the relative output is

$$y(t) = \ln \frac{m}{m - (m - 1)e^{-(t-t_0)/\tau_r}}.$$

The background b cancels. After the pulse, define

$$q = (m - 1) \left(1 - e^{-(t_1-t_0)/\tau_r} \right).$$

Then

$$r(t) = b \left[1 + qe^{-(t-t_1)/\tau_r} \right], \quad y(t) = -\ln \left[1 + qe^{-(t-t_1)/\tau_r} \right],$$

which is again independent of b .

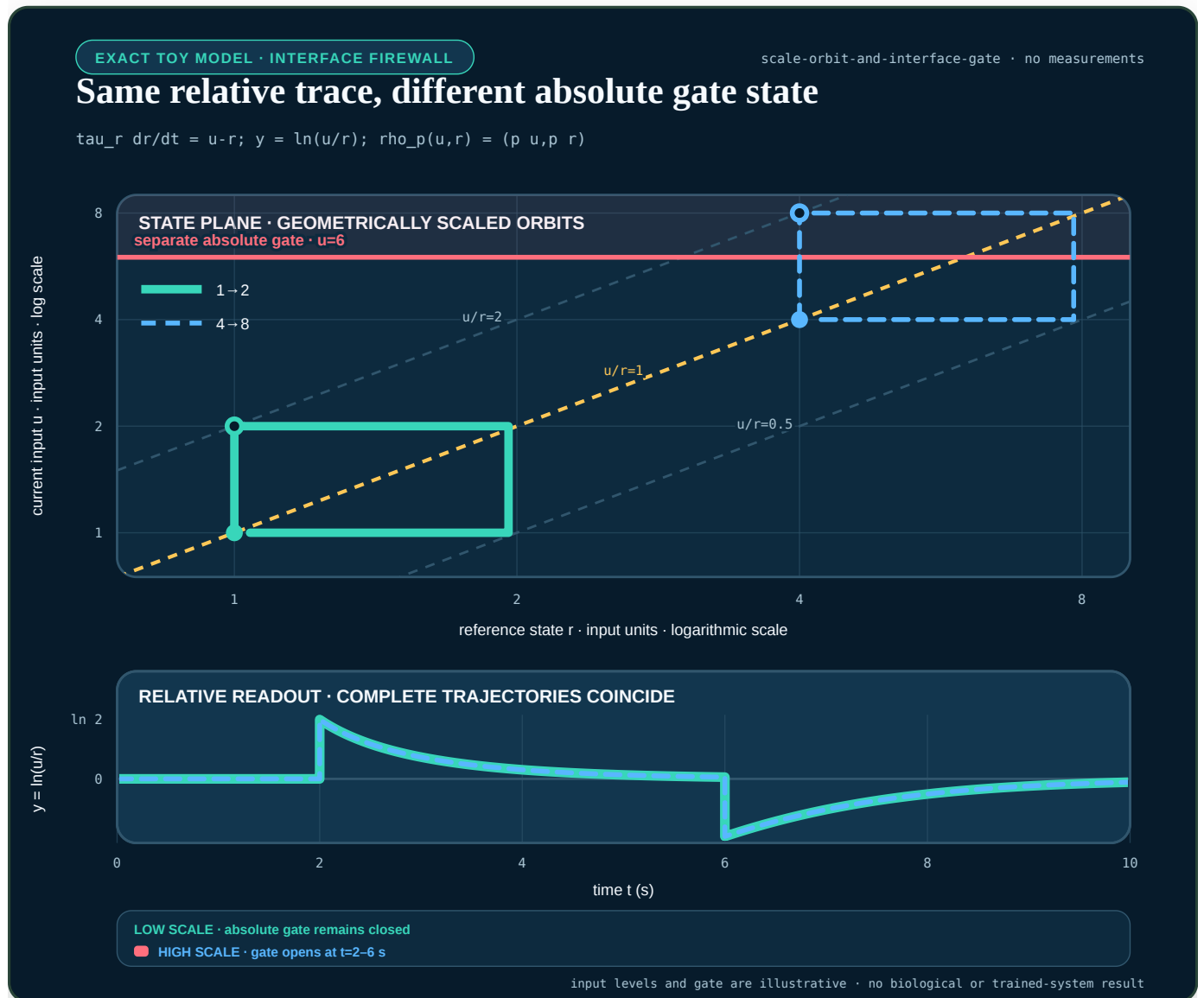


FIGURE 61 — Two geometrically scaled input/reference paths follow the same ratio contours and produce identical relative trajectories, while a separate absolute gate distinguishes them.

The figure uses $b \in \{1, 4\}$, $m = 2$, $\tau_r = 1.25$ s and an illustrative absolute gate at $u = 6$ input units. The relative traces coincide exactly in this model, but the $4 \rightarrow 8$ pulse crosses the gate while the $1 \rightarrow 2$ pulse does not. The gate is a mathematical counter-task, not a biological or safety threshold.

Weaker lookalikes

The constructions below are properties, not disjoint mechanism classes. A static ratio map with its reference supplied at the interface can satisfy exact trajectory symmetry, and endpoint return can coexist with equal peak. F-026 therefore keeps generator family as secondary provenance and evaluates cross-cutting properties separately while retaining their logical dependencies.

Finite-horizon endpoint return

A finite recorded trajectory can test return relative to each trajectory's own prestimulus output. On a frozen discrete horizon with endpoint tolerance ε_{end} , define

$$A_{\text{end},j,h,p} = 1[|y_{j,h,1,N} - y_{j,h,1,0}| \leq \varepsilon_{\text{end}} \wedge |y_{j,h,p,N} - y_{j,h,p,0}| \leq \varepsilon_{\text{end}}].$$

This finite-sample predicate is not general exact adaptation. A pulse can return because the input itself returned, and one forced final sample says nothing about causality or settling. Exact adaptation instead requires a declared sustained-input equilibrium or registered settling-tail condition; in the ideal infinite-horizon form,

$$\lim_{t \rightarrow \infty} y(t) = y_0.$$

Two systems can share y_0 while their peak, latency, width and tail differ.

Weber-like peak equality

Peak equality is measured as maximum absolute departure from each trajectory's own prestimulus value,

$$P_{j,h,p} = \max_k |y_{j,h,p,k} - y_{j,h,p,0}|, \quad P_{j,h,1} = \max_k |y_{j,h,1,k} - y_{j,h,1,0}|.$$

The first maximizing sample is the frozen latency tie-break. Equality of $P_{j,h,p}$ and $P_{j,h,1}$ does not constrain latency or the rest of the trajectory.

Static normalization

A static value u_t/r_t has no necessary causal rule for producing, aging, validating or resetting r_t . Its output can look relative while its reference uses future data or stale support.

Additive difference and derivative

For positive u and a frozen positive reference unit u_* with the same input unit,

$$\ln \frac{u(t)}{u_*} - \ln \frac{b}{u_*} = \ln \frac{u(t)}{b}$$

is invariant to a common multiplicative scale. In contrast, $u(t) - b$ is invariant to a common additive offset, and du/dt responds to a rate in input units per second. Limited step or ramp families can confound these statistics.

Approximate trajectory score

For output component j , registered input history $h \in \mathcal{H}$, held-out factor $p \in \mathcal{P}$, duration T seconds and frozen nonzero scale $s_{y,j}$, use

$$D_{j,h,p} = \sqrt{\frac{1}{T} \int_0^T \left\| \frac{y_{j,h,p}(t) - y_{j,h,1}(t)}{s_{y,j}} \right\|^2 dt}.$$

The score is dimensionless. The scale $s_{y,j}$, integration grid, time window, scale factors and uncertainty procedure must be frozen before confirmation. Report peak, latency, duration and tail discrepancies separately so a low integral error cannot hide one dangerous phase.

System-level classification uses the worst registered cell, not an unstated average:

$$D_{j,\max} = \max_{h \in \mathcal{H}, p \in \mathcal{P}} D_{j,h,p}.$$

With frozen tolerances $0 \leq \varepsilon_{\text{exact}} < \varepsilon_{\text{approx}}$, define the tolerance-qualified class

$$S_{\text{system},j} = \begin{cases} \text{exact}, & D_{j,\max} \leq \varepsilon_{\text{exact}}, \\ \text{approximate}, & \varepsilon_{\text{exact}} < D_{j,\max} \leq \varepsilon_{\text{approx}}, \\ \text{absent}, & D_{j,\max} > \varepsilon_{\text{approx}}. \end{cases}$$

Missing or rejected cells do not disappear from the maximum; they make the system classification unavailable and retain an abstention in its registered denominator.

If an actionable arm receives both complete output trajectories on this grid, $D_{j,h,p}$, paired trajectory match, finite-horizon endpoint return and peak equality are directly computable. Their errors measure numerical conformance and the resource cost of producing the values; they are not evidence that a representation learned the properties. A predictive interpretation needs a separately registered prospective observation cut, such as a withheld suffix, scale cell, or future intervention. That cut must state what the arm observes, what remains evaluator-only, and when its response freezes. Merely labelling a publicly enumerated scale as “withheld” does not create confirmatory secrecy.

The RSD-T01 development foundation therefore separates three objects:

1. a complete history $\times$ scale descriptor grid, including multiple scale cells per shared initialization and a predeclared prospective role;
2. valid scientific hostile worlds with explicit support membership and expected leakage behavior; and
3. malformed-record sentinels, which test fail-closed parsing and never enter the scientific denominator.

Its dimensionless scale sets are

$$\mathcal{P}_{\text{development}} = \{2, 4\}, \quad \mathcal{P}_{\text{prospective}} = \{8\}, \quad \mathcal{P} = \mathcal{P}_{\text{development}} \cup \mathcal{P}_{\text{prospective}}.$$

Every $p \in \mathcal{P}$ multiplies the same $p = 1$ reference trajectory and shares the registered initialization identity. “Prospective” is a public descriptor role here, not a concealed confirmation partition.

This descriptor foundation can test grid construction and aggregation while the predictive observation cut remains blocked. It has no comparator-result or claim authority.

For the current RSD-T01 contract the evaluator's property target is

$$\pi_j = (S_{\text{system},j}, \{A_{\text{end},j,h,p}\}_{h,p}, \{E_{\text{peak},j,h,p}\}_{h,p}, M_{\text{causal}}, \mathbf{Q}_{\text{membership},j}),$$

where $S_{\text{system},j} \in \{\text{exact}, \text{approximate}, \text{absent}\}$ is evaluated by $D_{j,\text{max}}$ across the complete frozen history $\times$ scale grid rather than inferred from one pair, $A_{\text{end},j,h,p}$ is the per-cell finite-horizon endpoint predicate, $E_{\text{peak},j,h,p}$ is the per-cell tolerance-qualified equality of $P_{j,h,p}$ and $P_{j,h,1}$, M_{causal} requires a machine-readable state equation or identifying intervention, and $Q_{\text{membership},j}$ is generator truth with one $\{\text{inside}, \text{outside}\}$ coordinate for each registered support axis: input domain, transformation family, instrument range, initialization contract, causal observation contract, and evaluation window. This vector prevents an additive transformation, clipped measurement, hidden reset, or future-aware normalizer from being collapsed into one ambiguous bit. An arm's support-detection decision is a separate output because detectability depends on its observation interface. The endpoint and peak matrices remain primary data. A compact summary may use only the frozen reducer `all` when every registered cell is true, `none` when every cell is false, and `partial` otherwise; it may not silently substitute a mean. These coordinates are asserted without using the generator-family name, but they are not logically independent: exact trajectory equality, for example, implies equal peak under the same peak definition.

The public generator-only smoke layer records a narrower per-world vector: paired-trajectory match, finite-horizon endpoint return, peak-amplitude equality, `causal_memory_status: unassessed`, and `support_membership: inside`. It does not promote one paired trace to a system symmetry certificate. That legacy scalar is a narrower v2 smoke field; it does not replace the axis-qualified vector in the scientific-grid foundation.

Property vector, not family label

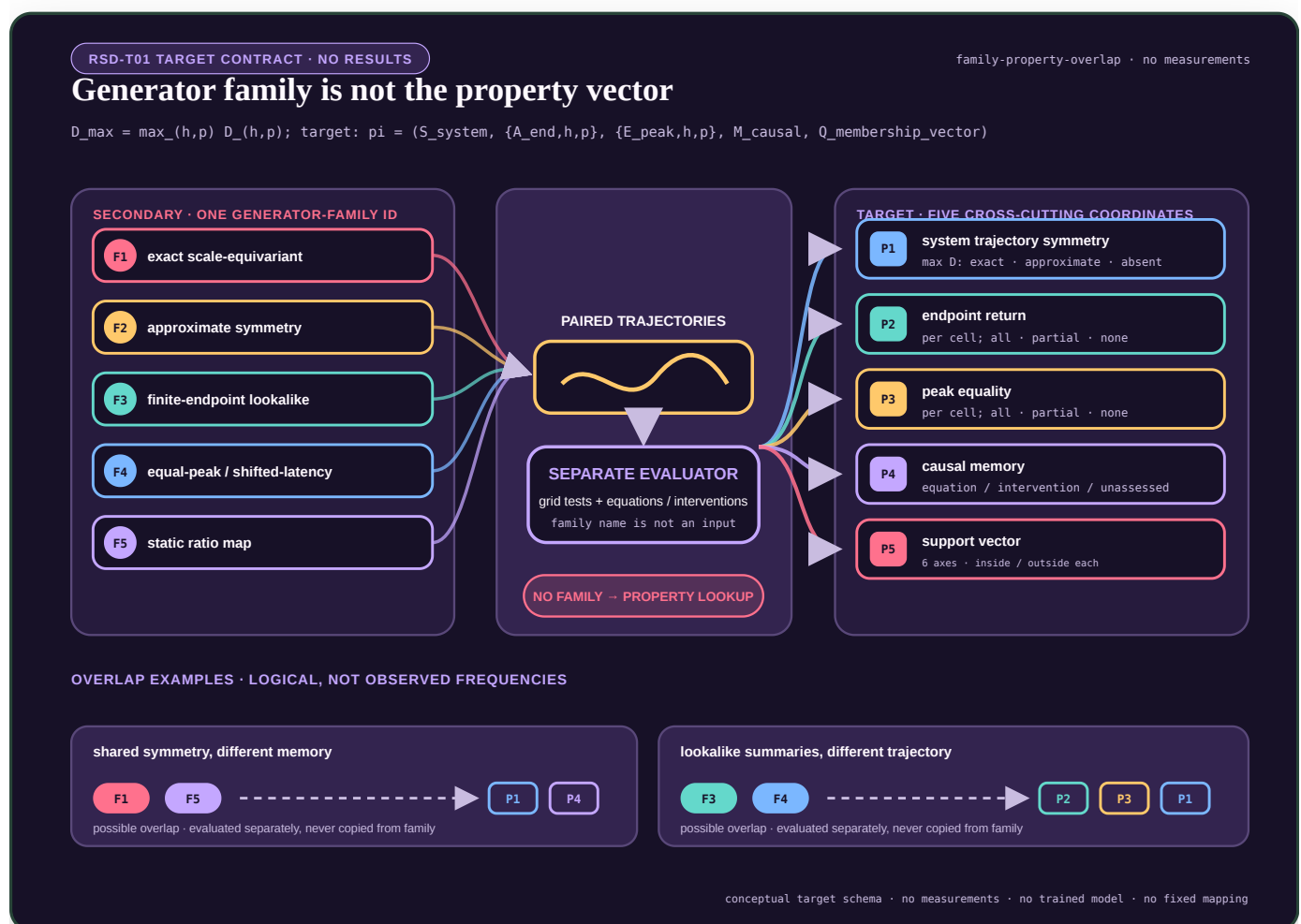


FIGURE 62 — Five generator-family labels converge through paired trajectories into a separate evaluator, which produces five cross-cutting property coordinates while retaining logical dependencies; two examples show that different families can share properties while differing on others.

The family ID remains a secondary synthetic diagnostic. The primary vector is evaluated from the registered trajectory grid, structural equations and interventions; it is never copied from a family-to-property lookup. The current smoke records only its directly observable subset and marks causal memory unassessed. The editable figure specification is the `rsd-t01-family-property-overlap` entry in `core-models.json`.

RSD-T02 uses the same rule at a deeper causal level. Its dedicated `intervention-qualified mechanism-equivalence` note constructs five exact matched-step recipes, scores separately certified structural properties and retains observational equivalence or abstention. It also keeps the Skataric--Nikolaev--Sontag fast-boundary-layer endpoint separate from the RMS score above: the source-qualified floor is a supremum-norm result on an ϵ -scaled temporal grid, while an integrated RMS discrepancy can vanish as the layer narrows.

Validity and failure boundaries

- 1. Positive support.** Log ratio is undefined at zero and changes meaning across sign. A hidden epsilon is not a scientific solution.
- 2. Saturation and clipping.** Multiplicative histories can become observationally identical for the wrong reason.
- 3. Additive shift.** A ratio-qualified model need not transfer across added backgrounds.
- 4. Reference age.** A stale or contaminated r_t changes the statistic even when the current u_t is valid.
- 5. Initialization.** Comparing trajectories from unmatched states does not test the stated symmetry.
- 6. Protocol support.** Steps alone do not certify pulses, ramps, stochastic histories or closed-loop action.
- 7. Interface.** Relative encoding upstream can become absolute or difference-based downstream.
- 8. Absolute target.** If the target depends on u rather than u/r , a relative-only representation is insufficient by construction.

Observation and recoverability

If an output is exactly invariant under a transformation, that transformed quantity can be structurally unidentifiable from that output interface under the declared initialized model. This is an observation-qualified statement, not universal destruction of information. Recoverability can change when one adds an output, fixes a parameter, changes initialization, or supplies a calibrated absolute observation.

F-026 therefore reports three axes separately:

1. robustness to nuisance multiplicative scale;
2. recoverability of every protected absolute quantity; and
3. complete observation, state, compute and maintenance cost.

Resource model

The authoritative report is a typed vector, not a sum of unlike units:

$$\mathbf{q} = (L_{\text{task}}, \mathbf{r}_{\text{risk}}, N_{\text{op}}, B_{\text{comm}}, B_{\text{state}}, T_{\text{wall}}, E_{\text{ws}}, N_{\text{ref}}).$$

Here L_{task} retains its task-native unit; $\mathbf{r}_{\text{risk}}$ retains one registered unit per risk component; N_{op} and N_{ref} are operation and reference-lifecycle counts; B_{comm} and B_{state} are bytes; T_{wall} is seconds; and measured workstation energy E_{ws} is joules. None is inferred from another.

If a frozen decision rule genuinely needs a scalar diagnostic, every component requires its own conversion weight into one declared decision unit:

$$J = \lambda_L L_{\text{task}} + \lambda_r^T \mathbf{r}_{\text{risk}} + \lambda_N N_{\text{op}} + \lambda_{Bc} B_{\text{comm}} + \lambda_{Bs} B_{\text{state}} + \lambda_T T_{\text{wall}} + \lambda_E E_{\text{ws}} + \lambda_U N_{\text{ref}}.$$

Reference updates, calibration, resets, selector execution, abstention, fallback, writes and artifacts must appear in the appropriate count, byte and time components. A biological FCD paper does not supply any of these AI conversion weights.

Test and retirement conditions

The relative translation remains viable only if it:

1. lowers held-out task loss or complete work against static, streaming, log-ratio, state-space and recurrent nulls;
2. preserves every absolute-critical counter-task or abstains before action;
3. survives non-step multiplicative histories inside declared support;
4. rejects additive, near-zero, saturated and stale-reference cases; and
5. retains its advantage after reference maintenance and fallback are charged.

Retire the translation if a simpler explicit transform or ordinary estimator matches it, if full-trajectory invariance collapses to peak equality, or if a calibrated absolute channel recreates the cost of keeping raw state everywhere.

Intervention-qualified mechanism equivalence

- **Purpose:** freeze the equations, property ontology, intervention interface, abstention rule and singular-boundary-layer metric for RSD-T02
- **Claims:** C-1541 and C-1560, C-1561, C-1562, C-1563, C-1564, C-1565, C-1566, C-1567, and C-1568
- **Evidence audits:** mechanism equivalence and intervention-qualified discrimination · population, identifiability, calibration, and null maturity
- **Decisions:** 0019 — score interventional properties, not generator names · 0020 — separate information cuts from scientific replication · 0021 — bind population inference to system lineages and instances
- **Experiment contract:** Fixture F-026, RSD-T02
- **Result state:** public-development equation and contract foundation; NO_RESULT, no implemented estimator comparison, no confirmation custody and no energy measurement

Two different questions

RSD-T02 contains two linked but mathematically different strata:

1. T02-MECH asks which structural or interventional properties can be distinguished after several synthetic systems are forced to have the same canonical step response.
2. T02-FLOOR asks whether an estimator recovers a source-qualified supremum-norm error floor in a singularly perturbed system.

The first is a model-discrimination problem. The second is a norm and temporal resolution problem. Combining them into one family label or one integrated trajectory score would make both endpoints ambiguous.

Notation and units

SYMBOL	MEANING	UNIT
t	physical time	seconds (s)
$uc(t)$	input on channel $c \in \{A, B\}$	input unit (U)
b	registered positive background	U
$vc(t) = uc(t)/b$	background-normalized input	dimensionless
F_*	canonical fold, fixed to 2	dimensionless
τ_*	matched response time constant	s
x, a, r_C, z	internal states	dimensionless
$\tilde{y}(t)$	internal output before a registered clamp	dimensionless
$y(t)$	reported output	dimensionless

SYMBOL	MEANING	UNIT
P	registered intervention panel	finite set
$D_{\infty}^{m,n,i}$	maximum output difference between recipes m, n under intervention i	dimensionless
$D_2^{m,n,i}$	RMS output difference	dimensionless
ϵ	fast/slow time-scale ratio	dimensionless
p	positive multiplicative scale factor	dimensionless
$E_{\infty}(\epsilon, p)$	maximum instantaneous FCD discrepancy	dimensionless
$E_2(\epsilon, p)$	integrated RMS discrepancy	dimensionless

The synthetic T02 input support uses $b \in \{0.5, 2, 8\} \text{ U}$, $\tau_{*} \in \{0.5, 1, 2\} \text{ s}$, and strictly positive normalized input. These are frozen benchmark design values, not biological estimates.

Exact matched-step construction

Every mechanism recipe starts at the steady normalized background $v = 1$ and receives the canonical step $v : 1 \rightarrow F_{*}$. All five recipes must produce

$$y_{*}(t) = e^{-t/\tau_{*}}, \quad t \geq 0.$$

This removes ordinary step fit as a discriminator by construction. The generator recipe remains evaluator-only provenance.

Input-driven incoherent feed-forward recipe

Use

$$\tau_{*}\dot{x} = v - x, \quad \tilde{y} = \frac{v - x}{F_{*} - 1}, \quad x(0^{-}) = 1.$$

The direct $v \rightarrow y$ path and delayed antagonistic $v \rightarrow x \rightarrow y$ path form the operational feed-forward structure. For a constant post-step input F_{*} ,

$$x(t) = F_{*} - (F_{*} - 1)e^{-t/\tau_{*}},$$

and therefore $\tilde{y}(t) = e^{-t/\tau_{*}}$.

This is a synthetic reduced recipe inspired by the functional structure. It is not asserted to be the molecular equation set of a particular organism.

Nonlinear output-feedback recipe

Let

$$\tilde{y} = \frac{v - a}{F_{*} - 1},$$

$$\tau_{*}\dot{a} = (F_{*} - 1)y + \kappa(v - 1)(v - F_{*})y^2, \quad a(0^{-}) = 1, \quad \kappa = 0.25.$$

Normally $y = \tilde{y}$. During the registered output-clamp intervention, $y = 0$ is forced inside the feedback edge while $\tilde{y}$ remains evaluator-visible after response freeze. The nonlinear term is zero at both $v = 1$ and $v = F_{*}$. On the canonical step,

$$\tau_* \dot{a} = F_* - a,$$

which gives the same $y_*(t)$. Away from that step and under the output clamp, the state update differs.

Channel-local receptor/reference memory

For $c \in \{A, B\}$, use

$$\tau_* \dot{r}_c = \chi_c(t) [v_c - r_c], \quad r_A(0^-) = r_B(0^-) = 1,$$

$$\tilde{y} = \frac{v_{c_{\text{active}}} - r_{c_{\text{active}}}}{F_* - 1}.$$

$\chi_c(t) = 1$ only for the active channel. The inactive channel retains its own reference. The first active-channel step again yields $y_*(t)$, while same-channel and cross-channel restimulation test state locality.

Static normalization plus an ordinary high-pass readout

Define the static affine fold transform

$$\phi_{\text{lin}}(v) = \frac{v - 1}{F_* - 1},$$

followed by

$$\tau_* \dot{z} = \phi_{\text{lin}}(v) - z, \quad \tilde{y} = \phi_{\text{lin}}(v) - z, \quad z(0^-) = 0.$$

The normalization reference $\bar{b}$ is static, but the complete system is not memoryless: the ordinary high-pass filter has causal state. On the canonical step $\phi_{\text{lin}} = 1$, so $z = 1 - e^{-t/\tau_*}$ and $\tilde{y} = y_*(t)$.

This recipe is input-output isomorphic to the reduced I1-FFL recipe under the registered affine interface. Their names must not be forced apart.

Explicit log difference plus the same readout order

For $v > 0$, define

$$\phi_{\text{log}}(v) = \frac{\ln v}{\ln F_*},$$

$$\tau_* \dot{z} = \phi_{\text{log}}(v) - z, \quad \tilde{y} = \phi_{\text{log}}(v) - z, \quad z(0^-) = 0.$$

The canonical step again has $\phi_{\text{log}} = 1$. Other folds and ramps distinguish the log transform from the affine transform. Nonpositive input is outside support and forces abstinence; no hidden numerical epsilon is inserted.

Matched-step certificate

For recipes m and n , background b , time constant τ_* and output samples t_k , define

$$D_{\text{step},\infty}^{m,n} = \max_k |y_m(t_k) - y_n(t_k)|.$$

A public-development pack is invalid unless:

1. every initialization residual is at most 10^{-12} ;
2. every registered recipe pair has $D_{\text{step},\infty}^{m,n} \leq 10^{-10}$;
3. the canonical step policy projection contains no recipe, equation, state, parameter or property identifier; and
4. every actionable arm receives the same projection hash.

The tolerances are benchmark design constants. They are not empirical biological margins and grant no result authority.

Property vector and equivalence

The primary target is the evaluator-certified vector

$$\pi = (T_{\text{drive}}, F_y, C_{\text{local}}, M_{\text{causal}}),$$

where:

1. $T_{\text{drive}} \in \{\text{affine-fold}, \text{log-fold}\}$;
2. F_y states whether the reported output participates in a state-update feedback edge;
3. C_{local} states whether reference state is channel-local;
4. $M_{\text{causal}} \in \{\text{true}, \text{false}, \text{unassessed}\}$ is certified from equations or an identifying intervention, never from a finite trace alone.

Nonlinear update form remains equation provenance in the v1 bank. It perfectly co-varies with the output-feedback coordinate across these five worlds, so the contract cannot honestly score it as a separately identified property without adding a counterworld that breaks that dependence.

For intervention panel P , observation map h , sample grid $\mathcal{T}$ and initialized recipes m, n , define

$$D_{\infty}^{m,n,i} = \max_{t_k \in \mathcal{T}_i} |y_m^i(t_k) - y_n^i(t_k)|,$$

$$D_2^{m,n,i} = \sqrt{\frac{1}{T_i} \int_0^{T_i} [y_m^i(t) - y_n^i(t)]^2 dt}.$$

A pair with different property vectors is numerically separated only when one frozen intervention has estimate $\widehat{D}_{\infty}^{m,n,i}$ and numerical error bound η satisfying

$$\widehat{D}_{\infty}^{m,n,i} - \eta \geq 10^{-3}, \quad \eta \leq 10^{-8}.$$

An analytic equivalence or a complete bounded finite-grid equivalence with $\widehat{D} + \eta \leq 10^{-10}$ may retain both recipes. A claimed analytic equivalence that conflicts with its numerical bound is invalid rather than silently unresolved. Every other case is unresolved.

The I1-FFL and affine high-pass recipes deliberately share one operational property vector and one full-panel equivalence class. Guessing between their hidden names is an error, not added accuracy.

The v1 construction certificates are scoped to $b = 2 \text{ U}$, $\tau_* = 1 \text{ s}$, a 24 s horizon, 64 output samples per second and binary64 RK4. Discontinuous commands use half-open intervals and the left limit for the final RK4 stage ending on an event; the next step starts from the right-limit command. Distances at internal steps $1/1024 \text{ s}$ and $1/2048 \text{ s}$ must differ by at most 10^{-12} . The conservative construction lower bounds are 0.23 for the clamp and cross-channel certificates and 0.07 for the affine-versus-log ramp certificate. These margin checks reproduce a synthetic construction; they are not empirical mechanism evidence or confirmation results.

Nested intervention panels

The full fixed panel contains exactly 26 episodes:

1. three canonical steps at the three backgrounds;
2. six repeated-pulse cells obtained from $w \in \{1/8, 1/2, 2\} \text{ s}$ and $\Pi \in \{1/2, 2, 8\} \text{ s}$ with $w < \Pi$;
3. eight ramps: linear or exponential, up or down, each lasting 0.5 or 4 s;
4. two opaque-state resets;
5. two opaque-state freezes;
6. one reported-output clamp;
7. two interrupted-ramp holds lasting 0.5 or 4 s; and
8. same-channel and cross-channel restimulation.

Every episode lasts 24 s. The noncanonical episodes use $b = 2 \text{ U}$ and $\tau_* = 1 \text{ s}$. Periodic pulses have $v(t) = 2$ on $[j\Pi, j\Pi + w)$ while that interval remains inside the episode and $v(t) = 1$ otherwise. For ramp duration d , let $s(t) = \min\{1, \max\{0, t/d\}\}$. Linear-in-fold ramps use

$$v(t) = 1 + s(t)(v_1 - 1),$$

and linear-in-log-fold ramps use

$$v(t) = \exp[s(t) \ln v_1],$$

with $v_1 = 2$ for up-ramps and $v_1 = 0.5$ for down-ramps. Reset occurs at 0.75 s; state freezes and the reported-output clamp occupy $[0.5, 1) \text{ s}$. Interrupted ramps pause after one second of active ramp progress and resume after the registered hold. Restimulation drives channel A on $[0, 1) \text{ s}$, returns both channels to one on $[1, 2) \text{ s}$, then drives A or B on $[2, 3) \text{ s}$. The machine contract carries these numbers in each episode descriptor.

Every recipe exposes two opaque handles. A single-state recipe receives an inert padding state, and a seed-derived hidden permutation maps states to handles. State count and semantic node names therefore do not identify the recipe.

The observation regimes are:

1. **00-MATCHED-STEP**: three background episodes and at most 4,611 sample rows per conditioned τ_* model instance. The construction runtime crosses $\tau_* = 0.5, 1, 2, \text{ s}$, so each recipe has nine executions and 13,833 rows; every varying structural coordinate requires abstention.
2. **01-FULL-PANEL**: all 26 episodes and at most 39,962 sample rows at the fixed $\tau_* = 1, \text{ s}$ construction scope; this is the full-panel regime.
3. **02-SELECT6**: the three step episodes plus at most six selected queries, no more than two privileged internal queries, and at most 13,833 sample rows; it is a secondary active-design regime.

The current code freezes these regimes and analytic construction certificates. An additive whole-system Stage 2 now implements all nine registered public-development policy-conformance references. These fixed policies close the executable feature-family matrix; they do not constitute trained estimators, calibrated posteriors, mature nulls, a claim-eligible run or a comparison.

Fixed whole-system policy-conformance references

Let the ordered packet be $P = (P_1, \dots, P_{35})$ with 53,795 sample rows. All nine active policies receive the same canonical bytes and common cap. Their responses are committed before `0-GRAPH` opens any member of P . The bands below were chosen after inspecting the five enumerated public construction worlds. They are therefore **construction-tuned protocol constants**, not fitted parameters or confirmation-calibrated decision limits.

Each packet is evaluated in a fresh Node child and one new hardened VM context. The self-contained bank evaluates the nine ordered policies in that context. The child receives one canonical LF JSON request and can read only the verified SHA-named policy bundle. The policy VM receives no process, filesystem, network, environment, clock, random or evaluator capability. Request, packet, configuration, bundle and runtime identities are bound into the returned receipt. Time, memory, request, stdout and stderr are capped, and any timeout, crash, malformed frame, replay or work-envelope violation becomes an ordered pre-evaluator abstention with no retry or same-process fallback. The runner atomically persists that outcome as a self-hashed `rsd-t02-arm-abstention.json`, replay-binds it on later invocations and forbids it from coexisting with the commitment or evaluator ledger. Commitment creation remains exclusive and file-synchronized before the raw evaluator ledger opens. The generator and evaluator are still statically loaded before their later file fingerprints, so concurrent repository mutation across that parent-module load boundary remains outside this public-development authority; the policy computation itself executes from the verified content-addressed bundle.

The six transform-policy references added in Stage 2b are deliberately small and causal. Let y_k be the reported output, u_k the active-channel input, u_0 the same-episode initial background, $r_k = u_k/u_0$, and $\Delta t = 1/64$ s.

`A-RAW` uses no engineered drive coordinate. It reads the intervention traces directly:

$$z_f = |y_{\text{CLAMP}}(1s)|, \quad z_c = \max_{2 \leq t \leq 3} |y_{\text{RESTIM}}(t)|,$$

and uses the reset/freeze distance h defined below. It declares a feedback edge for $z_f \geq 0.5$, no edge for $z_f \leq 0.45$, local channel state for $z_c \geq 0.9$, shared state for $z_c \leq 0.85$, memory for $h \geq 0.1$, and no memory for $h \leq 10^{-8}$. It always abstains on the drive transform.

`B-STATIC-DIV` evaluates the ramp at $t = 0.25$ s using

$$z_s = \frac{y(t)}{r(t) - 1}.$$

It declares log-fold for $z_s \geq 0.98$, affine-fold for $z_s \leq 0.92$, and otherwise abstains. `B-LOG-RATIO` instead uses the positive-domain coordinate

$$z_\ell = \frac{y(t)}{\log_2 r(t)},$$

declaring log-fold for $z_\ell \geq 0.85$, affine-fold for $z_\ell \leq 0.8$, and abstaining when the logarithm is undefined or the evidence lies in the gap.

`B-DIFFERENCE` uses only the first 64 ramp increments. With $\dot{y}_k = (y_k - y_{k-1})/\Delta t$ and $\dot{r}_k = (r_k - r_{k-1})/\Delta t$, its projection coefficient is

$$z_d = \frac{\sum_{k=1}^{64} \dot{y}_k \dot{r}_k}{\sum_{k=1}^{64} \dot{r}_k^2}.$$

It declares affine-fold for $z_d \geq 0.78$, log-fold for $z_d \leq 0.77$, and otherwise abstains.

B-STREAM processes each selected trace in chronological order. For $\alpha = \exp[-\Delta t/(0.25 \text{ s})]$, it updates

$$\delta_k = y_k - m_{k-1}, \quad m_k = \alpha m_{k-1} + (1 - \alpha) y_k, \quad v_k = \alpha v_{k-1} + (1 - \alpha) \delta_k^2,$$

and scores the causal standardized innovation $z_k = |\delta_k|/(\sqrt{v_k} + 1)$. Clamp-release evidence at one second uses true/false thresholds 0.38/0.30; cross-channel restimulation at two seconds uses 0.94/0.85. Intermediate evidence abstains.

C-DUAL requires all three drive votes from z_s , z_ℓ , and z_d to be present and identical. It then carries the supported raw intervention decisions for the other three coordinates. Missing or discordant drive votes produce an abstention; the frozen fallback count is zero.

For **B-STATE-SPACE**, define the fold $r_k = u_k/u_0$ and two fixed drives

$$\phi_{\text{aff}}(r) = r - 1, \quad \phi_{\text{log}}(r) = \log_2 r.$$

For each drive, the reference recursion and one-step reported-output residual are

$$x_{k+1} = x_k + \Delta t [\phi(r_k) - x_k], \quad e_k = y_k - [\phi(r_k) - x_k].$$

With $S_j = n^{-1} \sum_k e_{j,k}^2$, the margin $m = S_{\text{log}} - S_{\text{aff}}$ declares affine for $m \geq 10^{-5}$, log for $m \leq -10^{-5}$, and otherwise abstains. Its memory signature is

$$h = \max\{\|y_{\text{RESET-H0}} - y_{\text{RESET-H1}}\|_\infty, \|y_{\text{FREEZE-H0}} - y_{\text{FREEZE-H1}}\|_\infty\}.$$

It declares memory present for $h \geq 0.1$, absent for $h \leq 10^{-8}$, and otherwise abstains. Feedback-edge and channel-local coordinates are always outside this reference's bounded scope.

For **B-RECURRENT**, the causal state update is

$$x_{k+1} = \alpha x_k + (1 - \alpha) y_k, \quad \alpha = \exp[-\Delta t/(0.25 \text{ s})].$$

The absolute innovation at reported-output clamp release declares a feedback edge for values at least 0.45, no edge for values at most 0.35, and otherwise abstains. A separate state per observed channel gives the cross-channel restimulation innovation; values at least 0.9 declare local state, values at most 0.85 declare shared state, and intermediate values abstain. Drive and causal-memory coordinates remain outside this reference's scope.

C-MECHANISM-BANK uses four direct, frozen source-shaped signatures: linear up-ramp output at 0.25 s (log at least 0.48, affine at most 0.46), absolute clamp-release output (feedback at least 0.5, absent at most 0.45), maximum absolute cross-restimulation output on $[2, 3]$, s (local at least 0.9, shared at most 0.85), and h above (memory at least 0.1, absent at most 10^{-8}). Every indifference band forces abstention. The five candidate equation identities and their bytes are charged separately from the four distinct joint property-prior vectors and their bytes.

The semantic output is the joint set

$$\mathcal{V}(P) = \{v(M) : M \in \mathcal{M}, v_q(M) = \hat{v}_q \text{ for every declared coordinate } q\},$$

with duplicate vectors removed but their compatible hypothesis IDs retained. $\mathcal{V}(P)$ must be nonempty, and every declared marginal must agree with every member of $\mathcal{V}(P)$. This prevents independently plausible marginals from forming an impossible property combination.

Resource accounting is three-part: (1) shared acquisition of 35 episodes, 53,795 rows, 197 input commands, two resets, two freezes, one output clamp, one channel switch and two state writes; (2) policy construction/prior artifacts, threshold provenance, equations, vectors, labels and tuning; and (3) actual per-inference work. Every policy is charged 12 declared traversal operations per row, or 645,540 before its specific operations. Common caps are 10^6 scalar operations, 2,000 transcendental evaluations, 128 retained-state bytes, 4,096 influential-parameter bytes, 16 MiB scratch, 256 KiB combined policy and configuration artifacts, and zero fallbacks. Actual counts are not padded to the caps. Across the nine references, charged scalar work ranges from 645,544 (B-STATIC-DIV and B-LOG-RATIO) to 688,576 (B-STATE-SPACE), including the common packet traversal. Wall time and joules remain null.

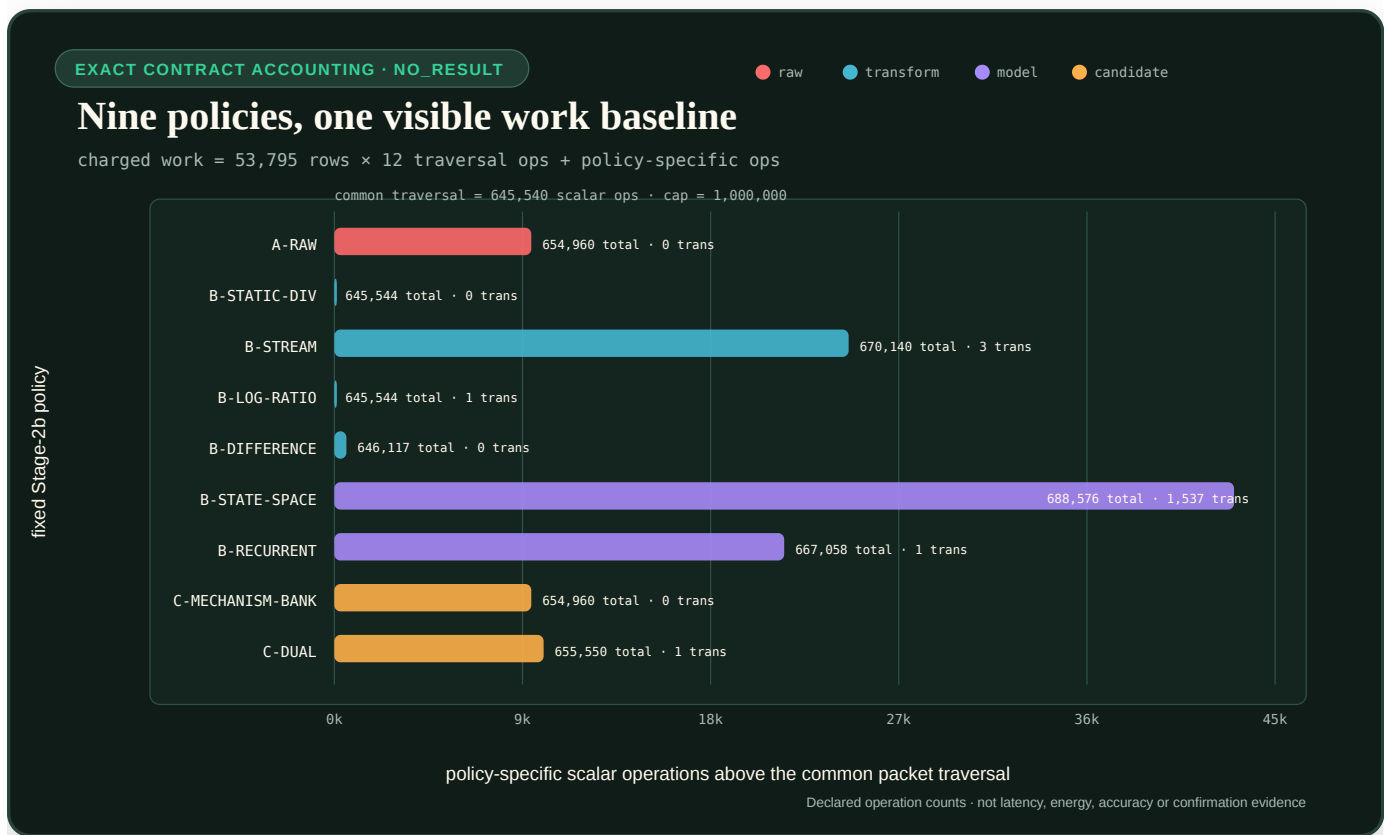


FIGURE 63 — Exact policy-specific scalar work above the shared packet-traversal baseline.

The plot exposes the common charge and the remaining policy-specific work without turning the declared scalar-operation model into a wall-time, energy or architecture-ranking claim.

The repeated-pulse grid is retained because the Rahi evidence makes refractory stabilization and period skipping useful one-sided signatures in a different bounded model class. The current five-world bank does not instantiate that signature: its registered feedback nonlinearity is zero at both square-pulse levels. C-1561 is specified separately in the repeated-stimulus topology-signature contract; it is not a claim supported by these five-world construction tests.

Prospective fit, calibration and evaluation cut

The Stage-3 design partitions the ordered 64-seed public pack once: ordered positions 1--32 are `fit`, 33--48 are `calibration`, and 49--64 are `evaluation`. The literal seed labels remain the frozen values `1540001`--`1540064`; the position numbers are not substitute seeds. Parameters and fit-only model selection stop at the first boundary; probability calibration plus support and abstention thresholds stop at the second; evaluation is one-pass frozen inference and scoring. Every future artifact must bind the preceding artifact and the exact partition identity.

That split controls access, not replication. In the present generator, a seed selects one of only two hidden permutations of two opaque state handles. The permutation can swap which internal coordinate a reset or freeze targets, but it does not sample a new equation, parameter set, input history or noisy system. The 64 labels therefore cannot be analyzed as 64 independent systems, and the public split has no comparison or power authority. This is the scoped experimental-unit problem described by Hurlbert (1984), bibliography key `hurlbert1984pseudoreplication`, not a statement that seeds can never be valid units in other generators.

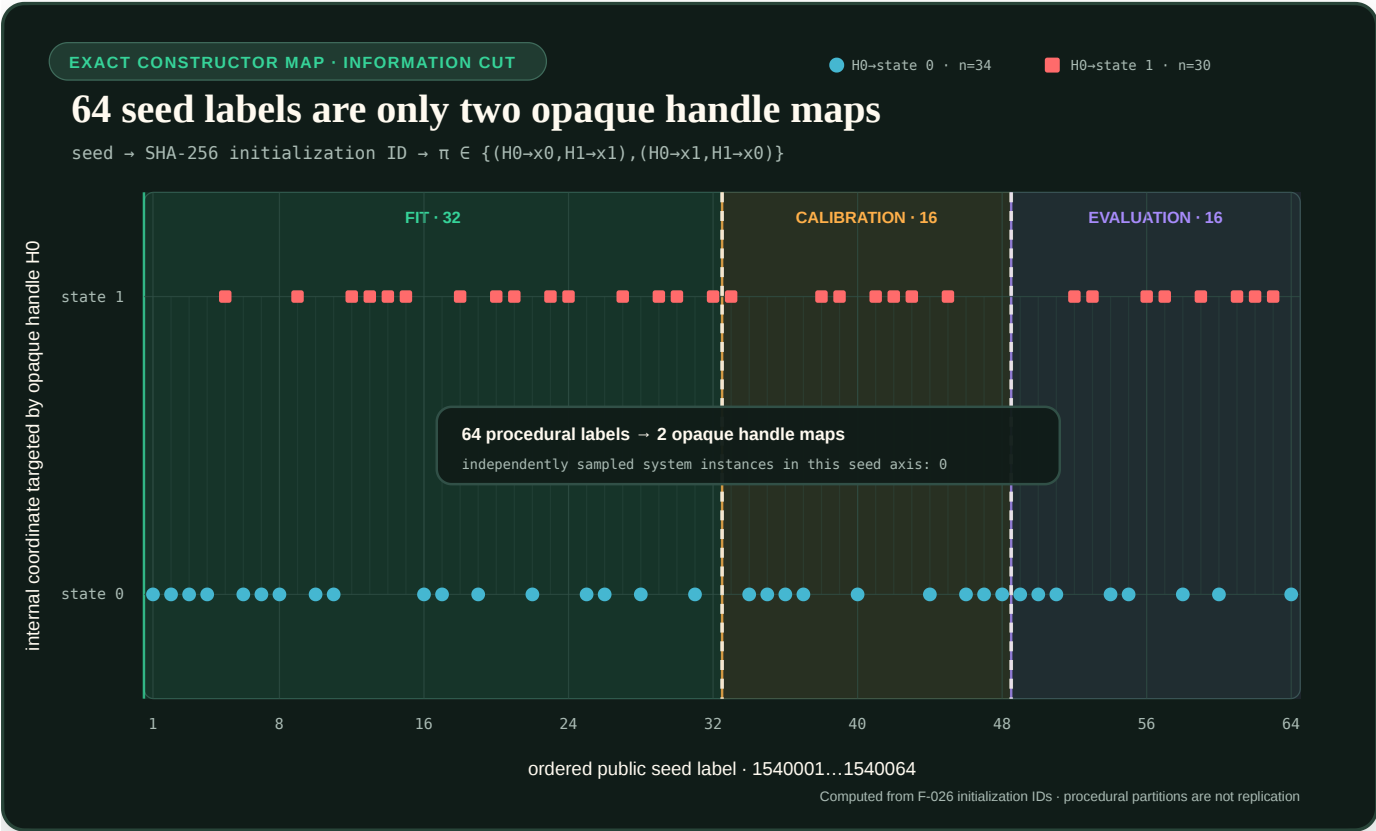


FIGURE 64 — The 64 procedural seed labels map to only two opaque state-handle permutations.

The plotted points are computed from the checked-in initialization-ID and opaque-permutation functions. The colored regions show the access cut; they do not add independent system variation.

The two generic references become mature nulls only after trainable causal state-space and compact recurrent estimators are implemented against the same fixed-parameter packet schema, with no direct plant state, recipe or equation access. Their construction, selection, calibration, failures and fallbacks enter the resource ledger. A later confirmatory comparison needs independently generated held-out system instances and an outer system-family holdout; neither exists in the present five-world bank.

For that later design, the fixed primary endpoint families are mean property log loss in nats and mean dimensionless decision loss. Coverage, selective risk, reliability, compatible-vector coverage and the resource vector are reported alongside them. The earlier Stage-3 wording grouped the two candidate-versus-generic-null contrasts within each endpoint. The later population contract supersedes that weaker boundary and uses the sequentially rejective

procedure of Holm (1979), bibliography key `holm1979sequential`, once across all four fixed endpoint-by-comparator hypotheses at familywise $\alpha = 0.05$. Sample size must be powered before private confirmation seeds are created; 16 public evaluation labels are not assumed sufficient.

The closed machine form is `rsd-t02-stage3-design.json`, validated by `rsd-t02-stage3-design.mjs`.

Prospective system population and outer-family boundary

The information cut above remains valid, but it is not a population design. The next contract uses the following nesting, from inferentially broad to repeated measurement:

design $\supset$ structural lineage $\supset$ family $\supset$ independent instance $\supset$ packet $\supset$ episode $\supset$ realization.

A procedural seed is only a replay key. A family is one frozen equation template plus a declared parameter distribution. An instance is one accepted parameter vector drawn independently from that family; it is the unit for a fixed-family estimand. Episodes, rows, property coordinates, solver refinements and noise realizations remain nested measurements and never increase the reported independent n .

This distinction also invalidates a tempting reuse of the current packet. Its nine O0 executions cross $\tau_* = 0.5, 1, 2$ s, while its 26 O1 executions hold $\tau_* = 1$ s. One population instance must bind one parameter vector—including one time constant—across every episode. The old mixed- τ packet remains a construction-conformance artifact; a population packet must be regenerated per fixed parameter vector.

The first defensible claim mode is a **fixed finite family panel**. Visible development families and separately sealed outer-confirmation and outer-transfer families are split by structural lineage, not by recipe label or instance. Related equations, derivations, code siblings and full-panel-equivalent recipes share a leakage group and cannot cross partitions. An outer result therefore generalizes only to the prospectively frozen weighted panel, not to arbitrary future mechanisms. A family-superpopulation claim requires a separate frozen probabilistic family grammar and powers on independently drawn families or lineages.

For coordinate q , instance i , family f and arm a , first aggregate the coordinate loss inside the instance,

$$\ell_{a f i} = \sum_q w_q L_{a f i q}, \quad d_{f i}^{a, b} = \ell_{a f i} - \ell_{b f i}.$$

Then form the equal-family weighted contrast

$$\bar{\Delta}^{a, b} = \sum_f w_f \bar{d}_f^{a, b}, \quad \bar{d}_f^{a, b} = \frac{1}{n_f} \sum_i d_{f i}^{a, b}, \quad w_f = \frac{1}{F}.$$

Resampling occurs over instances within the fixed families; rows are never resampled as independent systems. The prospective power artifact must freeze the smallest effect of interest, alpha, power, family heterogeneity, family and instance counts, abstention coverage, failure disposition and the calculation implementation hash before private responses. The experiment-level success rule applies Holm's procedure across all four fixed endpoint-by-comparator hypotheses at $\alpha = 0.05$; controlling two contrasts separately inside each endpoint would not close the cross-endpoint multiplicity boundary.

For equal retained counts in each of F fixed families, first derive the minimum effective count for one lower-tail contrast:

$$n_{\text{eff}} = \max \left\{ 2, \left\lceil \frac{\sum f \sigma_f^2}{F^2} \left(\frac{z_{1-\alpha/4} + z_{1-\beta}}{\delta} \right)^2 \right\rceil \right\}.$$

Here σ_f^2 is the development-evaluation variance of the paired system-instance contrast in family f , $\delta > 0$ is the minimum relevant improvement magnitude in the endpoint unit, and $1 - \beta$ is target power. The $\alpha/4$ term protects the most conservative first Holm step.

If r is the prospective probability of pre-response invalid generation and γ is the registered experiment-level retention assurance, the plotted planned count is instead

$$n_{\text{plan}} = \min \left\{ m \in \mathbb{N} : \Pr[\text{Binomial}(m, 1 - r) < n_{\text{eff}}] \leq \frac{1 - \gamma}{F} \right\}.$$

The Bonferroni allocation on the right guarantees at least γ retention assurance across the F family strata without assuming their attrition events are independent. Runtime failures retain their registered in-denominator penalty. The support-coverage floor remains a separate gate and is not silently converted into another sample-size multiplier.

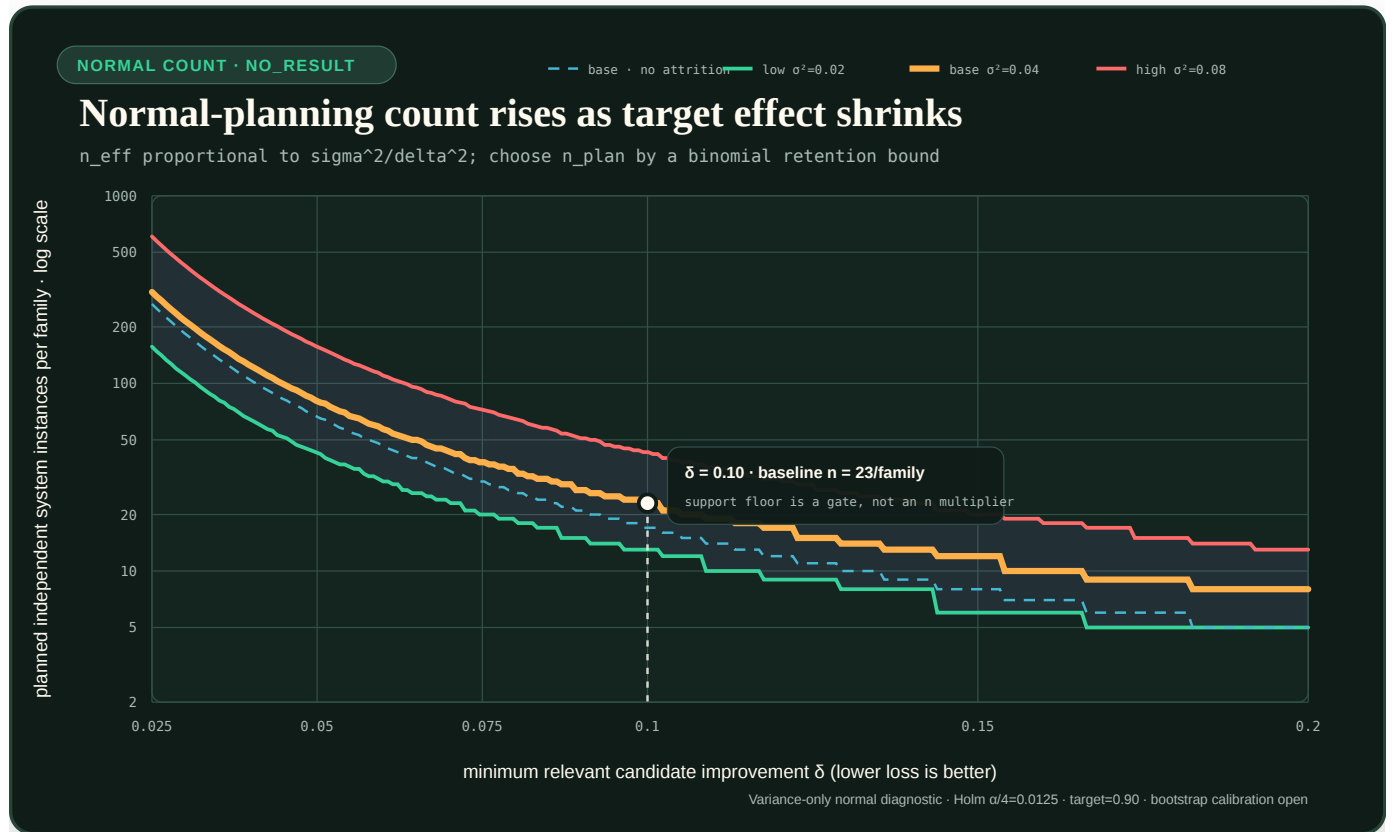


FIGURE 65 — Illustrative prospective independent-system count versus minimum relevant effect under three variance profiles.

The figure is a sensitivity map, not a power result. Its variance profiles are illustrative, so it cannot freeze a sample size. The checked-in calculator requires a development-evaluation variance-artifact hash and exposes every unit, approximation, attrition rule and blocker. A syntactically valid hash does not verify the artifact bytes, role or review, so the calculator always keeps the power-plan release gate open until those bindings are independently validated. The normal calculation also does not establish power for the final bootstrap- t analyzer: zero-standard-error bootstrap resamples create a data-dependent lower bound

$$p_{\text{min,data}} = \frac{B_0 + 1}{B + 1},$$

where B is the registered resample count and B_0 is the number of zero-standard-error resamples conservatively counted as extreme. The analyzer records this bound for every hypothesis and closes its resolution gate only when it can reach the first Holm threshold. A future frozen power artifact must therefore simulate the pilot transcripts through the exact analyzer; variance alone is insufficient.

Synthetic transcript calibration of the exact analyzer

The public diagnostic now performs that simulation step for four declared synthetic scenarios without treating the resulting frequencies as scientific power. For k events in R Monte Carlo replicates it reports the smoothed point diagnostic

$$\hat{p}_{MC} = \frac{k+1}{R+1}, \quad \Delta p_{MC} = \frac{1}{R+1}.$$

The point smoothing and interval have different jobs. The Wilson score interval is computed for the observed binomial count k/R ; applying Wilson to the artificial pair $(k+1, R+1)$ would give the wrong coverage target and can exclude zero even when a deterministic hostile produces no events.

The generated unit identity uses a DGP-only fingerprint. It binds scenario baselines, attrition, failure probabilities, contrast means and covariances, the simulation key and registered family order. Confidence level, total replicate ceiling, bootstrap count, alpha and endpoint penalties remain in the full configuration/report identity but do not reorder a common generated prefix or change its finite-bootstrap point decision.

The canonical configuration uses $R = 99$ and $B = 1000$. Its declared synthetic null yields 6 any-rejection events, while the declared minimum-relevant-effect scenario yields 55. The null Wilson Monte Carlo interval, 0.028--0.126, spans the 0.05 reference; the alternative interval, 0.457--0.650, is far below the illustrative 0.90 target. The two two-instance hostiles each fail the analyzer's bootstrap-resolution gate in all 99 replicates. This rejects plan acceptance; it does not estimate future model power.

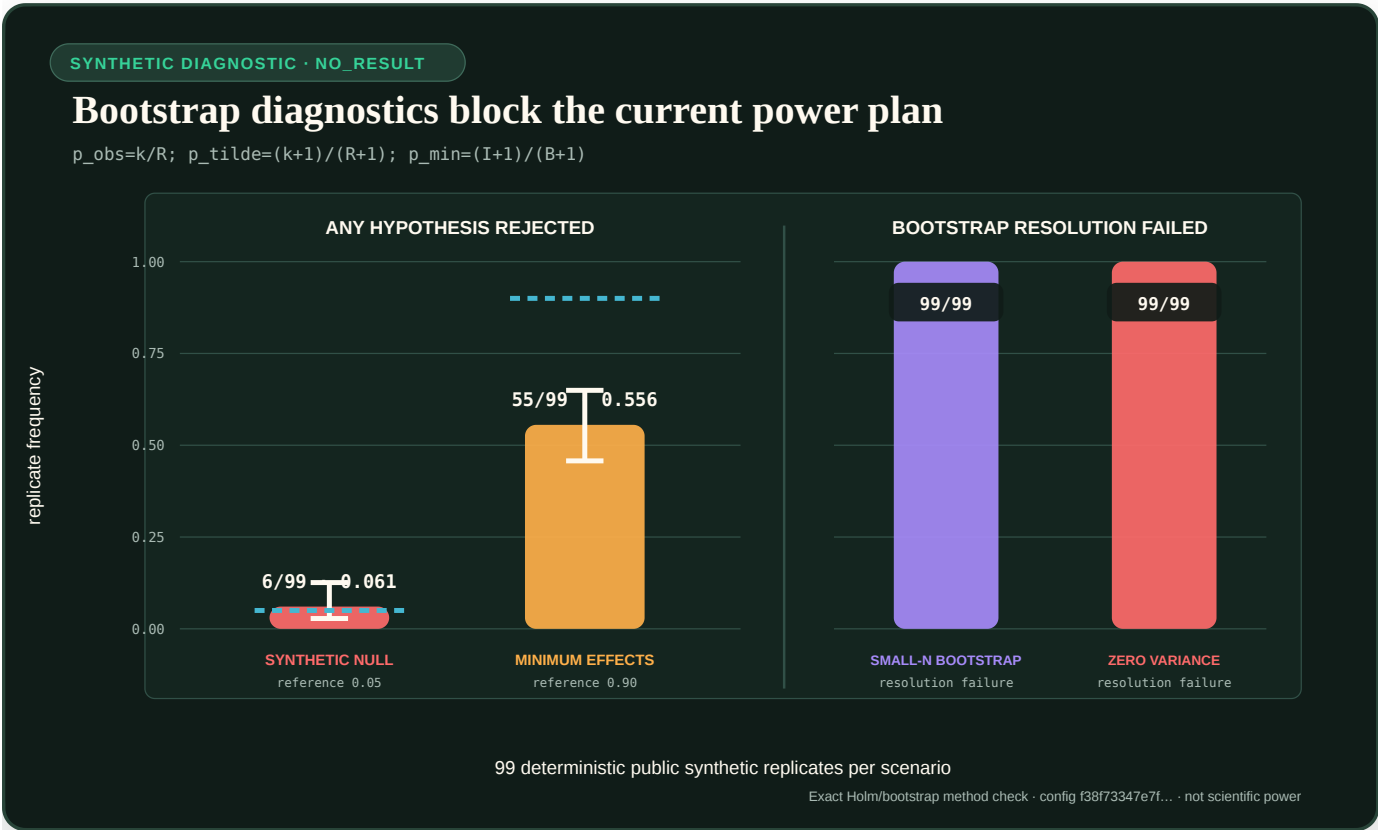


FIGURE 66 — Observed any-rejection frequencies with Wilson Monte Carlo intervals for the declared synthetic null and alternative, beside bootstrap-resolution failure frequencies for two two-instance hostiles.

The executable contract is `rsd-t02-pilot-transcript-calibration.json`, implemented by `rsd-t02-pilot-transcript-calibration.mjs`. Closure is narrow: reviewed real pilot bytes and role, an analyzer release hash, jointly frozen effects, target, resampling key and failure penalties, and an accepted planned-count calibration remain open.

The causal-memory coordinate is not primary-scorable in this first population design because the current family bank contains no valid memory-negative lineage. It can become primary only after both values have prospective, lineage-diverse coverage. Outer family templates and truth remain encrypted or evaluator-custodied until the model, calibration, thresholds, analyzer, resource caps and power plan are frozen. A bare public hash is a commitment, not secrecy for a small family search space.

The closed prospective machine form is `rsd-t02-population-design.json`, validated by `rsd-t02-population-contract.mjs`.

Exact public fixed-instance construction

The public family registry currently contains the five named equation families only. Four generator-conformance coordinates are crossed with every family, producing 20 metadata artifacts. This count tests deterministic construction; it is not a powered sample size.

For family f , draw index j , parameter key k , and HMAC attempt a , let

$$w_{f,j,k,a} = \text{U64BE} [\text{HMAC}_{K_{\text{public}}} (\text{canon}(d, H\mathcal{F}, Hf, f, vf, j, k, a))_{0:8}].$$

$H\mathcal{F}$ hashes the ordered scientific family definitions only; coverage policy, custody state, packet metadata and generator authority are excluded. Hf binds the selected family, including its declared equation- template digest. K_{public} is committed replay material rather than a secret. The sole sampled parameter is an integer time constant

$$\tau_{\mu s} \sim \text{DiscreteUniform}\{500000, \dots, 2000000\}.$$

With $K = 1,500,001$ possible integers and

$$L = 2^{64} - (2^{64} \bmod K),$$

the generator rejects $w \geq L$ before applying modulo reduction, then uses

$$\tau_{\mu s} = 500000 + (w \bmod K), \quad \tau_s = \frac{\tau_{\mu s}}{10^6} \text{ s.}$$

The complete parameter document stores exact numerator, denominator and unit objects. Its digest, the fixed nuisance-interface digest and complete certificate-set digest, including the equation-template digest, enter the canonical system identity. The full registry, population-design bytes and model-source bytes remain separate provenance bindings. The draw index is recorded in the receipt but not in that identity. Every generated packet lists the same parameter digest and τ_s on all 26 unique episodes. A distinct episode-protocol digest binds every schedule, the horizon, integration step, output rate, input bounds, units and interpreter semantics into the packet ID without contaminating the system ID. The generator packet itself contains no trajectories or policy response. A separate fixed-instance conformance runner now materializes its 26 trajectories and causal view. An additive overlay binds a content-addressed 26-projection abstention bundle, executes it in a fresh restricted child, semantically replays the nine responses, and durably resumes an owner-bound fixed-instance ledger. A compact population runner traverses all 20 unique public instances and receipts 520 episodes, 799,240 transcript rows and 180 arm invocations. The integrated execution release composes both layers: one identity-keyed durable instance directory per system and one bounded outer record only after its nine-arm summary is

complete and current. Restart tests cover the post-instance/pre-outer crash window and full-panel reopen without duplicating a scientific unit. The outer records contain no endpoints or causal payloads, and the release still does not execute the trained candidate or null policies.

The coverage function counts distinct structural lineages per property value, collapsing the full-panel-equivalent l1-FFL and affine high-pass siblings into one lineage. The frozen minimum is two:

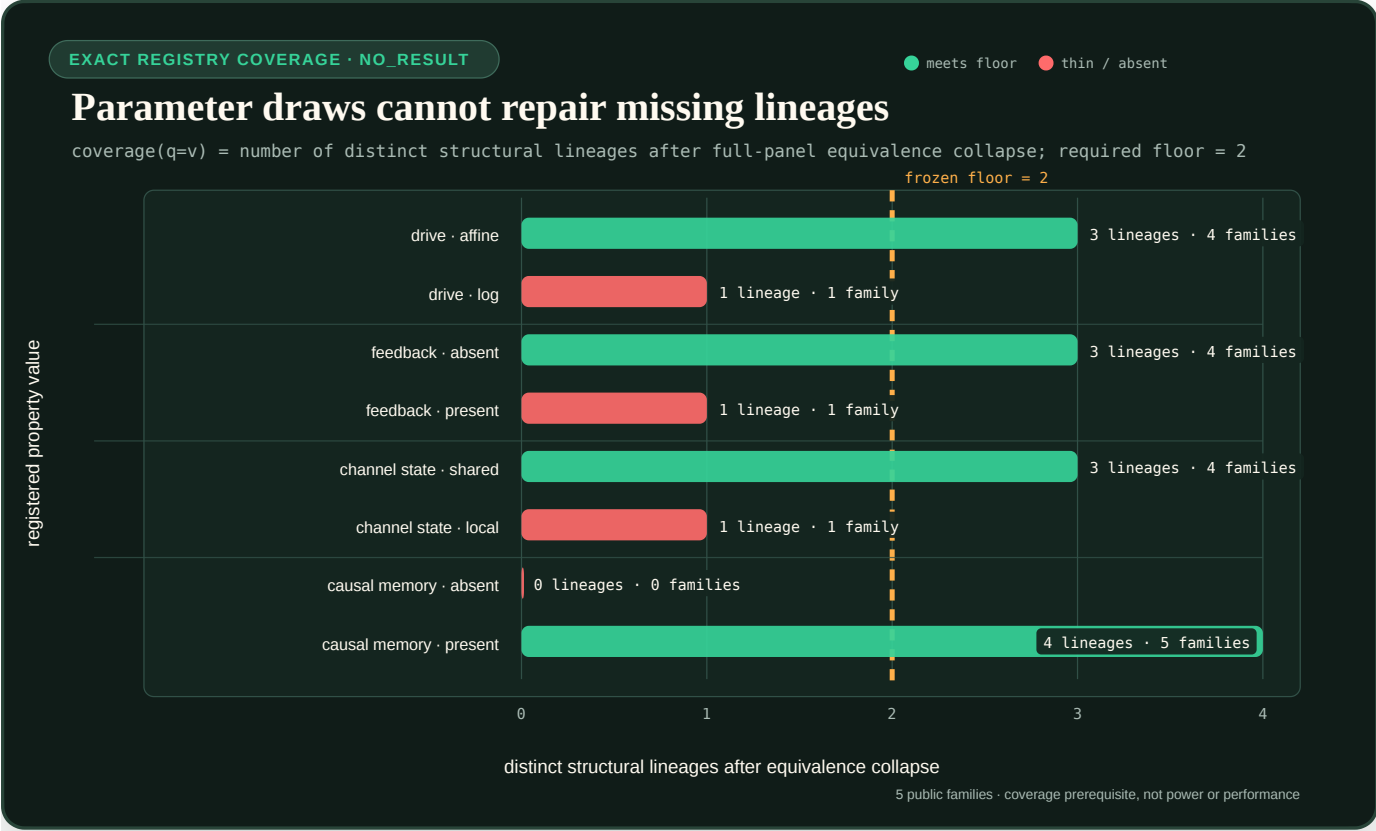


FIGURE 67 — The current public family registry has four property values below the two-lineage floor.

log-fold, feedback-present and channel-local-present each have one lineage; memory-negative has zero. More draws from the current equations cannot close those structural gaps. The closed machine artifacts are the family registry, instance plan, and generator.

Generic-null maturity is a state machine

The executable B-STATE-SPACE and B-RECURRENT policies in the older 35-projection construction bank remain fixed level-one references. Separate deterministic trainable implementations now provide a causal latent state-space prototype and a compact GRU-style prototype. They consume the fixed-instance causal view only through a post-validation adapter and occupy level two; they do not replace the older policy responses. The maturation sequence is:

1. fixed conformance reference;
2. trainable public prototype;
3. fit-frozen development estimator;
4. calibrated development comparator;
5. confirmation-frozen mature null; and
6. confirmation-evaluated run state.

Only level 5 satisfies the population gate. Both current trainable prototypes are at level 2. They emit normalized value posteriors for all three primary coordinates, identifiability probabilities, one coherent joint posterior, support status, a deterministic decide-or-abstain action, reason codes and a typed work ledger, but their probabilities are uncalibrated and their models, resource caps and source/runtime identity are not comparison-frozen. The machine implementations are the `prototype module`, `post-validation adapter`, and `maturation design`.

For instance i and property q , a prospective common objective family is

$$J(\theta) = \sum_{i,q} a_q [\text{BCE}(I_{iq}, s_{iq}) + I_{iq} \text{CE}(\pi_{iq}, p_{iq})] - \lambda_E \sum_i \log \left(\sum_{v \in \mathcal{E}_i} q_i(v) \right) + \lambda_P L_{\text{pred}}(\theta) + \lambda_R R(\theta).$$

θ is the trainable parameter vector, i indexes system instances, and q indexes active property coordinates. Let $\mathcal{V}$ be the finite active joint property-vector domain. For each i , the joint posterior $q_i : \mathcal{V} \rightarrow [0, 1]$ is normalized by $\sum_{v \in \mathcal{V}} q_i(v) = 1$, while the evaluator supplies a nonempty set $\emptyset \neq \mathcal{E}_i \subseteq \mathcal{V}$ of compatible vectors. Zero posterior mass on all of $\mathcal{E}_i$ gives an infinite negative log-mass penalty.

$I_{iq} \in \{0, 1\}$ is certified identifiability and $s_{iq} \in [0, 1]$ its predicted probability. If $I_{iq} = 1$, π_{iq} is the unique property truth and p_{iq} a normalized posterior over the registered values of coordinate q . If $I_{iq} = 0$, π_{iq} is not required and the masked expression is defined as $I_{iq} \text{CE}(\pi_{iq}, p_{iq}) = 0$. $L_{\text{pred}}(\theta)$ and $R(\theta)$ are respectively the auxiliary causal-prediction loss and frozen regularizer. They, BCE, CE, and the log-mass term are normalized to dimensionless quantities.

The fit-only training weights satisfy $a_q \geq 0$ and $\sum_q a_q = 1$; they are not the common endpoint-aggregation weights $w_q \geq 0$, $\sum_q w_q = 1$, which freeze before development evaluation. The coefficients $\lambda_E > 0$ and $\lambda_P, \lambda_R \geq 0$ are dimensionless and selected inside fit. Predictive horizon, optimizer and stopping rule remain fit-only choices, while the deterministic trial tie-break freezes before any trial outcome exists. Until the causal-memory activation condition passes, $\mathcal{E}_i$ and q_i span the three primary coordinates only; activating the fourth coordinate requires a new contract version, head and calibration.

The exact six-level status, freeze order, common resource requirements and three separate gate scopes are frozen in the `null-maturation design`. The two trainable-prototype gates and the parameterized isolated durable runner gate are satisfied; seven intrinsic null-maturity gates remain open. Two of ten comparison-release gates—the registry and generator—are also satisfied; the measured-energy meter gate is conditionally applicable only when an energy claim is requested. It is not counted among the 20 mandatory gates for a non-energy comparison. The current non-energy total is therefore five satisfied and 15 open. Affected fitting remains blocked by incomplete lineage coverage, absent sealed outer-family templates, and the absent instance-level fit/calibration/development-evaluation assignment. The exact local runtime closure for the promoted infrastructure gate is recorded in the `parameterized runner release`.

Calibration and abstention

Probability quality is evaluated with logarithmic loss as a proper scoring rule in the sense reviewed by Gneiting and Raftery (2007), bibliography key `gneiting2007scoring`. The separate decision loss below encodes this fixture's abstention costs; it is not silently folded into calibration.

For property coordinate q , let $\mathcal{E}_q$ be the set of values compatible with the registered observation packet and define

$$I_q = 1[|\mathcal{E}_q| = 1].$$

When $I_q = 1$, let π_q denote the unique element of $\mathcal{E}_q$.

An arm returns a probability $\widehat{P}(I_q = 1)$, a posterior over property values, including posterior mass $\widehat{P}(\pi_q)$ on that unique compatible value, and either `decide` or `abstain`. With probabilities clipped at 10^{-12} , the calibration loss in nats is

$$L_{\text{cal},q} = -I_q \ln \widehat{P}(I_q = 1) - (1 - I_q) \ln [1 - \widehat{P}(I_q = 1)] - I_q \ln \widehat{P}(\pi_q).$$

When $I_q = 0$, the final masked term is defined to be exactly zero; π_q and $\widehat{P}(\pi_q)$ are not evaluated.

The decision loss is dimensionless:

$$L_{\text{dec},q} = \begin{cases} 0, & I_q = 1 \text{ and the decision is correct,} \\ 0.25, & I_q = 1 \text{ and the arm abstains,} \\ 0, & I_q = 0 \text{ and the arm abstains,} \\ 1, & \text{wrong decision or declaration on a non-singleton set.} \end{cases}$$

Coverage, selective risk and reliability remain separate. Sensitivity analyses later vary the identifiable-case abstention cost to 0.1 and 0.5; they do not replace the primary loss.

The fast boundary layer needs the right norm

The source-qualified stratum uses the nondimensional input-degradation model

$$\tau_s \dot{x} = \bar{u} - x, \qquad \epsilon \tau_s \dot{y} = x - \bar{u} y,$$

with $\tau_s = 1$ s, $\bar{u}_0 = 1$, $\bar{u}_* = 2$, $x(0) = \bar{u}_0$ and $y(0) = 1$. The scaled member uses $\bar{u}_p = p \bar{u}$, $x_p(0) = p \bar{u}_0$ and $y_p(0) = 1$.

The primary finite-grid truth is

$$D_{\infty,\epsilon,p} = \sup_{0 \leq t \leq 8 \tau_s} |y_{\epsilon,p}(t) - y_{\epsilon,1}(t)|.$$

For this registered construction, the associated fast initial-value systems give

$$M_p = \left| 1 - \frac{\bar{u}_0}{\bar{u}_*} \right| |p - 1| p^{p/(1-p)} > 0, \qquad p \neq 1,$$

and a source-shaped finite- ϵ bound has the form

$$D_{\infty,\epsilon,p} \geq M_p - \epsilon \widetilde{N}_p.$$

The protected construction does not infer this asymptotic statement from a finite sweep. It checks the declared generator and bound.

An RMS score measures something else:

$$D_{2,\epsilon,p} = \sqrt{\frac{1}{8 \tau_s} \int_0^{8 \tau_s} |y_{\epsilon,p}(t) - y_{\epsilon,1}(t)|^2 dt}.$$

The exact physical-time metric counterexample

$$e_{\epsilon}(t) = e^{-t/(\epsilon\tau_s)}$$

has

$$\|e_{\epsilon}\|_{\infty} = 1, \quad \text{RMS}(e_{\epsilon}) = \sqrt{\frac{\epsilon\tau_s}{2T} \left(1 - e^{-2T/(\epsilon\tau_s)}\right)} \longrightarrow 0.$$

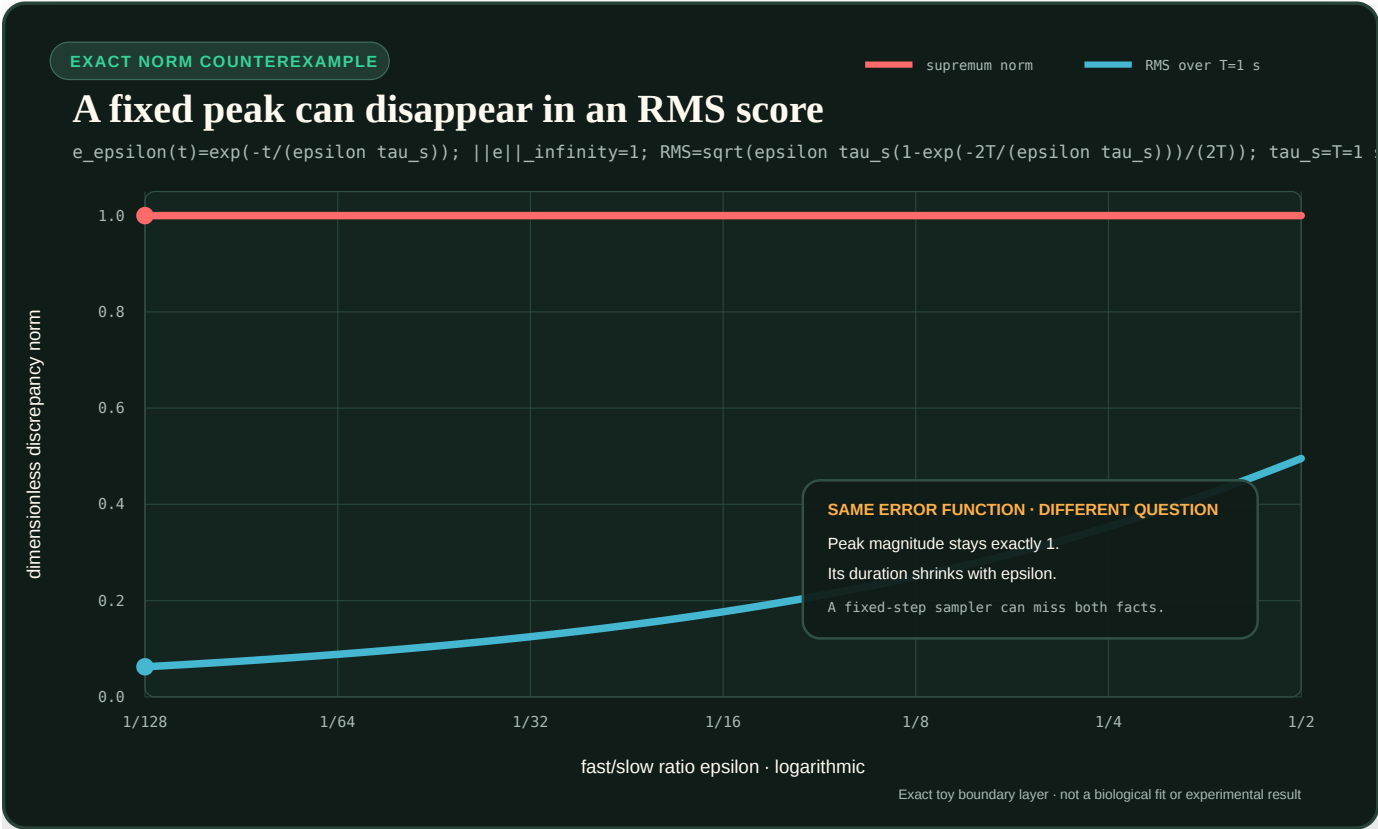


FIGURE 68 — A boundary-layer peak remains fixed in the supremum norm while its RMS value falls with epsilon.

The figure is an exact toy norm comparison, not a biological fit or a run of the source-shaped generator.

Fast- and slow-time sampling

The protected grid freezes

$$\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\},$$

$$p \in \{0.5, 2, 4, 8, 20\}.$$

The critical initial-layer time is

$$t_{\epsilon,p} = \epsilon\tau_s \frac{\ln p}{(p-1)u_*}.$$

Every cell samples the union

$$\left\{\epsilon\tau_s\frac{j}{128} : j = 0, \dots, 1024\right\} \cup \left\{\tau_s\frac{k}{64} : k = 0, \dots, 512\right\}.$$

The first grid resolves the shrinking boundary layer; the second retains the eight-second slow response. Deduplication leaves at most 1,537 paired rows per cell.

The floor stratum crosses three equation-defined models, seven epsilon values and five scale factors. Besides the singular system above, both controls share

$$\tau_s\dot{x} = \bar{u} - x.$$

The exact-equivariance control reports

$$y_0 = \frac{\bar{u}}{x} - 1.$$

Because the scaled member has $\bar{u}_p = p\bar{u}$ and $x_p = px$, its discrepancy is identically zero. The regular-perturbation control reports

$$y_\epsilon = \frac{\bar{u}}{x} - 1 + \epsilon(\bar{u} - \bar{u}_0).$$

Its scaled-versus-unscaled discrepancy is $\epsilon|p - 1||\bar{u} - \bar{u}_0|$ and therefore tends to zero. The three registered models are consequently:

1. the source-shaped singular construction;
2. an exact-equivariance zero-floor control; and
3. a regular-perturbation control whose discrepancy tends to zero.

That is 105 public-development cells and at most 161,385 paired rows. RMS is diagnostic only and cannot establish or refute the supremum floor.

Information firewall

Every actionable arm receives only causal inputs, reported outputs, masks, opaque intervention commands, timestamps and units. Before the arm response is frozen it does not receive:

1. recipe or equation identity;
2. semantic state names or the hidden handle permutation;
3. parameters or evaluator properties;
4. future samples or future-derived normalization;
5. the equivalence or separation certificate; or
6. continuous evaluator truth.

The actionable registry has nine roles:

1. `A-RAW` ;
2. `B-STATIC-DIV` ;
3. `B-STREAM` ;
4. `B-LOG-RATIO` ;
5. `B-DIFFERENCE` ;
6. `B-STATE-SPACE` ;

7. `B-RECURRENT` ;
8. `C-MECHANISM-BANK` ; and
9. `C-DUAL` .

`0-GRAPH` is evaluator-only and excluded from parity, tuning, promotion and resource rankings. Public source code does not create confirmation secrecy; a claim-eligible run later needs separately committed sealed seed mapping and custody.

Typed acquisition and computation cost

Do not collapse intervention access and execution into one score. Every arm retains at least:

1. episodes and sample rows;
2. serialized observation bytes and input commands;
3. internal resets, freezes and output clamps;
4. channel switches and state writes;
5. scalar operations and transcendental evaluations;
6. retained-state and parameter bytes;
7. tuning trials and wall seconds; and
8. later measured joules, when a calibrated physical protocol exists.

The foundation suggests future caps of one CPU thread, binary64 arithmetic, 16 retained scalars, 512 trainable scalars and 32 tuning trials. They are not active parity claims because actionable algorithms and a complete arm-level parity/resource ledger are absent.

Support and fail-closed cases

Each scientific case retains the six independent support axes already frozen for RSD-T01:

1. input domain;
2. transformation family;
3. instrument range and temporal resolution;
4. initialization;
5. causal observation; and
6. evaluation window.

Valid scientific hostiles include additive offset, near-zero input, clipping, hidden reset, future-aware normalization, channel-state contamination and boundary-layer censoring. Parser, checksum, order or unit failures remain malformed sentinels outside the scientific denominator.

Missing, duplicate, rejected or mixed-initialization transcripts force system abstention and remain visible. A fixed slow-time sampler that misses the protected fast layer is a scientific failure, not a missing-data deletion.

Current authority and kill rules

The machine registry states:

```
{
  "authority": "contract-foundation-only",
  "partition": "public-development",
  "information_cut_status": "registered-projection-no-secret-custody",
```



```
"comparison_authority": false,  
"result_authority": "NO_RESULT"  
}
```

The registry is the foundation authority, not the execution result. A separate deterministic bounded public-development runtime now consumes the T02-MECH registry to generate all O0/O1 construction episodes, enforce the policy firewall and response commitment, reconstruct evaluator truth, retain typed acquisition/construction/inference ledgers, and validate append-only resume. Its additive Stage 2 commits all nine fixed whole-system policy-conformance responses before evaluator access, with zero inactive placeholders. Every event and analysis remains `NO_RESULT`; trained or calibrated estimators, mature nulls, comparisons, claim eligibility, O2, T02-FLOOR execution, confirmation, workstation measurement and energy conclusions are absent.

The future T02 comparison is killed if any of the following occurs:

1. step fit, recipe name or graph access supplies the primary answer;
2. a declared separating intervention lacks a pairwise certificate;
3. an arm is rewarded for guessing inside an observational equivalence class;
4. privileged access is unequal or missing from the cost vector;
5. RMS or a fixed slow grid substitutes for the registered supremum endpoint;
6. a finite epsilon sweep is presented as proof of an asymptotic theorem; or
7. a mechanism-specific bank cannot beat the generic state-space/recurrent null under the same projection and budget.

The current foundation and construction runtime can test equations, exact step matching, operational equivalence, registry closure, schedule semantics, abstention aggregation, replay integrity and temporal-grid coverage. They cannot support architecture superiority, natural-mechanism attribution, workstation readiness or energy efficiency.

Horizon-qualified learning outcomes

Scope

This note makes acquisition, retention, transfer, fluency, calibration, motivation, and effort separate outputs. It supports the learning-science audit and the experiment contracts for Candidates 004 and 019. Its evidence boundaries are C-627–C-658.

Immediate and retained change

For learner/model i , skill item j , and policy m , let $Y_{ijm}(t)$ be a score in a declared score unit. Acquisition and retained change are

$$A_{ijm} = Y_{ijm}(t_{\text{post}}) - Y_{ijm}(t_{\text{pre}}),$$

$$R_{ijm}(\Delta) = Y_{ijm}(t_{\text{post}} + \Delta) - Y_{ijm}(t_{\text{pre}}),$$

where retention horizon Δ is seconds, task events, or another declared clock. A policy can win on A and lose on $R(\Delta)$.

Transfer is indexed

Let $d \in \{0, 1, 2, 3\}$ identify preregistered strata: trained form, near variant, changed representation/context, and novel causal composition. Against baseline m_0 ,

$$T_m(d) = \mathbb{E}[Y^{\text{transfer}} \mid m, d] - \mathbb{E}[Y^{\text{transfer}} \mid m_0, d].$$

$T_m(d)$ uses score units and is reported for every d . A single mean called “generalization” cannot show which cue, mapping, or composition transferred.

For elapsed time τ seconds, c correct responses, and n attempts,

$$F = \frac{c}{\tau} \quad [\text{correct tasks s}^{-1}], \quad e = 1 - \frac{c}{n} \quad [1].$$

Fluency requires the (F, e) frontier, not speed alone. Calibration retains item-level confidence before feedback and reports a proper score such as

$$\text{BS} = \frac{1}{N} \sum_{k=1}^N (p_k - y_k)^2,$$

which is dimensionless for probability p_k and binary outcome y_k .

Skill-local scheduling state

For skill j , scheduler state is

$$\mathcal{L}_{j,t} = (\hat{a}, \hat{r} \Delta_{1:\Delta_Q}, \hat{T}_{0:3}, \hat{F}, \hat{c}, \sigma, \chi, h, v),$$

where $\hat{a}$ is acquisition, $\hat{r}$ retention by target horizon, $\hat{T}$ transfer by stratum, $\hat{F}$ fluency, $\hat{c}$ calibration, σ support/scaffold state, χ confusability/context, h intervention history, and v state/model version. These are estimates with uncertainty, not hidden truth labels.

The next event may retrieve, restudy, vary, compare, explain, fade support, or stop. Difficulty is admissible only when processing succeeds often enough to produce information; failure without interpretable feedback is not useful effort by definition.

Complete cost

Keep raw cost components

$$\mathbf{c}_m = (\tau, N_e, N_r, N_f, N_h, B, E, H_T, H_L),$$

where τ is learner time in seconds; N_e, N_r, N_f, N_h are exposure, retrieval, feedback, and hint counts; B is stored bytes; E is joules; and H_T, H_L are teacher and learner effort in person-seconds. Variable-time mastery, adaptive scheduling, or interactive teaching cannot claim efficiency while receiving uncharged extra attempts or attention.

Falsification boundary

The scheduler loses if tuned spaced repetition, fixed expanding intervals, ordinary knowledge tracing, blocked/random/interleaved schedules, worked examples with fixed fading, hard-example mining, or fixed curriculum match the delayed retention–transfer–cost frontier. A teaching channel loses if a versioned artifact plus tests preserves equal capability across learner/model turnover at lower combined effort.

Material commitment and realized service

Scope

This note separates information, policy, commitment, physical state, and realized service for resource-bearing AI systems. It instantiates the supply-chain/operations audit and the soil/crop multi-resource audit without treating inventory, routing, buffers, or biological co-limitation as new AI principles. Its evidence boundaries are C-627 and C-659–C-678.

Conservation before optimization

For location or module i over interval t , physical inventory obeys

$$I_{i,t+1} = I_{i,t} + \text{Rec}_{i,t} + T_{i,t}^{\text{in}} - S_{i,t} - T_{i,t}^{\text{out}} - X_{i,t},$$

where every term is a count, byte quantity, mass, or another single declared unit: on-hand I , receipts Rec , inbound/outbound transfer T , issued work S , and expiry/damage X . The explicit receipt symbol avoids reusing R for both receipts and reservations. Forecasts, requests, allocations, and record corrections do not appear as physical flow unless an observed transition links them to it. Filtration, internal recovery, reintroduction, final egress, and storage change likewise remain separate ledger terms (C-1491). A third-body interface inventory also requires generation, transport, transformation, reincorporation, and escape closure before retain/remove optimization (C-1502).

With accepted but unfinished work,

$$B_{i,t+1} = [B_{i,t} + D_{i,t}^{\text{acc}} - S_{i,t}]^+,$$

where backlog B , accepted demand D^{acc} , and completion S use the same work unit. Lost requests require a separate state because they do not remain in B and may disappear from later demand records.

Resource form and deliverability

For location i , resource type k , and interval t , retain a compartment state

$$\mathbf{x}_{ik,t} = (s^{\text{bound}}, s^{\text{lab}}, s^{\text{sol}}, s^{\text{pipe}}, Q^{\text{internal}})_{ik,t}^{\text{T}}.$$

s^{bound} is bound or slow-release stock, s^{lab} is a labile or exchangeable pool excluding the solution pool, s^{sol} is material in solution or its transported analogue, s^{pipe} is in-transit material inside the declared boundary, and Q^{internal} is an already delivered, remobilizable internal pool. The operational compartments are mutually exclusive and use one declared mass, count, or task-native resource unit. Total stock and solution concentration are derived without adding unlike or overlapping states:

Aggregate stock or inflow therefore cannot establish local arrival, uptake, useful consumption, or deficit at the receiver's support (C-1488).

$$S_{ik,t}^{\text{tot}} = \mathbf{1}^T \mathbf{x}_{ik,t}, \quad C_{ik,t}^{\text{sol}} = \frac{S_{ik,t}^{\text{sol}}}{V_{i,t}^{\text{sol}}}.$$

V^{sol} is the declared carrier or solution volume. Reachable stock $S_{ik}^{\text{reachable}}(t; \Phi)$ and usable stock $U_{ik}^{\text{usable}}(t; \Phi)$ are deadline-qualified functions of this state and the transport/release model, not extra compartments that can be double counted. $\mathbf{1}$ is the all-ones vector and ^{T} denotes transpose.

The compartment balance is

$$\mathbf{x}_{ik,t+1} = \mathbf{T}_{ik,t}(\theta_{i,t}, pH_{i,t}, \Theta_{i,t}, \beta_{i,t}) \mathbf{x}_{ik,t} + \mathbf{a}_{ik,t} - \mathbf{e}_{ik,t} - \mathbf{\ell}_{ik,t},$$

where $\mathbf{T}_{ik,t}$ is a dimensionless transition operator for release, sorption/desorption, dissolution, mineralization, and other form changes; θ is volumetric water content, pH is the declared logarithmic activity measure, Θ is temperature on a declared scale, and β is the declared biological or process state. $\mathbf{a}$ is externally added material, $\mathbf{e}$ is observed service consumption or withdrawal across the boundary, and $\mathbf{\ell}$ is boundary loss. The three vector flow terms are interval-integrated amounts in the same resource unit as $\mathbf{x}$. A resource unit occupies one compartment; transition columns conserve it unless a declared transformation or loss is represented explicitly. Uptake into Q^{internal} is an internal transition; it must not also be subtracted as a boundary flow. Internal form changes can conserve an element while changing its deliverability. A lubricant, buffer, cache, or reserve used as an interface mediator remains a typed, regime-qualified resource rather than an automatically beneficial label (C-1499).

Observed uptake over Δt is bounded separately from stock:

$$U_{ik,t} \leq \min \left\{ (J_{ik,t}^{\text{adv}} + J_{ik,t}^{\text{diff}}) \Delta t, V_{ik,t}^{\text{receiver}} \Delta t, D_{ik,t}^{\text{phys}} \Delta t, S_{ik}^{\text{reachable}}(t; \Phi) \right\},$$

where U is observed uptake or receiver acceptance, J^{adv} and J^{diff} are non-overlapping advective and diffusive delivery rates, V^{receiver} is receiver uptake or processing capacity, and D^{phys} is physiological or task demand. Δt is the interval duration in seconds. All four right-hand quantities use the same resource amount: the first three are rates in resource units per second integrated over Δt , and the fourth is already an amount reachable inside window Φ . This equation is a service-accounting bound, not a universal soil, crop, or receiver model.

State contract

For service commitment c , retain

$$\mathcal{K}_c = (F_v, D, A, O, K^{\text{res}}, \mathbf{I}, P, C, Q, L, \Phi, \mathcal{F}, Y, M),$$

where F_v is forecast plus vintage, D observed request, A admission decision, O order/release, K^{res} reservation/frozen commitment, $\mathbf{I}$ typed on-hand and pipeline inventory by form, age, and condition, P policy version, C qualified capacity and common-cause groups, Q route/queue/qualification state, L lead-time distribution, Φ service stage or deadline window, $\mathcal{F}$ jointly feasible service set, Y delivered outcome, and M service-measurement definition. Every field carries event, availability, and decision times when they differ.

The tuple is deliberately not a scalar "available resource." A valid record does not create physical stock; a route plan does not complete transport; a shipment does not prove correct or timely service. The renal accounting boundary supplies an independent instance of why intermediate throughput is not final service (C-1491).

Jointly executable service

Let service class j require $\nu_{jk} \geq 0$ units of resource k per service unit. At location i , the currently feasible service set is

$$\mathcal{F}_i(t) = \left\{ \mathbf{q} \geq 0 : \sum_j \nu_{jk} q_j \leq U_{ik}^{\text{usable}}(t) \text{ for every required } k, \quad q_j = 0 \text{ outside } \Phi_j \right\}.$$

q_j is service units executable within window Φ_j , and U_{ik}^{usable} is the resource amount deliverable within that same window. This bound detects stranded single-resource reservations. It is not a claim that all production functions are fixed-proportion: an experiment must declare partial substitution, internal storage, transformation, and toxicity when they are possible.

For a two-resource factorial experiment, keep the interaction contrast in the same units as the outcome:

$$\Delta_{k\ell} = Y_{11} + Y_{00} - Y_{10} - Y_{01}.$$

A positive $\Delta_{k\ell}$ is super-additivity in that experimental context. It does not by itself distinguish simultaneous, independent, or serial co-limitation, nor identify a transport or allocation mechanism.

Service and recovery vectors

Report service as a vector

$$\mathbf{s} = (f_{\text{unit}}, f_{\text{order}}, f_{\text{OTIF}}, f_{\text{bundle}}, W_{50}, W_{95}, B, L_s, Z, E, C_m, H, W_m),$$

where fill/service fractions f are dimensionless, delays W_{50}, W_{95} are seconds, f_{bundle} is the fraction of commitments whose complete typed resource bundle was executable in its window, backlog B and lost demand L_s use work units, Z is stranded reserved resource in a declared resource unit, lifecycle energy E is joules, monetary cost C_m is currency, human work H is person-seconds, and material waste W_m is mass or count. Resource-specific Z and W_m coordinates remain separate when units differ. Coordinates remain visible even when a declared policy uses weights.

Recovery ends at horizon T_R only after backlog, reserve, and next-event service are measured:

$$\mathcal{R}(T_R) = (t_{\text{nominal}}, t_{\text{backlog}=0}, t_{\text{reserve}}, s_{\text{second}}, E, C_m, H, W_m).$$

Nominal throughput is one timestamp, not the whole recovery state.

For multi-resource systems, the second-event vector also reports residual stock by form, missed service windows, and losses. Late abundance does not retroactively repair a commitment that failed inside its declared window.

Falsification boundary

The held contract loses if typed event sourcing plus inventory reconciliation, queueing/flow models, base-stock or multi-echelon control, stochastic/robust optimization, receding-horizon planning with frozen commitments, and explicit service metrics match its decisions and cost. For soil/crop-derived claims the null stack also includes factorial response surfaces, mechanistic transport-plus-uptake models, balanced-nutrient models, and calibrated crop system simulators. Pooling, supplier count, information sharing, just-in-time operation, co-limitation, structural proliferation, and closed loops receive no default efficiency or resilience credit.

Measurement contracts, uncertainty, and invalidation

Scope

This note defines the minimum mathematics needed to turn an indication into a decision-bearing result. It connects the metrology audit, Candidate 014, and the energy model. It does not replace a domain-specific measurement procedure.

Result record

For one measured quantity, retain

$$\mathcal{M} = (Y, \hat{y}, u_c, U, k, \mathcal{Q}, \mathcal{P}, \mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{D}, \mathcal{V}, t_0, t_1),$$

where:

SYMBOL	MEANING	UNIT OR DECLARATION
Y	measurand: quantity intended to be measured	physical unit or named reference scale
$\hat{y}$	value estimate attributed to Y	same unit as Y
u_c	combined standard uncertainty	same unit as Y
U	expanded uncertainty	same unit as Y
k	coverage factor such that $U = k u_c$	dimensionless
$\mathcal{Q}$	object, population, component, state, location, and conditions	typed metadata
$\mathcal{P}$	measurement procedure and model	versioned identifier
$\mathcal{C}$	calibration chain, references, corrections, and validity scope	dependency record
$\mathcal{E}$	environment and influence quantities	values with units
$\mathcal{S}$	sampling, selection, missingness, and association contract	probability/model record
$\mathcal{D}$	decision rule, tolerance, guard band, and risk allocation	versioned rule
$\mathcal{V}$	data, software, instrument, certificate, and transformation provenance	dependency graph
t_0, t_1	start and end of the supported interval	seconds on a named time basis

The tuple is incomplete when a field that can alter the downstream decision is missing. A confidence score alone cannot represent these distinct dependencies (C-519–C-524).

Measurement model and dimensional validity

Let

$$Y = f(X_1, \dots, X_n),$$

where input quantity X_i has unit $[X_i]$ and Y has unit $[Y]$. Every additive term produced by f must have unit $[Y]$. Software, coefficients, numerical precision, and preprocessing that implement f are part of $\mathcal{P}$. A coefficient of friction is therefore not portable without its interface pair, load/motion state, environment, procedure, support, and history (C-1497).

For estimates x_i , first-order covariance propagation gives

$$u_c^2(\hat{y}) = \sum_{i=1}^n \sum_{j=1}^n c_i c_j u(x_i, x_j), \quad c_i = \left. \frac{\partial f}{\partial X_i} \right|_{x_1, \dots, x_n}.$$

c_i has unit $[Y]/[X_i]$ and covariance $u(x_i, x_j)$ has unit $[X_i][X_j]$, so every summand has unit $[Y]^2$. Removing off-diagonal terms asserts independence; it is not a harmless simplification when calibration, clock, environment, preprocessing, prior, or training data are shared (C-525, C-533). For nonlinear or discontinuous models, propagate sampled input distributions and report empirical interval coverage.

Error, correction, and coverage

Given reference value y_{ref} ,

$$e = \hat{y} - y_{\text{ref}}, \quad c = -\hat{e}, \quad \hat{y}_{\text{corr}} = \hat{y} + c.$$

Error e , correction c , and corrected estimate $\hat{y}_{\text{corr}}$ have unit $[Y]$. The estimated correction remains uncertain. Expanded uncertainty is

$$U = k u_c,$$

where k is dimensionless. A stated k does not by itself establish a coverage probability; the distributional method and achieved coverage must be reported (C-524, C-526).

Repeatability and reproducibility

For replicate result y_{lors} from location l , operator o , run r , and system s , a variance-component null is

$$y_{lors} = \mu + Ll + Oo + Rr(lo) + Ss + \epsilon_{lors}.$$

μ and every random effect have unit $[Y]$; their variances have unit $[Y]^2$. The design declares which factors are fixed or random and which are confounded. A short run with one system estimates repeatability, not broad reproducibility (C-527, C-528).

Decision rules and guard bands

Let specification require $Y \leq T$, where T has unit $[Y]$. A simple guarded acceptance rule is

$$\text{accept if } \hat{y} + gU \leq T,$$

where guard multiplier $g \geq 0$ is dimensionless. Increasing g generally reduces false acceptance while increasing false rejection. Choose g from a declared loss or risk allocation; do not hide that choice inside “confidence.” The rule version, tolerance version, uncertainty method, and cost owner are dependencies of the decision (C-531).

Drift and recalibration

Represent indication drift relative to the last accepted calibration as

$$d(t) = y_{\text{check}}(t) - y_{\text{ref}}(t),$$

where d , check-standard result y_{check} , and reference value y_{ref} share unit $[Y]$. The review policy is a function

$$\pi_{\text{cal}} = \pi(h_t, r_t, u_t, c_t, e_t),$$

where h_t is calibration/check history, r_t is decision risk, u_t is current uncertainty, c_t is calibration cost, and e_t is equipment and environment state. Inputs keep their native units; π returns a categorical action such as continue, check, restrict, recalibrate, or quarantine. A calendar interval alone is not evidence of stability (C-532).

Dependency invalidation

Let a directed acyclic graph $G = (V, E)$ contain calibration, raw-data, software, model, transformation, result, and decision versions. Edge $(a, b) \in E$ means that node b depends on node a . If dependency a changes or fails review, its invalidation cone is

$$\mathcal{I}(a) = \{v \in V : a \rightsquigarrow v\},$$

where $a \rightsquigarrow v$ denotes a directed path and $\mathcal{I}(a)$ is a set of version identifiers. Each affected node is re-evaluated, superseded, restricted, or withdrawn; provenance alone does not decide which action is correct (C-534, C-535).

Energy-measurement instantiation

For sampled power P_m at time t_m with interval Δt_m ,

$$\hat{E} = \sum_{m=1}^M P_m \Delta t_m,$$

where P_m is W, Δt_m is s, and $\hat{E}$ is J. The model must name the electrical boundary, voltage/current/phase calibration, bandwidth, anti-aliasing, clock alignment, integration rule, missing-sample policy, warm-up, idle allocation, retries, useful outputs, and facility attribution. Uncertainty propagates through the integration model with shared meter, coefficient, and clock covariance retained.

Compare candidate C and baseline B only after both satisfy the same quality, risk, latency, and workload envelope. For paired run r ,

$$\Delta e_r = \frac{EC,r}{N_{q,C,r}} - \frac{EB,r}{N_{q,B,r}},$$

where energies are J , N_q is a qualified-event count, and Δe_r is J/qualified event. Report absolute values, paired effect, coverage, and all failed or rejected runs. A software counter or narrow device boundary may rank systems differently from calibrated end-to-end measurement; the direction is an empirical question (C-536).

Falsification conditions

Reject the cross-layer composition as a distinct systems contribution if:

1. a complete conventional metrology, statistics, and content-addressed provenance stack matches its coverage and stale-decision frontier;
2. it improves results only by receiving extra sensors, calibration runs, compute, storage, or analyst time;
3. dependencies are recorded but do not trigger correct downstream action;
4. uncertainty intervals become narrower while empirical coverage worsens;
5. shared dependencies are counted as independent evidence; or
6. the metadata and review burden costs more than the errors or work it avoids.

Budgeted memory lifecycle

Purpose

Turn selective replay, consolidation, and forgetting into a costed decision problem. This note specifies a testable controller; it does not claim that its score is biologically implemented or optimal.

Memory items and actions

At maintenance cycle t , the candidate memory set is $\mathcal{M}_t$. Each item $i \in \mathcal{M}_t$ may receive one action

$$a_{i,t} \in \mathcal{A} = \{\text{defer}, \text{replay}, \text{merge}, \text{externalize}, \text{weaken}, \text{delete}\}.$$

The action is versioned. `weaken` and `delete` are invalid unless an external provenance record or explicit retention-policy exception exists.

Expected utility

For item i and action a , estimate

$$G_{i,a,t} = \widehat{\Delta L}_{i,a,t} + \lambda_R \widehat{\Delta R}_{i,a,t} + \lambda_S \widehat{\Delta S}_{i,a,t} - \lambda_D \widehat{D}_{i,a,t},$$

where:

- $\widehat{\Delta L}_{i,a,t}$ is the predicted reduction in future task loss relative to `defer` (dimensionless loss units);
- $\widehat{\Delta R}_{i,a,t}$ is the predicted reduction in a declared risk metric (risk units);
- $\widehat{\Delta S}_{i,a,t}$ is the predicted improvement in schema-consistent transfer (dimensionless score units);
- $\widehat{D}_{i,a,t}$ is predicted destructive-action harm, including rare memory loss and provenance failure (harm units);
- λ_R has units loss/risk;
- λ_S has units loss/score; and
- λ_D has units loss/harm.

The terms cannot be added before the conversion coefficients and evaluation sets are declared. Novelty, uncertainty, reward surprise, familiarity, and interference are features used to estimate these outcomes, not interchangeable units of utility.

Resource model

For each candidate action, predict:

- $E_{i,a,t}$ — energy in joules;
- $B_{i,a,t}$ — data moved in bytes;
- $T_{i,a,t}$ — wall time in seconds; and
- $W_{i,a,t}$ — slow-model optimizer updates (count).

The scheduler selects binary variables $x_{i,a,t} \in \{0, 1\}$:

$$\max_x \sum_{i \in \mathcal{M}_t} \sum_{a \in \mathcal{A}} x_{i,a,t} (G_{i,a,t} - \lambda E_{i,a,t})$$

subject to

$$\sum_a x_{i,a,t} = 1 \quad \forall i,$$

$$\sum_{i,a} x_{i,a,t} E_{i,a,t} \leq E_t^{\max}, \quad \sum_{i,a} x_{i,a,t} B_{i,a,t} \leq B_t^{\max},$$

$$\sum_{i,a} x_{i,a,t} T_{i,a,t} \leq T_t^{\max}, \quad \sum_{i,a} x_{i,a,t} W_{i,a,t} \leq W_t^{\max}.$$

λE has units loss/joule. The four budgets prevent a scheduler from appearing efficient by hiding memory traffic, time, or optimizer work behind a single energy proxy.

This is a multiple-choice multidimensional knapsack problem when predictions are fixed. The first experiment need not solve it exactly; greedy, learned, and standard replay priorities are compared under the same budgets.

The illustrative [single-item price envelope](#) shows how the upper admissible action can change with λE . Its gains, costs, and selected actions are hypothetical and are not a fitted policy.

Reconsolidation gate

Retrieval produces a candidate branch rather than mutating the stored item in place. Let $m_{i,t}$ be a calibrated mismatch statistic and θ_t a gate:

$$z_{i,t} = 1[m_{i,t} > \theta_t].$$

If $z_{i,t} = 0$, the item remains read-only. If $z_{i,t} = 1$, an update branch is allowed but promotion still requires regression and provenance tests. Because the exact biological prediction-error boundary is disputed under C-040, θ_t is an experimental variable, not a biological constant.

Required ablations are:

1. always read-only;
2. always writable on retrieval;
3. fixed mismatch threshold;

- 4. calibrated risk-dependent threshold; and
- 5. explicit source-version change rather than inferred mismatch.

Schema-fit gate

Let $s_{i,t}$ be schema-fit estimated on a held-out structural probe and $q_{i,t}$ be shortcut risk. A candidate may receive a cheaper merge path only when

$$s_{i,t} \geq \tau_s \quad \text{and} \quad q_{i,t} \leq \tau_q.$$

The thresholds τ_s and τ_q are calibrated on validation streams with both compatible items and deceptive near-matches. Training loss alone is not a schema-fit measure.

Measurement protocol

Report both realized and predicted quantities:

QUANTITY	UNIT	BOUNDARY
replay compute energy	J	accelerator or node, declared per run
memory traffic	byte	host–device and device memory reported separately
maintenance time	s	wall clock, including scheduler
optimizer work	update count	slow-model updates only
retention	task metric by item age	fixed evaluation stream
adaptation	events or seconds to threshold	after declared change point
obsolete intrusion	error count/rate	outdated association probes
destructive error	item count/rate	deleted state later shown necessary
provenance recovery	fraction	deleted items reconstructable from source

Calibration error between predicted G and realized outcome is itself a core result. A scheduler that makes good-looking choices only after seeing future tests is invalid.

Falsification

The memory-lifecycle controller fails its first test if a conventional single-priority or reservoir baseline matches its quality–risk–energy frontier, if scheduling overhead consumes the saved replay budget, or if destructive errors exceed the declared safety bound.

Mission-profile-qualified device reliability contract

This note defines the quantitative boundary for Fixture F-008. It operationalizes the durable result of the semiconductor device and circuit reliability audit: accepted service must be evaluated against the **actual mission profile** of a variable and aging physical population, with mechanism-qualified extrapolation, explicit correction and retirement, and complete lifecycle ledgers.

- **Status:** fixture mathematics; no new principle or candidate
- **Comparison unit:** one preregistered service interval and its physical cohort
- **Primary rule:** never infer recovery, reliability, or efficiency from task score, monitor output, accelerated stress, or component energy alone

Physical identity and time base

For service episode e at sample time t , seal the identity envelope

$$I_{e,t} = (l, w, d, b, p, h, s, e, t),$$

where l identifies a fabrication lot, w a wafer, d a die, b a physical block or array, p a package and board path, h a hardware and firmware version, s a site or facility, e an episode, and t elapsed time [s] from a declared clock origin. All identifiers are immutable byte strings. Replacement, repair, reprogramming, remapping, firmware change, or calibration creates a new versioned identity link; it does not overwrite history.

Let $\Delta t_n = t_{n+1} - t_n$ [s] be sample interval n . The actual mission profile over interval $[0, T_e]$ is

$$M_e = \left\{ u_n, V_n, f_n, T_n, J_n, a_n, \phi_n, \dot{D}_n^{\text{ion}}, c_n, r_n, \Delta t_n \right\}_{n=0}^{N_e-1},$$

where u_n is workload class [class], V_n supply or terminal voltage [V], f_n clock or operation rate [Hz], T_n measured absolute temperature [K], J_n current density [A m^{-2}], a_n switching or access activity [dimensionless], ϕ_n particle flux [$\text{particle m}^{-2} \text{s}^{-1}$], $\dot{D}_n^{\text{ion}}$ ionizing dose rate [Gy s^{-1}], c_n cooling state [state], r_n route and protection state [state], and N_e is the number of intervals [interval]. $T_e = \sum_n \Delta t_n$ is episode duration [s]. Commanded voltage, nominal temperature, or benchmark label cannot substitute for the measured histories.

The workload record is further resolved as

$$u_n = (q_n, N_n^{\text{op}}, B_n^{\text{move}}, N_n^{\text{write}}, y_n),$$

where q_n is requested service type [type], N_n^{OP} is operation count [operation], B_n^{move} is bytes moved [byte], N_n^{write} is write or program count [write], and y_n is the required quality and safety envelope [contract].

Latent physical state and observable evidence

Let the latent device state be

$$z_n = (\theta, D_n^{\text{perm}}, R_n^{\text{rev}}, W_n, \mathcal{F}_n, Q_n^{\text{res}}),$$

where θ is the time-zero physical parameter vector in declared native units, D_n^{perm} is cumulative irreversible damage [damage unit], R_n^{rev} is reversible degradation [damage unit], W_n is consumed write, cycle, or stress endurance [cycle or declared wear unit], $\mathcal{F}_n$ is latent fault state [state], and Q_n^{res} is remaining repair, spare, timing, thermal, and correction reserve [declared reserve unit]. A dimensionless normalized representation is allowed only after every component scale is fixed.

The physical transition law is

$$z_{n+1} = g_k(z_n, M_{e,n}, \alpha_k, \Gamma_k, \xi_n),$$

where g_k is a mechanism-qualified transition model, k indexes a physical mechanism, $M_{e,n}$ is the mission-profile slice, α_k is a vector of mechanism parameters in declared native units, Γ_k contains interaction coefficients, and ξ_n is process noise in the units of z . A model that mixes mechanisms must identify k or explicitly carry a mixture state.

Observed telemetry is

$$o_n = h(z_n, M_{e,n}, \beta_{v_n}) + \epsilon_n, \quad \epsilon_n \sim p_{\epsilon, v_n},$$

where o_n is the observation vector in sensor-native units, h is the measurement function, β_{v_n} is calibration state under calibration version v_n , and p_{ϵ, v_n} is the version-qualified noise law. Calibration state has covariance Σ_n^β in squared native units.

Define observation availability $m_{n,j} \in \{0, 1\}$ for channel j and censoring bounds $L_{n,j}$ and $U_{n,j}$ in the channel's native unit. The observation record is

$$O_n = (o_n, m_n, L_n, U_n, v_n, \Sigma_n^\beta, a_n^{\text{ev}}),$$

where $a_n^{\text{ev}} = t_n - t_n^{\text{last qualified}}$ is evidence age [s]. Adaptive physical interfaces require the same explicit latent inventory, mode, hysteresis, depletion, health, evidence-age, and fallback state; an “adaptive” label is not an observation (C-1504). For a right-censored lifetime $t_i^{\text{f}} > C_i$, unit i contributes

$$\mathcal{L}_i = S(C_i \mid M_i),$$

where C_i is censor time [s], S is survival probability [dimensionless], and M_i is the unit's observed mission profile. A failed unit with failure time t_i^{f} [s] and classified mechanism k_i contributes

$$\mathcal{L}_i = \lambda_{k_i}(t_i^{\text{f}} \mid M_i) S(t_i^{\text{f}} \mid M_i),$$

where λ_{k_i} is mechanism-specific hazard [s^{-1}]. Missing and censored records are represented in the likelihood; they are not imputed as healthy observations.

Hierarchical variation, yield, and leakage control

For parameter q measured at lot l , wafer w , die d , and block b , use the hierarchical decomposition

$$\theta_{l,w,d,b,q} = \mu_q + L_{l,q} + W_{l,w,q} + D_{l,w,d,q} + B_{l,w,d,b,q} + \eta_{l,w,d,b,q},$$

where μ_q is the population mean, $L_{l,q}$ the lot effect, $W_{l,w,q}$ the wafer effect, $D_{l,w,d,q}$ the die effect, $B_{l,w,d,b,q}$ the local block effect, and $\eta_{l,w,d,b,q}$ measurement residual, all in the native unit of q . Spatial covariance and gradients are modeled explicitly when present.

Let $A_{i,q} = 1$ when physical unit i satisfies acceptance criterion q and $A_{i,q} = 0$ otherwise. Joint accepted yield is

$$\hat{Y}_{\text{joint}} = \frac{1}{N_{\text{fab}}} \sum_{i=1}^{N_{\text{fab}}} \prod_{q=1}^Q A_{i,q},$$

where N_{fab} is the number of fabricated units [unit], Q is the number of jointly required criteria [criterion], and $\hat{Y}_{\text{joint}}$ is dimensionless. Failed, untestable, unpackageable, and discarded dies remain in N_{fab} .

For a die area $A_{\text{die}} [\text{m}^2]$ and random killer-defect density D_0 [defect m^{-2}], the Poisson yield null is

$$Y_P = \exp(-A_{\text{die}} D_0),$$

where Y_P is dimensionless. More flexible clustering models may replace this null only with held-out wafer and lot evidence.

Competing mechanisms and mission-profile damage

For K competing mechanisms, total hazard is

$$\lambda(t \mid M) = \sum_{k=1}^K \lambda_k(t \mid M),$$

and survival through time t is

$$S(t \mid M) = \exp \left[- \int_0^t \lambda(\tau \mid M) \, \text{d}\tau \right],$$

where λ and every λ_k have unit s^{-1} , t and τ have unit s, and S is dimensionless. Cause-specific cumulative incidence is

$$F_k(t \mid M) = \int_0^t S(\tau^- \mid M) \lambda_k(\tau \mid M) \, \text{d}\tau,$$

where F_k is dimensionless and τ^- denotes the instant before τ .

For a monotone damage proxy, define

$$D_k(t) = \int_0^t r_k(V(\tau), T(\tau), J(\tau), a(\tau), \phi(\tau), u(\tau)) \, d\tau,$$

where r_k is mechanism- k damage rate [damage unit s^{-1}] and D_k is accumulated damage [damage unit]. This integral is evaluated on actual telemetry, not on mean voltage or mean temperature. Endpoint-matched wear histories can therefore carry different mechanism, transition, repair, and remaining-service state (C-1500).

The equal-mean mission-history illustration uses hypothetical Arrhenius parameters to visualize this nonlinearity; it is not calibrated device damage.

An Arrhenius acceleration factor between use temperature T_u [K] and stress temperature T_s [K] is

$$AFT = \exp \left[\frac{E_a}{k_B} \left(\frac{1}{T_u} - \frac{1}{T_s} \right) \right],$$

where E_a is activation energy [eV], k_B is Boltzmann's constant [eV K^{-1}], and AFT is dimensionless. The electromigration lifetime null is

$$t_{50} = A_{EM} J^{-n_{EM}} \exp \left(\frac{E_{a,EM}}{k_B T} \right),$$

where t_{50} is median failure time [s], A_{EM} has the compound unit required to yield seconds, J is current density [A m^{-2}], n_{EM} is dimensionless, and $E_{a,EM}$ is activation energy [eV]. The fitted range, waveform, geometry, and failure criterion travel with every estimate.

Radiation-induced upset rate for sensitive regions j is

$$RSEU = \sum_j \int_0^\infty \phi_j(E) \sigma_j(E) \, dE,$$

where E is particle energy [J or eV, declared consistently], $\phi_j(E)$ is differential flux [particle $\text{m}^{-2} \text{s}^{-1} \text{energy}^{-1}$], $\sigma_j(E)$ is upset cross-section [m^2/bit or m^2/device], and $RSEU$ is upset rate [$\text{bit}^{-1} \text{s}^{-1}$ or $\text{device}^{-1} \text{s}^{-1}$].

For mechanisms k and r , interaction departure is

$$\Delta_{k,r}(M) = L_{k+r}(M) - L_k(M) - L_r(M) + L_0(M),$$

where L_{k+r} is loss under combined stress, L_k and L_r are losses under each stress alone, and L_0 is unstressed loss, all in the same task or physical unit. $\Delta_{k,r} = 0$ is the additive null; the sign and uncertainty of $\Delta_{k,r}$ must be reported rather than absorbed into an unspecified "aging" variable.

Accelerated-test support and extrapolation

Let x^{stress} be the vector of stress covariates in their normalized, preregistered coordinates and let $\mathcal{S}_{\text{train}}$ be the support of the accelerated-test design. Define support distance

$$d_{\text{sup}}(x) = \inf_{x' \in \mathcal{S}_{\text{train}}} \|x - x'\|_{\Sigma^{-1}},$$

where Σ is a fixed covariance or scale matrix, $\|v\|_{\Sigma^{-1}} = \sqrt{v^\top \Sigma^{-1} v}$ is dimensionless Mahalanobis distance, and d_{sup} is dimensionless. A prediction is out of support when $d_{\text{sup}}(x) > d_{\text{max}}$ for preregistered dimensionless threshold d_{max} .

For a nominal $(1 - \alpha)$ survival interval $[\hat{S}_i^L(t), \hat{S}_i^U(t)]$, empirical interval coverage is

$$\hat{C}_S(t) = \frac{1}{N_{\text{hold}}} \sum_{i=1}^{N_{\text{hold}}} 1[S_i(t) \in [\hat{S}_i^L(t), \hat{S}_i^U(t)]],$$

where α , S_i , and $\hat{C}_S$ are dimensionless, N_{hold} is held-out unit count [unit], and 1 is the indicator function. Mechanism transitions, failure-analysis disagreement, or false-safe predictions invalidate extrapolation even when aggregate error is small.

Thermal, electrical, and wear coupling

For thermal node vector $T(t)$ [K], the lumped electrothermal null is

$$C_{\text{th}} \frac{dT}{dt} + G_{\text{th}}(T - T_{\text{amb}}) = P(t),$$

where C_{th} is thermal-capacitance matrix [J K^{-1}], G_{th} is thermal-conductance matrix [W K^{-1}], T_{amb} is ambient-temperature vector [K], and $P(t)$ is dissipated power vector [W]. Routing comparisons use measured spatial $T(t)$ and $P(t)$.

Dynamic switching energy for operation class q is

$$E_{\text{dyn},q} = N_q \alpha_q C_q V_q^2,$$

where N_q is operation count [operation], α_q is activity factor [dimensionless], C_q is effective switched capacitance [F/operation], V_q is voltage [V], and $E_{\text{dyn},q}$ is energy [J]. Leakage, regulation, clocking, memory, transfer, monitoring, correction, thermal control, and idle energy are separate terms.

For physical element j , normalized wear evolves as

$$w_j(t + \Delta t) = w_j(t) + \frac{\Delta x_j(t)}{X_j^{\text{end}}},$$

where w_j is dimensionless consumed endurance, Δx_j is stress, write, or cycle increment [wear unit], and X_j^{end} is measured endurance capacity [same wear unit]. Element j is exhausted when $w_j \geq 1$, unless a stricter registered threshold applies.

Separate native margin, reversible recovery, and compensation as

$$m_j^{\text{obs}}(t) = m_j^0 - d_j^{\text{perm}}(t) - d_j^{\text{rev}}(t) + c_j^{\text{comp}}(t),$$

where m_j^{obs} , m_j^0 , permanent loss d_j^{perm} , reversible loss d_j^{rev} , and compensation c_j^{comp} share the same physical margin unit, such as volts or seconds. A reduction in d_j^{rev} is recovery; an increase in c_j^{comp} is adaptation. They are never scored as the same event.

Fault geometry and the soft/hard firewall

Let every fault event carry type

$$f_i = (g_i, \ell_i, \tau_i, p_i, c_i, x_i),$$

where g_i is spatial geometry [bit, word, bank, chip, route, or domain], ℓ_i is persistence [s], τ_i is occurrence time [s], p_i is physical or injected provenance [class], c_i is common-cause identifier [class], and x_i is external-side-effect state [state].

The firewall outcome is one of

$$Y_i^{\text{fw}} \in \{\text{CE}, \text{DUE}, \text{SDC}, \text{MC}, \text{ESC}\},$$

where CE is corrected error, DUE is detected uncorrectable error, SDC is silent data corruption, MC is miscorrection, and ESC is escaped unsafe side effect. Each is counted in events [event]. For N_{tx} protected transactions,

$$R_y = \frac{N_y}{N_{\text{tx}}},$$

where N_y is count [event] of firewall outcome y , and R_y is rate [event/transaction]. SDC and ESC are never merged into average task loss.

For independent per-bit upset probability p_b during scrub interval Δt_s [s] and codeword length n_c [bit], the probability of more than one upset is

$$P_{>1} = 1 - (1 - p_b)^{n_c} - n_c p_b (1 - p_b)^{n_c - 1},$$

where $P_{>1}$ and p_b are dimensionless. This is only a null: burst, adjacent, chip, decoder, timing, permanent, and common-cause faults require their measured geometry.

Evidence-age-qualified control authority

Let m_n^{lb} be a conservative lower bound on timing, voltage, memory, or analog margin in its native unit. Let the proposed operating point consume margin c_n^{op} in the same unit and let reserve requirement r_n^{min} share that unit. Authority is admissible only when

$$m_n^{\text{lb}} - c_n^{\text{op}} \geq r_n^{\text{min}}, \quad a_n^{\text{ev}} \leq a_{\text{max}}, \quad x_n \in \mathcal{V}_n,$$

where a_n^{ev} and maximum evidence age a_{max} are seconds, x_n is current operating covariate vector, and $\mathcal{V}_n$ is the validated operating envelope. Failure of any condition invokes a preregistered safe operating point or stops acceptance.

Let $P_{\text{esc}}(a)$ be probability [dimensionless] that action a causes an escaped protected failure during one transaction. A controller action is permitted only if

$$P_{\text{esc}}(a \mid O_{0:n}, M_{0:n}) \leq \epsilon_{\text{esc}},$$

where ϵ_{esc} is the preregistered per-transaction risk limit [dimensionless]. The bound includes monitor, regulator, clock, policy, and fallback faults rather than conditioning them away.

Analog and in-memory computation state

For programmed conductance matrix G^0 [S], the effective matrix at time t is

$$G(t) = G^0 + \Delta G^{\text{prog}} + \Delta G^{\text{drift}}(t, T) + \Delta G^{\text{cycle}} + \Delta G^{\text{stuck}},$$

where every ΔG term is in siemens [S] and separately denotes programming error, time- and temperature-dependent drift, cycling variation, and stuck-cell error. For input-voltage vector v [V], ideal current is $i = Gv$ [A]. Measured output is

$$\tilde{i} = Q_{\text{ADC}}(\Psi_{\text{wire}}(G, v, T) + n_{\text{ana}}),$$

where Ψ_{wire} maps conductance and voltage to current while including wire and peripheral effects, n_{ana} is analog noise [A], and Q_{ADC} is the converter map from amperes to digital code [code].

Hardware-aware training distribution $P_{\text{train}}(\delta)$ over nonideality vector δ is compared with held-out physical distribution $P_{\text{test}}(\delta)$. The support test uses the previously defined d_{sup} ; confident acceptance outside support is scored separately as silent failure.

Repair, spares, yield, and retirement state

For unit i , lifecycle state is

$$s_i(t) = (a_i, b_i, r_i, q_i, w_i, v_i),$$

where a_i is availability [dimensionless], b_i remaining spare capacity [block or byte], r_i cumulative repair count [repair], q_i current service qualification [class], w_i wear vector [dimensionless], and v_i version record [version]. A repair updates s_i and its provenance; it never resets fabrication yield or prior embodied cost.

Let C_i^{repair} be repair cost in a declared vector of joules, kilograms, person-hours, currency, and downtime seconds. Let V_i^{future} be expected accepted future service [accepted-service unit]. Repair is economically or environmentally admissible only under the registered componentwise budget and risk constraints; a scalar ratio may be reported as

$$\rho_i^{\text{repair}} = \frac{V_i^{\text{future}}}{E_i^{\text{repair}} + E_i^{\text{future}}} \quad [\text{accepted service/J}],$$

where E_i^{repair} and E_i^{future} are repair and future operational energy [J]. Material, labor, risk, and time remain separate ledgers.

Hard retirement indicator is

$$R_i^{\text{hard}} = 1[U_i = 1 \vee P_{\text{esc},i}^U > \epsilon_{\text{esc}} \vee m_i^{\text{lb}} < m_i^{\text{min}} \vee q_i \notin Q_{\text{safe}} \vee b_i < b_i^{\text{min}} \vee v_i \notin \mathcal{V}_i],$$

where U_i is an uncontained or unclassifiable fault indicator [dimensionless], $P_{\text{esc},i}^U$ is the upper confidence bound on escape probability [dimensionless], m_i^{min} is minimum physical margin in the same unit as m_i^{lb} , Q_{safe} is the set of qualified service classes, b_i^{min} is minimum reserve in the same unit as b_i , and $\mathcal{V}_i$ is the set of accepted versions and validity states. When $R_i^{\text{hard}} = 1$, the unit cannot accept protected work. Economic or average-quality gains cannot override this rule.

Accepted service and complete lifecycle ledgers

For transaction j , define acceptance

$$A_j = 1[q_j \in \mathcal{Q}_j \wedge \ell_j \leq \ell_j^{\max} \wedge c_j \in \mathcal{C}_j \wedge y_j^{\text{fw}} \in / \{\text{SDC}, \text{MC}, \text{ESC}\}],$$

where q_j is measured quality in its native unit, $\mathcal{Q}_j$ is the accepted quality set, ℓ_j is latency [s], $\ell_j^{\max}$ is latency limit [s], c_j is calibration and constraint state [state], $\mathcal{C}_j$ is its accepted set, and y_j^{fw} is firewall outcome. Accepted service is

$$S_{\text{acc}} = \sum_{j=1}^{N_{\text{tx}}} A_j \omega_j \quad [\text{accepted service}],$$

where ω_j is registered service value [service unit/transaction]. Report also the unweighted accepted transaction count $\sum_j A_j$ [transaction].

Operational energy is

$$E_{\text{op}} = E_{\text{compute}} + E_{\text{memory}} + E_{\text{move}} + E_{\text{convert}} + E_{\text{monitor}} + E_{\text{correct}} + E_{\text{cal}} + E_{\text{cool}} + E_{\text{idle}} + E_{\text{recover}},$$

where every term is measured in joules [J] at the declared boundary. Lifecycle energy is

$$E_{\text{life}} = E_{\text{fab}} + E_{\text{package}} + E_{\text{test}} + E_{\text{op}} + E_{\text{repair}} + E_{\text{replace}} + E_{\text{eol}},$$

where fabrication, packaging, test, operation, repair, replacement, and end-of-life terms are joules [J] allocated by a published rule. Failed dies, spares, calibration, replacement inventory, and facility overhead remain in scope.

Material and work ledgers are vectors

$$\mathbf{M}_{\text{life}} = (m_1, \dots, m_R), \quad \mathbf{H}_{\text{life}} = (h_1, \dots, h_P),$$

where m_r is mass [kg] of material category r , R is category count, h_p is labor [person-hour] for role p , and P is role count. Carbon dioxide equivalent C_{life} [kg CO₂e] is reported separately with inventory version, geography, time, allocation, and uncertainty.

Energy intensity of accepted service is

$$\eta_E = \frac{E_{\text{life}}}{S_{\text{acc}}} \quad [\text{J/accepted service}],$$

and is undefined when $S_{\text{acc}} = 0$. Energy intensity never replaces the firewall, material, work, availability, latency, or tail-risk outcomes. A coupon-level friction or wear reduction cannot promote without this mission-qualified accepted-service and lifecycle transfer (C-1505).

Matched budget and Pareto comparison

Every arm a receives componentwise budget vector

$$\mathbf{B}_a = (N_{\text{fab}}, A_{\text{silicon}}, N_{\text{sens}}, N_{\text{spare}}, N_{\text{cal}}, M_{\text{label}}, N_{\text{sim}}, N_{\text{train}}, P_{\text{peak}}, E_{\text{life}}, T_{\text{wall}}, B_{\text{store}}, H_{\text{human}}, M_{\text{material}}, R_{\text{risk}}),$$

where the components are fabricated units [unit], silicon area [m²], sensors [sensor], spares [block], calibration observations [observation], labels [label], simulation calls [call], training operations [operation], peak power [W], lifecycle energy [J], wall time [s], stored bytes [byte], human work [person-hour], material mass [kg], and risk allowance [declared risk unit]. Arm a is feasible only if

$$\mathbf{B}_a \preceq \mathbf{B}^{\max},$$

where $\preceq$ means every component is within its preregistered ceiling in the same unit. Removed ablation components do not donate their budgets elsewhere.

The protected outcome vector is

$$\mathbf{Y}_a = (S_{\text{acc}}, R_{\text{SDC}}, R_{\text{ESC}}, Q_{0.99}^{\text{lat}}, A_{\text{avail}}, \widehat{Y}_{\text{joint}}, E_{\text{life}}, \mathbf{M}_{\text{life}}, \mathbf{H}_{\text{life}}, C_{\text{life}}, N_{\text{repair}}, N_{\text{replace}}),$$

where $Q_{0.99}^{\text{lat}}$ is 99th-percentile latency [s], A_{avail} is availability [dimensionless], N_{repair} is repair count [repair], and N_{replace} is replacement count [replacement]; the other components were defined above. Pareto dominance is assessed componentwise after preregistering beneficial directions and hard constraints.

For paired held-out mission e , candidate-minus-null effect on scalar outcome y is

$$\Delta_{e,y} = y_{e,\text{cand}} - y_{e,\text{null}},$$

where $\Delta_{e,y}$ has the unit of outcome y . Report hierarchical intervals grouped by lot, wafer, die, site, workload family, and future time; random-record splits are diagnostic only.

Ten-track measurement map

AUDIT TRACK	REQUIRED QUANTITATIVE CONSTRUCT	DECISIVE HELD-OUT UNIT
E-SEMI-01	$\widehat{Y}_{\text{joint}}$, hierarchy, false accept/reject, post-aging yield	lot, wafer, die, block, future time
E-SEMI-02	$S(t \mid M)$, F_k , AF_T , d_{sup} , coverage, censoring	use-like low stress and mechanism transition
E-SEMI-03	$T(t)$, $P(t)$, $D_k(t)$, w_j , accepted-service lifecycle frontier	unseen spatial workload and cooling regime
E-SEMI-04	f_i , Y_t^{fw} , scrub age, common-cause identity	withheld geometry and persistence class
E-SEMI-05	m^{lb} , a^{ev} , P_{esc} , fallback	monitor, controller, regulator, and compound fault
E-SEMI-06	exact-state boundary, SDC/ESC, verification and fallback cost	distribution, objective, and structured-error shift
E-SEMI-07	$G(t)$, $\tilde{i}$, all peripheral energy, yield, endurance	operator family, device, time, reuse, temperature
E-SEMI-08	$P_{\text{train}}(\delta)$, d_{sup} , calibration, abstention	lot, nonideality, correlation, drift-age combination
E-SEMI-09	w_j , value, reconstruction cost, movement and metadata	skewed, shifting, burst, and adversarial writes
E-SEMI-10	S_{acc} , E_{life} , $\mathbf{M}_{\text{life}}$, $\mathbf{H}_{\text{life}}$, retirement	inventory, electricity, workload, repair, replacement sensitivity

Statistical and retirement contract

The confirmatory analysis preregisters cohort sizes from power or precision targets, all exclusion rules, multiplicity control, censoring model, calibration method, hierarchical grouping, uncertainty propagation, and the direction and minimum relevant magnitude of each effect. Report medians, tails, intervals, per-device traces, failure maps, and unfavorable regimes; do not pool mechanisms or populations merely to obtain significance.

The cross-candidate composition is retained only if, on sealed held-out mission profiles and within $\mathbf{B}^{\max}$:

1. it improves at least one preregistered accepted-service or lifecycle outcome beyond the complete mature null by the minimum relevant magnitude;
2. no hard firewall, coverage, calibration, availability, or retirement limit is violated;
3. the effect survives hierarchy-aware analysis, mechanism and inventory sensitivity cases, and removal of any unnecessary candidate mechanism; and
4. every claimed gain remains after calibration, correction, recovery, failed units, spare consumption, repair, replacement, material, and human work are charged.

Failure invokes the narrowest applicable response: remove the unsupported component, reduce authority, derate or repurpose a qualified unit, or set $R_i^{\text{hard}} = 1$. No result in this contract allocates a new principle or candidate identifier.

Multiscale reduction contract

- **Status:** pre-implementation mathematical and accounting contract
- **Claims:** C-1526--C-1529
- **Source audit:** applied multiscale reduction
- **Fixture:** F-024
- **Result state:** an AMR-T01 development-only smoke runner exists; no sealed confirmation or transfer run, reference-workstation result, measured effect, or energy result exists

This note fixes notation, units, exact identities, approximation boundaries, and resource ledgers for four different reduction methods. It does not assert that a biological or artificial workload has short memory, a slow manifold, a valid homogenized limit, or closed coarse variables.

1. Notation and units

SYMBOL	MEANING	UNIT
t, s, T, t_h	physical, integration, burst, and healing times	s
x, y, z	resolved, unresolved, and full states	component-specific; U in the linear example
$\alpha, \beta, \gamma, \lambda$	linear rate coefficients	s^{-1}
$K(\tau)$	linear memory kernel	s^{-2} when x, y share U
$\eta(t)$	unresolved-initial-condition transient	$U \text{s}^{-1}$
P, Q	projection and complementary projection, $Q = I - P$	dimensionless operators
$\mathcal{L}$	Liouville generator	s^{-1}
τ_f, τ_s	fast and slow reference times	s
$\varepsilon = \tau_f / \tau_s$	timescale ratio	dimensionless
f, g	fast and slow vector fields after time scaling	same state units as x and y , respectively
X_0	declared trajectory-magnitude scale for x	U
S_0	critical manifold $f(x, y, 0) = 0$	state space
σf	minimum magnitude of fast spectral real part	dimensionless in fast time; s^{-1} if dimensionalized
u, U	fine and coarse states	component-specific
Q_c	compression/restriction from fine to coarse state	typed operator
$R(U, \xi)$	reconstruction/lift with admissible lift index ξ	fine state
Φ_T^f, Φ_T^c	fine and induced coarse propagators over T	typed maps
L, x, ℓ	slab length, coordinate, and microperiod	m
$T_h(x)$	temperature field; subscript avoids confusion with burst time	K
κ	thermal conductivity	$\text{W m}^{-1} \text{K}^{-1}$
q	volumetric heat source	W m^{-3}
$J = -\kappa dT_h/dx$	heat flux	W m^{-2}
N_*	event or operation count	count

SYMBOL	MEANING	UNIT
B_*	attributed data traffic or storage	byte
$t_{\text{cpu}}, t_{\text{wall}}$	measured process and elapsed times	s
M_{RSS}	peak resident-set size	byte

Subscripts identify objects; they do not silently change units. Every implementation artifact must store a unit string beside dimensional fields. Dimensionless normalization constants remain in the manifest rather than being discarded after preprocessing.

2. Exact resolved memory in a coupled linear system

2.1 Full model

Consider

$$\dot{x}(t) = -\alpha x(t) + \beta y(t), \quad \dot{y}(t) = \gamma x(t) - \lambda y(t), (1)$$

with $[x] = [y] = U$ and $[\alpha] = [\beta] = [\gamma] = [\lambda] = \text{s}^{-1}$. Each right-hand term has unit $U \text{ s}^{-1}$. Stability of the two-dimensional linear system is guaranteed in the fixture by

$$\alpha + \lambda > 0, \quad \alpha \lambda - \beta \gamma > 0. (2)$$

Equation (2) is a generator restriction, not a biological or deployment limit.

2.2 Eliminate the unresolved state

Multiplying the second equation by $e^{\lambda t}$ gives

$$\frac{d}{dt} (e^{\lambda t} y(t)) = \gamma e^{\lambda t} x(t). (3)$$

Integrating from 0 to t and multiplying by $e^{-\lambda t}$ yields

$$y(t) = e^{-\lambda t} y_0 + \gamma \int_0^t e^{-\lambda(t-s)} x(s) \, ds. (4)$$

Substitution into the first equation gives the exact resolved equation

$$\dot{x}(t) = -\alpha x(t) + \underbrace{\beta e^{-\lambda t} y_0}_{\eta(t)} + \int_0^t \underbrace{\beta \gamma e^{-\lambda(t-s)}}_{K(t-s)} x(s) \, ds. (5)$$

The three contributions in (5) are instantaneous resolved dynamics, a transient determined by unresolved initial state, and a memory convolution. They are exact for (1); no stochastic premise is required.

2.3 Dimensional analysis

The kernel has

$$[K] = [\beta \gamma] = \text{s}^{-2}.$$

Since $[x(s) \, ds] = U \text{ s}$,

$$\left[\int_0^t K(t-s)x(s) ds \right] = U s^{-1}.$$

Also $[\eta] = [\beta y_0] = U s^{-1}$. Thus every term in (5) has the unit of $\dot{x}$. An implementation that stores K as a dimensionless attention weight without its time discretization is not implementing (5).

For uniform step Δt , a left-rule discrete memory approximation is

$$M_n = \Delta t \sum_{j=0}^{n-1} K((n-j)\Delta t)x_j, (6)$$

where $\Delta t K$ has unit s^{-1} . A learned discrete coefficient $w_j = \Delta t K(j\Delta t)$ therefore has unit s^{-1} , not unit one.

2.4 Markov and finite-memory approximations

If λ is large relative to the resolved variation rate, x is smooth on $1/\lambda$, and the initial layer has decayed, then

$$\int_0^t e^{-\lambda(t-s)}x(s) ds = \frac{x(t)}{\lambda} + O(\lambda^{-2}\dot{x}) + O\left(\frac{X_0}{\lambda}e^{-\lambda t}\right). (7)$$

Here X_0 is a declared bound or reference scale with the same unit U as x . All three terms in (7) therefore have unit $U s$; the final term records the exponentially decaying initial-layer contribution rather than adding a dimensionless remainder to a dimensional integral.

This gives the leading Markov approximation

$$\dot{x} \approx \left(-\alpha + \frac{\beta\gamma}{\lambda} \right) x. (8)$$

Equation (8) is conditional on timescale separation and initial-layer control. It is not obtained by merely setting $y = 0$. A finite window W instead drops

$$R_W(t) = \int_0^{\max(0, t-W)} K(t-s)x(s) ds, (9)$$

whose error requires a bound on both kernel tail and trajectory magnitude. A short fitted window is not self-validating.

The original two-state system (1) is an exact augmented-state representation. It is the strongest compact null for this example: a finite-memory model must not claim mathematical novelty merely for approximating a system that the latent state represents exactly.

3. General projection identity and its boundary

Let the autonomous fine dynamics be $\dot{z} = F(z)$ and let observables evolve under the Liouville generator

$$\mathcal{L}A(z) = F(z) \cdot \nabla A(z). (10)$$

For a declared projection P on observables, $Q = I - P$, and sufficient regularity for the semigroups below, the Dyson identity gives one common Mori–Zwanzig form:

$$\frac{d}{dt}e^{t\mathcal{L}}A = e^{t\mathcal{L}}P\mathcal{L}A + e^{tQ\mathcal{L}}Q\mathcal{L}A + \int_0^t e^{(t-s)\mathcal{L}}P\mathcal{L}e^{sQ\mathcal{L}}Q\mathcal{L}A ds. (11)$$

Depending on convention, equivalent rearrangements place propagators or projectors differently. The implementation must record the chosen convention; pieces from incompatible conventions cannot be combined by label.

In (11):

1. $e^{t\mathcal{L}}P\mathcal{L}A$ is the propagated Markov term;
2. $e^{tQ\mathcal{L}}Q\mathcal{L}A$ is the orthogonal-dynamics term; and
3. the integral is the memory term.

If A has unit U , then $\mathcal{L}A$ and the first two right-hand terms have unit $U s^{-1}$. Inside the integral, two generators contribute $U s^{-2}$ and integration restores $U s^{-1}$. Calling the orthogonal-dynamics term “noise” does not establish independence, stationarity, Gaussianity, or zero mean.

The 1961 nonlinear transport derivation must be read with the 1972 correction for omitted fluctuations. This note relies on the directly checked linear derivation (1)–(5) for the fixture and treats (11) only within its declared operator assumptions.

4. Fast--slow geometry and normal hyperbolicity

4.1 Scaling

Start with reference times $\tau_f, \tau_s > 0$:

$$\tau_f \frac{dx}{dt} = f(x, y, \varepsilon), \quad \tau_s \frac{dy}{dt} = g(x, y, \varepsilon), \quad \varepsilon = \frac{\tau_f}{\tau_s}. (12)$$

Here f and g have the same state units as x and y respectively. With slow time $\theta = t/\tau_s$, (12) becomes

$$\varepsilon x' = f(x, y, \varepsilon), \quad y' = g(x, y, \varepsilon), (13)$$

where prime means $d/d\theta$ and the vector fields carry the corresponding state units. They become dimensionless only after an additional, explicitly recorded state normalization. With fast time $r = t/\tau_f$,

$$\dot{x}_r = f(x, y, \varepsilon), \quad \dot{y}_r = \varepsilon g(x, y, \varepsilon). (14)$$

4.2 Critical and slow manifolds

At $\varepsilon = 0$, the critical manifold is

$$S_0 = \{(x, y) : f(x, y, 0) = 0\}. (15)$$

For a compact submanifold $K \subset S_0$, define the fast spectral margin

$$\sigma f(K) = \inf_{(x, y) \in K} \min_{\mu \in \text{spec } D_x f(x, y, 0)} |\text{Re } \mu|. (16)$$

K is normally hyperbolic when $\sigma f(K) > 0$ and the splitting required by the theorem is uniform. Under the requisite smoothness and compactness, Fenichel persistence supplies a locally invariant S_ε near K for sufficiently small ε . The theorem is local to the qualified compact set and sufficiently small, not all, ε .

4.3 Boundary at a fold

For the dimensionless fold normal form used in [AMR-T02](#),

$$\varepsilon x' = y - x^2, \quad y' = -1, (17)$$

$S_0 = \{y = x^2\}$. The fast Jacobian is

$$D_x f = -2x. (18)$$

The branch $x > 0$ is attracting in fast time, the branch $x < 0$ is repelling, and at $(x, y) = (0, 0)$ the spectral margin vanishes. Ordinary normal-hyperbolic persistence therefore does not cover the fold point. Special fold/canard analysis can describe additional behaviour under additional hypotheses; it does not retroactively make (16) positive.

An operational gate may estimate a lower margin, an invariance residual and an off-manifold distance, but that estimate is a diagnostic, not a theorem. Its false-acceptance and abstention rates must be tested against a full stiff solver, including initial layers and held-out fold approaches.

5. Compression, reconstruction, and local micro queries

Let X_f and X_c be typed fine and coarse state spaces. Define

$$Q_c : X_f \rightarrow X_c, \quad R : X_c \times \Xi \rightarrow X_f, (19)$$

where Q_c is compression/restriction and $R(U, \xi)$ is a reconstruction or lift indexed by admissible unresolved detail ξ . Exact consistency would be

$$Q_c R(U, \xi) = U. (20)$$

Numerical consistency instead records

$$e_{QR}(U, \xi) = \|Q_c R(U, \xi) - U\|_{X_c} (21)$$

and compares it with a frozen tolerance. Equation (20) or small (21) does not guarantee dynamical closure.

Suppose a macro scheme advances

$$U_{n+1} = \mathcal{M}_{\Delta T}(U_n; D(U_n)), (22)$$

but the macro datum $D(U)$, such as a constitutive flux, is unavailable. HMM uses a constrained micro problem on support $\omega(U)$:

$$u_{m+1}^\xi = \mathcal{S} \delta t(u_m^\xi; U, \omega, b, \xi), \quad \widehat{D}(U) = \mathcal{A}(\{u_m^\xi\}_{m=m_h}^{m_h+m_a}), (23)$$

where b is the micro boundary rule, m_h is relaxation/healing count, m_a is averaging count and $\mathcal{A}$ is the declared estimator. The error must be decomposed, at minimum, as

$$e_{\text{total}} \leq e_{\text{macro}} + e_{\text{micro}} + e_{\text{boundary}} + e_{\text{relax}} + e_{\text{sampling}} + e_{QR}.(24)$$

The terms need not be statistically independent and may not be combined in quadrature without a proof. A local solve is not cheap by definition; support, unknowns, coefficient queries, iterations, retries and transfers all enter the ledger.

For the one-dimensional heat track,

$$-\frac{d}{dx} \left(\kappa(x/\ell) \frac{dT_h}{dx} \right) = q(x).(25)$$

dT_h/dx has unit K m^{-1} , so the inner flux has unit W m^{-2} and its divergence W m^{-3} , matching q . For $\kappa(\xi) = \kappa_0[2 + \sin(2\pi\xi + \phi)]$, the one-dimensional periodic effective coefficient is the harmonic mean

$$\kappa_{\text{eff}} = \left(\int_0^1 \frac{d\xi}{\kappa(\xi)} \right)^{-1} = \kappa_0 \sqrt{3}.(26)$$

This analytic result is the strongest null for the single-sinusoid family, not information that may be hidden from a comparator.

6. Lift--evolve--restrict and equation-free closure

Given a fine propagator Φ_T^f , define the lift-specific coarse map

$$\Phi_T^c(U; \xi) = Q_c \Phi_T^f(R(U, \xi)).(27)$$

With healing t_h followed by measurement burst T_b , one derivative estimator is

$$\widehat{F}_c(U; \xi) = \frac{Q_c \Phi_{t_h+T_b}^f(R(U, \xi)) - Q_c \Phi_{t_h}^f(R(U, \xi))}{T_b}.(28)$$

For an admissible lift set Ξ_U , define post-healing disagreement

$$D_{\text{lift}}(U; t_h, T_b) = \max_{\xi_i, \xi_j \in \Xi_U} \left\| \widehat{F}_c(U; \xi_i) - \widehat{F}_c(U; \xi_j) \right\|_{X_c}.(29)$$

A small value is evidence only at the tested state, lift family, healing time, burst, ensemble size and norm. Persistent disagreement falsifies the proposed closure at that horizon. Agreement can be falsely induced by using identical lifts, an insensitive restriction, an overlong burst that crosses the decision horizon, shared random numbers without independent variance checks, or a norm that suppresses the omitted coordinate.

A projective step

$$U_{n+1} = U_n + \Delta T \widehat{F}_c(U_n).(30)$$

must record stability rejections and fallback. Work spent on an abandoned projective step remains work.

7. Error, work, traffic, and microstep ledgers

7.1 Accuracy records

Every seed, world and arm records protected errors before aggregation:

$$e_{L^2}(a) = \left(\frac{1}{T} \int_0^T \|Q_{C^Z a}(t) - Q_{C^Z O}(t)\|_2^2 dt \right)^{1/2}, (31)$$

or the protocol's spatial analogue, plus maximum error, flux/event error, coverage, abstention, failure and closure disagreement. Numerical quadrature and interpolation rules are frozen. An absent output receives the terminal loss; it is not omitted.

7.2 Compute ledger

For every arm, record nonnegative integer counters:

COUNTER	REQUIRED CONTENT
$N_{\text{fine,RHS}}$	all fine-model RHS evaluations, including rejected solver steps
$N_{\text{coarse,RHS}}$	all explicit coarse-model RHS evaluations
N_{lin}	linear iterations, including failed and setup iterations
N_{micro}	replica-, cell-, or particle-microsteps, not merely burst calls
N_{macro}	accepted and rejected macrosteps, separated in raw fields
N_{lift}	reconstruction/lifting calls for every replica and retry
N_{restrict}	compression/restriction calls
N_{heal}	microsteps executed only for healing/relaxation
N_{query}	coefficient, simulator, or oracle queries permitted to the arm
N_{retry}	retries and fallback invocations

A declared operation proxy is

$$W_{\text{proxy}} = c_f N_{\text{fine,RHS}} + c_c N_{\text{coarse,RHS}} + c_l N_{\text{lin}} + c_m N_{\text{micro}} + c_L N_{\text{lift}} + c_Q N_{\text{restrict}}, (32)$$

where every coefficient c_* is a frozen measured instruction- or multiply-add-equivalent calibration with manifest hash. Track-specific named work may use a raw counter such as N_{micro} to avoid hiding a dominant cost inside weights. Report both. Do not tune c_* after results. Counters used together in (32) must represent disjoint work. If, for example, N_{micro} is a labelled subset of $N_{\text{fine,RHS}}$, the manifest sets one corresponding coefficient to zero and records the parent-- child relation; nested counters may not be charged twice.

7.3 Traffic and memory ledger

Attributed algorithmic traffic is

$$B_{\text{moved}} = B_{\text{state,read}} + B_{\text{state,write}} + B_{\text{history,read}} + B_{\text{history,write}} + B_{\text{coeff}} + B_{\text{lift}} + B_{\text{restrict}} + B_{\text{solver}} + B_{\text{artifact}}. (33)$$

Each component is counted from typed-array lengths times element widths plus serialized metadata bytes. Allocation alone is not traffic. Cache-line or hardware-counter estimates may be added as a separately labelled measurement; they may not replace (33) silently. Record:

1. logical bytes read and written by component;
2. artifact bytes written;
3. peak live algorithmic bytes from allocation instrumentation;

4. process $MRSS$;
5. CPU and wall seconds; and
6. implementation, runtime, operating-system and processor manifest hashes.

7.4 HMM and equation-free microstep identities

For HMM macro queries $q = 1, \dots, N_q$ with microcell unknown count n_q , relaxation steps h_q , averaging steps a_q and solver iterations i_q ,

$$N_{\text{micro}}^{\text{HMM}} = \sum_{q=1}^{N_q} n_q (h_q + a_q + i_q), (34)$$

unless the implementation supplies a more literal disaggregated counter. It is invalid to report only N_q .

For equation-free coarse steps $k = 1, \dots, N_K$, lifts $j = 1, \dots, n_{\xi,k}$, replicas $r = 1, \dots, n_{r,k}$ and executed fine steps m_{kjr} including healing, burst, rejected projection and fallback,

$$N_{\text{micro}}^{\text{EF}} = \sum_{k=1}^{N_K} \sum_{j=1}^{n_{\xi,k}} \sum_{r=1}^{n_{r,k}} m_{kjr}. (35)$$

Adaptive increases in lift count, replicas, healing, or burst length remain in (35). A fallback full solve is added, not substituted after deleting the failed coarse attempt.

7.5 Equal-resource and equal-information comparisons

For compared arms A/B/C, freeze:

1. identical resolved observations and timestamps;
2. identical coefficient or simulator-query authority;
3. identical public development worlds and private pack commitments;
4. identical tuning-call, CPU-thread, wall-time and memory caps;
5. no fine hidden state or analytic answer unique to the proposed arm;
6. identical terminal-loss handling; and
7. complete pre-fallback plus fallback accounting.

If an analytic null legitimately knows a coefficient formula, that knowledge is recorded as problem information and its storage/precomputation is stated. The proposed arm may not claim an information advantage by withholding a closed-form solution that is part of the declared problem family.

8. Validation invariants

Before any experiment result can be interpreted, a runner must prove:

1. equations (1) and (5) match to numerical tolerance on random stable systems;
2. discrete kernel units include Δt and converge under step halving;
3. zero coupling makes the exact memory term zero;
4. the fold gate reports loss of margin as $x \rightarrow 0$;
5. reconstruction error (21) is recorded for every lift/micro query;
6. the analytic harmonic mean (26) matches direct quadrature;

7. replica relabelling leaves restrictions and results invariant;
8. operation, microstep and byte ledgers reconcile with raw events;
9. failures and fallbacks remain in every aggregate; and
10. no field labelled energy or joules is emitted without an independently specified physical measurement protocol.

Until these invariants, the four fixture tracks, sealed confirmation, and transfer gates are complete, this note supports only a testable mathematical contract. It contains no performance result.

Notation and units

System variables

SYMBOL	MEANING	UNIT
x	one input event or task instance	declared by task
Q	task-quality measure	task-specific
R	risk, error, or miscalibration measure	task-specific
$Lp95$	95th-percentile end-to-end latency	seconds
E	energy over a declared interval	joules
P	average power over a declared interval	watts = joules/second
T	measurement duration	seconds
N	qualified events completed	count
B	bytes transferred across a named boundary	bytes
C	counted arithmetic or module work	declared operation
$g_i(x)$	gate for module i on event x	dimensionless
$c_i(x)$	cost proxy for module i	operations, bytes, or joules
z_t	latent state at time t	dimensionless vector
σ_t	predicted uncertainty for z_t	same scale as distance
s_t	normalized surprise	dimensionless

Symbol scope

The table above reserves unqualified project-wide symbols. A chapter may add subscripts or superscripts, but it must not reuse a reserved symbol for a different physical quantity. In particular, P remains power and s_t remains normalized surprise. Predictors use symbols such as $f\theta$, and an environment state uses a locally declared symbol such as u_t .

Time durations use seconds. Discrete sequence horizons use an explicit count such as T_τ or T_B rather than the unqualified duration T . Every local objective must state the unit of its value and of every coefficient that makes unlike terms commensurable.

Required qualifiers

FLOP, operation, token, event, and sample are not interchangeable. Every use must define:

- arithmetic precision;
- dense, effective-sparse, or executed count;
- training or inference;
- batch and sequence shape; and

- whether memory and communication are included.

Boundary labels

- E_{device} : accelerator package only.
- E_{node} : accelerator, CPU, memory, local storage, and node cooling.
- E_{cluster} : nodes plus network and shared infrastructure.
- E_{facility} : cluster energy multiplied by a contemporaneous, attributable facility overhead.

Results must not move between boundaries by implication.

Occupancy-qualified spatial transition

This note defines the narrow built-environment stress track for Candidate 001. Its purpose is to expose state and cost that disappear when a topology edit is represented as an instantaneous graph operation.

The contract applies only when physical or service topology changes while people, critical services, or hazardous inventory remain inside the affected dependency boundary. If the asset can be isolated and a fixed approved sequence dominates, ordinary change control is the relevant baseline.

Typed transition state

For transition version v and step k , keep the state tuple

$$X_{v,k} = (G, Y, O, R, F, S, U, H, C, E)_{v,k}.$$

The fields are deliberately not collapsed:

- G : surveyed physical and service topology, including uncertainty;
- Y : required service by user group, location, and time;
- O : observed or forecast occupancy, ability, familiarity, and assistance;
- R : ordinary, accessible, construction, emergency, responder, and goods routes, including capacity and closures;
- F : detection, alarm, suppression, compartment, smoke and heat state, and scenario-specific evacuation timing;
- S : structural, utility, control, and indoor-environment state;
- U : ownership, permits, operating restrictions, acceptance authority, stop-work power, and their expiry;
- H : material and work state—installed, removed, stored, in transit, reusable, damaged, waste, temporary, and irreversible;
- C : cost, energy, carbon, water, waste, and disruption accounts; and
- E : evidence support, timestamp, version, calibration, coverage, uncertainty, and invalidation dependencies.

A plan or model is an input to E ; it is not a substitute for G , O , F , or S . Likewise, an actuator command belongs to the control record while verified equipment response belongs to S .

Intermediate-state admissibility

Let $\mathcal{H}$ be the preregistered set of hard constraints. The step gate is

$$\mathcal{A}_{v,k} = 1 \left[\bigwedge_{h \in \mathcal{H}} g_h(X_{v,k}) \geq 0 \right] 1[E_{v,k} \text{ is sufficient and current}] 1[U_{v,k} \text{ permits the step}],$$

where every indicator is dimensionless. Each g_h retains its own physical unit and threshold; the conjunction does not add seconds, newtons, contaminant concentration, and route availability into a fictitious scalar score.

For fire scenario s , one necessary margin is

$$m_{s,v,k}^{\text{egress}} = ASET_{s,v,k} - RSET_{s,v,k} - M_s \quad [\text{s}],$$

with $m_{s,v,k}^{\text{egress}} \geq 0$. Here $ASET$ is available safe egress time, $RSET$ is required safe egress time, and M_s is the frozen scenario margin, all in seconds. Passing this one margin does not establish structural, accessibility, utility, or environmental admissibility.

For user group g , define accessible route availability as

$$a_{g,v,k} = 1[\exists \text{ usable route for } g] 1[q_{g,v,k} \geq q_g^{\min}],$$

where a is dimensionless, route capacity q and its minimum $q^{\min}$ use the same declared unit such as persons per second, and “usable” includes the group’s mobility and assistance requirements. A graph path that is blocked, untenable, too narrow, or unsupported does not satisfy the predicate.

The controller may execute step k only when $\mathcal{A}_{v,k} = 1$. Otherwise it must abstain or enter a pre-approved restricted, decanted, isolated, or safe-stop state. It cannot average one hard violation against an energy saving.

Evidence-age gate

For evidence item j , let measurement time be t_j , current decision time be t , and maximum permitted age under the current hazard and work state be $\tau_j(X_{v,k})$, all in seconds. Its freshness indicator is

$$f_j(t, X_{v,k}) = 1[t - t_j \leq \tau_j(X_{v,k})] 1[d_j(X_{v,k}) = 0],$$

where d_j is a dimensionless invalidation flag. Moving a wall, isolating a circuit, changing occupancy, impairing suppression, or changing the sensor support can set $d_j = 1$ even before the clock expires. Required evidence is current only when $\prod_{j \in \mathcal{J}_{v,k}} f_j = 1$.

Service and transition burden

For user group g over transition interval $[t_0, t_1]$, service loss is

$$L_g = \int_{t_0}^{t_1} [Q_{g,\text{ref}}(t) - Q_g(t)]_+ dt.$$

If Q_g is usable service places, L_g is place-hours; if Q_g is flow in persons per hour, L_g is persons. The service unit must be declared, and the vector $\mathbf{L} = (L_1, \dots, L_G)$ remains visible so an aggregate cannot hide loss concentrated on one group.

Transition burden is reported as a vector rather than an automatic weighted sum:

$$\mathbf{B} = (\mathbf{L}, T_{\text{transition}}, E_{\text{transition}}, GWP_{\text{transition}}, C_{\text{transition}}, W_{\text{material}}, N_{\text{defect}}, N_{\text{hard violation}}).$$

The components use, respectively, declared service-unit-hours, seconds, joules, kilograms CO₂e, currency, kilograms, a defect count, and a hard-violation count. Any scalarization must publish its weights, units, and decision authority. The primary feasibility condition is $N_{\text{hard violation}} = 0$ for every protected scenario—not a favorable mean.

Fast reversible allocation and slow structural adaptation therefore retain different action, deficit, build, carry, and stranded-capacity coordinates (C-1496).

Commitment and recovery

Let $K_{v,k}$ be the set of material or legal commitments already made. A software rollback is physically available only if a tested predecessor state $X_{v,k}^-$ remains reachable under the current $K_{v,k}$, resources, authority, and time. Define

$$rv_{v,k} = \mathbf{1}[X_{v,k}^- \in \mathcal{R}(X_{v,k}, K_{v,k})],$$

where r is dimensionless and $\mathcal{R}$ is the set of reachable, verified states. When $rv_{v,k} = 0$, the record must name a reachable safe-stop or forward-recovery state; calling the transition reversible is false.

Recovery after a demand-regime reversal is part of the structural decision, not an optional endpoint after the selected regime has already been scored (C-1496).

Recovery is complete only after target service, evidence validity, open-defect limits, and reserve are restored. For each service component m ,

$$T_m^{\text{recover}} = \inf\{t \geq t_f : Q_m(t : t + \Delta) \geq Q_m^{\text{target}} \wedge R_m(t : t + \Delta) \geq R_m^{\text{min}}\} - t_f,$$

where t_f is fault or interruption time, Δ is the frozen sustainment window, Q and reserve R use declared service units, and T^{recover} is seconds. Reopening without reserve restoration is reported separately.

Equal-budget comparison and rejection

The residual contract is compared with the complete ordinary stack: surveyed as-built records, code and professional review, permits, impairment control, configuration management, construction sequencing, commissioning, post-occupancy evaluation, and lifecycle asset management. All arms receive the same asset, staff time, information, sensors, professional review, material, construction window, compute, maintenance, and whole-life budget.

Reject or merge the residual when any of the following holds:

1. the asset can be isolated and a fixed approved sequence matches outcomes;
2. a protected-user accessibility or life-safety constraint is violated;
3. the advantage disappears when stale evidence, failed actuators, occupancy changes, transition duration, irreversible work, and common causes are represented;
4. mature change control plus commissioning matches the protected-service and lifecycle frontier; or
5. apparent recovery omits displaced exposure, defects, reserve restoration, or the next disruption.

The editable transition diagram is occupancy-qualified-spatial-transition.mmd. The evidence boundary is developed in the built-environment audit.

Operator-qualified active acoustic inference

This note formalizes Fixture F-009 from the acoustics, hearing, and auditory-scene analysis audit. It supplies a hostile comparison boundary for Candidate 002, Candidate 006, Candidate 007, Candidate 009, Candidate 012, and Candidate 014.

Episode, operator, and action identity

For receiver r , source s , microphone or ear channel m , and episode e , preserve

$$\mathcal{A}_e = (X_e, S_e, E_e, R_e, H_e, O_e, C_e, T_e, U_e, B_e),$$

where:

- X_e is geometry, medium, boundaries, temperature in kelvins, relative humidity as a dimensionless fraction, flow in metres per second, occupancy, and time-varying room/material state;
- S_e is every target and interfering source identity, waveform, position in metres, velocity in metres per second, directivity, emission time in seconds, and source-level record;
- E_e is commanded and realized active emission: waveform, spectrum, duration in seconds, acoustic energy in joules, aim in degrees, repetition, and dose;
- R_e is receiver/body/array position and orientation, aperture in metres, morphology, feasible motion, transducer response, gain, health, and saturation;
- H_e is prior acoustic exposure, room/source history, adaptation, training, previous emissions and motions, and feedback with timestamps;
- O_e is the observation operator: impulse responses, transfer functions, spatial/time support, sample rate in samples per second, quantization in bits, clock, latency in seconds, preprocessing, missingness, and selection;
- C_e is calibration identity, uncertainty, reference pressure and distance, instrument class, traceability, and data vintage;
- T_e is the target construct, literal outcome, deadline in seconds, and loss or utility in declared units;
- U_e is the independent unit: waveform, frame, event, source, room, receiver, body, array, device, site, or population; and
- B_e is the complete ceiling in samples, events, bytes, seconds, person-hours, joules, emissions, exposure, unsafe events, sensors, actuators, replacements, and opportunity.

For method q and literal outcome k , the estimand is

$$Q_{q,k}(\mathcal{A}) = \mathbb{E}[Y_k \mid do(q), \mathcal{A}],$$

where Y_k uses the registered unit for outcome k . A contrast against baseline b does not isolate q if aperture, calibration, room response, source distribution, observation receipt time, emission or motion authority, training scenes, or any binding resource differs without intervention.

Pressure, level, exposure, and clock contract

For acoustic pressure $p(t)$ in pascals over interval T in seconds,

$$p_{\text{rms}} = \sqrt{\frac{1}{T} \int_0^T p^2(t) dt}, \quad L_p = 20 \log_{10} \left(\frac{p_{\text{rms}}}{p_0} \right),$$

where p_{rms} is RMS pressure in pascals, $p_0 = 20 \mu\text{Pa}$ is the reference pressure in air, and L_p is in decibels relative to $20 \mu\text{Pa}$. Frequency weighting, time weighting, band, location, orientation, and calibration are fields, not implied defaults.

Sound exposure and its level are

$$E_p = \int_0^T p^2(t) dt, \quad L_E = 10 \log_{10} \left(\frac{E_p}{E_0} \right), \quad E_0 = p_0^2 t_0,$$

where E_p is in pascal-squared seconds, L_E is in decibels relative to E_0 , and reference duration $t_0 = 1$ s. Exposure is not acoustic emission energy, electrical energy, or a universal risk model.

For channel clock c , correct a device timestamp by

$$\bar{t}_{c,n} = t_{c,n}^{\text{dev}} - \delta_{c,v},$$

where device time $t_{c,n}^{\text{dev}}$, corrected time $\bar{t}_{c,n}$, and version- v offset $\delta_{c,v}$ are in seconds; c is the channel-clock index, n is the dimensionless sample index, and v is the dimensionless calibration version. Retain residual clock uncertainty $\sigma_{c,v}$ in seconds, drift in seconds per second, capture time, receipt time, synchronization method, and calibration interval. At decision time t , only observations with receipt time no later than t are available.

Propagation, rooms, and time-varying mixing

At channel m , use the time-varying mixture

$$y_m(t) = \sum_{s=1}^S \int h_{m,s}(t, \tau) x_s(t - \tau) d\tau + n_m(t),$$

where received waveform $y_m(t)$, source pressure $x_s(t)$, and noise $n_m(t)$ are in pascals; S is dimensionless source count; t and delay τ are in seconds; and transfer kernel $h_{m,s}(t, \tau)$ is in reciprocal seconds. A time-invariant room convolution is a special case. Source/receiver motion, changing boundaries, temperature, flow, occupancy, and adaptive emitters make the operator time-dependent.

For an ideal free-field point source with unchanged directivity and negligible absorption,

$$\Delta L_p = 20 \log_{10} \left(\frac{r_1}{r_2} \right),$$

where distances r_1, r_2 are in metres and ΔL_p is in decibels. The comparison must expose directionality, near-field terms, boundaries, atmospheric absorption, scattering, and receiver orientation when present.

For a diffuse-field Sabine approximation,

$$T_{60} \approx 0.161 \frac{V}{A},$$

where T_{60} is decay time in seconds, room volume V is in cubic metres, equivalent absorption area A is in square metres, and 0.161 has units seconds per metre. Report measured impulse responses and uncertainty by band, source, receiver, and support; this approximation is not an operator identity.

Masking, filterbank, compression, and events

A real gammatone-like channel is

$$g_j(t) = a_j t^{n_j-1} e^{-2\pi b_j t} \cos(2\pi f_j t + \phi_j), \quad t \geq 0,$$

where centre frequency f_j and bandwidth b_j are in hertz, filter order n_j is dimensionless, phase ϕ_j is in radians, time t is in seconds, and a_j has units s^{-n_j} so g_j is in reciprocal seconds. Fixed FFT, wavelet, mel, ERB, gammatone/gammachirp, modulation, and learned filterbanks remain competing representations.

A local normalized compression fit is

$$\frac{z}{z_0} = \left(\frac{x}{x_0} \right)^\gamma, \quad 0 < \gamma \leq 1,$$

where magnitudes x, x_0 share one input unit, z, z_0 share one output unit, and exponent γ is dimensionless. Register frequency, level, state, attack/release time in seconds, saturation, and distortion; one exponent does not describe an adaptive cochlea or compressor globally.

For yes/no target detection,

$$d' = \Phi^{-1}(P_{\text{hit}}) - \Phi^{-1}(P_{\text{false alarm}}),$$

where both P terms are dimensionless probabilities, Φ^{-1} is the inverse standard-normal cumulative distribution, and sensitivity d' is dimensionless. Keep decision criterion, energetic overlap, informational uncertainty, spatial release, source grouping, and task history separate.

For event encoder threshold θ_j in the unit of channel response $u_j(t)$,

$$e_{j,n} = (j, t_{j,n}, u_j(t_{j,n}), v_\theta, v_c) \quad \text{when} \quad |u_j(t_{j,n}) - u_j(t_{j,n-1})| \geq \theta_j,$$

where j is channel identity, event time $t_{j,n}$ is in seconds, n is the dimensionless event index, $e_{j,n}$ is the retained event record, $t_{j,n-1}$ is the previous retained-event time in seconds, and v_θ, v_c are dimensionless threshold and clock versions. Event rate is in events per second. Report missed sustained signals, false events, timestamp error, decoder cost, bytes, task information, and measured joules; event sparsity is not an energy unit.

ITD, ILD, aperture, correlation, and beamforming

For sound speed c_{snd} in metres per second and path-length difference Δr in metres,

$$\Delta t = \frac{\Delta r}{c_{\text{snd}}},$$

where interaural or interchannel time difference Δt is in seconds. Clock offset, phase ambiguity, multipath, source extent, and head/array transfer functions are part of its uncertainty.

For left/right RMS pressures p_L, p_R in pascals,

$$\text{ILD} = 20 \log_{10} \left(\frac{p_L}{p_R} \right),$$

where interaural level difference is in decibels and is band-, time-, source-, and orientation-qualified. ITD, ILD, spectral cues, and head motion remain separate intervention axes.

For far-field array aperture D in metres, frequency f in hertz, and wavelength $\lambda = c_{\text{snd}}/f$ in metres,

$$\Delta\theta \sim \frac{\lambda}{D},$$

where angular resolution $\Delta\theta$ is in radians. Geometry, beampattern, SNR, estimator, bandwidth, and resolution definition determine the coefficient; no learned estimator restores spatial frequencies excluded by the physical support without additional prior information.

For signals $y_1(t), y_2(t)$ in pascals, define generalized cross-correlation

$$R_{12}^{(\Psi)}(\tau) = \int_{-\infty}^{\infty} \Psi(f) Y_1(f) Y_2^*(f) e^{i2\pi f\tau} df, \quad \hat{\tau} = \arg \max_{\tau \in \mathcal{T}} R_{12}^{(\Psi)}(\tau),$$

where Y_1, Y_2 are Fourier transforms in pascal-seconds, frequency f is in hertz, delay τ and estimate $\hat{\tau}$ are in seconds, $\mathcal{T}$ is the feasible delay set in seconds, superscript $*$ is complex conjugation, and weighting $\Psi(f)$ has reciprocal pascal-squared-second-squared units so $R_{12}^{(\Psi)}$ is in hertz after integration. The peak must carry a calibrated multimodal delay distribution under reverberation, not only an argmax.

For array snapshot $\mathbf{y} \in \mathbb{C}^M$ in pascals, dimensionless steering vector $\mathbf{a}$, and noise covariance $\mathbf{R}_n = \mathbb{E}[\mathbf{nn}^H]$ in pascal-squared units,

$$\mathbf{w}_{\text{MVDR}} = \frac{\mathbf{R}_n^{-1} \mathbf{a}}{\mathbf{a}^H \mathbf{R}_n^{-1} \mathbf{a}}, \quad z = \mathbf{w}_{\text{MVDR}}^H \mathbf{y},$$

where M is microphone count, superscript H is conjugate transpose, $\mathbf{w}_{\text{MVDR}}$ is dimensionless, and output z is in pascals. Steering mismatch, covariance error, short sample support, correlated sources, motion, and room change receive sealed interventions.

Reverberation, separation, and source identity

For direct component $d_m(t)$, early-reflection component $r_m^{\text{early}}(t)$, late component $r_m^{\text{late}}(t)$, and noise $n_m(t)$, all in pascals,

$$y_m(t) = d_m(t) + r_m^{\text{early}}(t) + r_m^{\text{late}}(t) + n_m(t).$$

The early/late boundary is a registered time in seconds relative to the direct arrival. Waveform dereverberation, speech compensation, perceived distance, localization, and downstream intelligibility are distinct outcomes.

For estimated source $\hat{\mathbf{s}}$ and reference source $\mathbf{s} \in \mathbb{R}^N$ with the same amplitude unit and sample count N ,

$$\mathbf{s}_{\text{target}} = \frac{\langle \hat{\mathbf{s}}, \mathbf{s} \rangle}{\|\mathbf{s}\|_2} \mathbf{s}, \quad \text{SI-SDR} = 10 \log_{10} \frac{\|\mathbf{s}_{\text{target}}\|_2^2}{\|\hat{\mathbf{s}} - \mathbf{s}_{\text{target}}\|_2^2},$$

where $\langle \cdot, \cdot \rangle$ is the waveform inner product and SI-SDR is in decibels. Report unknown source count, permutation, causal identity, localization, intelligibility, calibration, perceptual/task quality, and residual mixture separately.

For predicted scene outcome z_n and causally available acoustic observation o_n , proper log loss is

$$L_q = -\frac{1}{N_e} \sum_{n=1}^{N_e} \log_2 p_q(z_n | o_n),$$

where N_e is independent event count and L_q is in bits per event. Use it for detection, source count, assignment, localization bins, range bins, and abstention under held-out operators; do not pool incompatible outcomes.

Active emission, dose, and closed-loop action

For monostatic emission at t_{emit} and echo receipt at t_{recv} , both in seconds,

$$\hat{r} = \frac{c_{\text{snd}}}{2} (t_{\text{recv}} - t_{\text{emit}}),$$

where estimated range $\hat{r}$ is in metres and sound speed c_{snd} is in metres per second. Clock uncertainty, target motion, refraction, multipath, transducer ringing, waveform ambiguity, association, and detector threshold must propagate to range uncertainty.

For acoustic output power $P_{\text{ac}}(t)$ in watts over emission duration T_e in seconds,

$$E_{\text{emit}} = \int_0^{T_e} P_{\text{ac}}(t) dt,$$

where E_{emit} is acoustic joules. Electrical input, transduction loss, body/sensor motion, sensing, compute, cooling, detectability, interference, and exposure remain separate axes.

At decision time t , a complete acoustic policy is

$$(a_t^{\text{emit}}, a_t^{\text{body}}, a_t^{\text{sensor}}, a_t^{\text{gain}}, a_t^{\text{task}}) = \pi_q(\mathcal{H}_t, \hat{X}_t, \hat{O}_t, \hat{U}_t, \mathcal{A}_t^{\text{safe}}, B_t),$$

where a_t^{emit} contains waveform, level, spectrum, aim, and timing; a_t^{body} contains pose or head/body motion in metres, radians, and seconds; a_t^{sensor} contains aperture, sample rate, and channel selection; a_t^{gain} is dimensionless or in declared decibels; a_t^{task} is the downstream command in its native unit; $\mathcal{H}_t$ is causally received history; $\hat{X}_t$ is scene state; $\hat{O}_t$ is operator/calibration state; $\hat{U}_t$ is uncertainty; $\mathcal{A}_t^{\text{safe}}$ is the feasible action envelope; and B_t is remaining budget in matched units. Attention- or efferent-like gain receives causal credit only when intervening on it changes its registered endpoint.

Hardware, lifecycle energy, and equal budgets

Lifecycle energy is

$$E_q^{\text{life}} = E_q^{\text{data}} + E_q^{\text{train}} + E_q^{\text{emit}} + E_q^{\text{sense}} + E_q^{\text{infer}} + E_q^{\text{move}} + E_q^{\text{comm}} + E_q^{\text{store}} + E_q^{\text{cal}} + E_q^{\text{maint}} + E_q^{\text{emb}},$$

where every term is in joules over one service interval and denotes data acquisition, training, acoustic emission, sensing, inference, physical motion, communication, storage, calibration, maintenance, and amortized embodied energy. DSP, FPGA, ASIC, neuromorphic, CPU, GPU, and analog front ends are measured at identical physical boundaries and workload deadlines.

Human effort is

$$H_q^{\text{human}} = H_q^{\text{design}} + H_q^{\text{record}} + H_q^{\text{label}} + H_q^{\text{listen}} + H_q^{\text{cal}} + H_q^{\text{tune}} + H_q^{\text{monitor}} + H_q^{\text{repair}},$$

where every H term is in person-hours and roles are reported separately.

The complete cost vector is

$$\mathbf{C}_q = (N_{\text{sample}}, N_{\text{event}}, N_{\text{step}}, N_{\text{query}}, N_{\text{byte}}, T_{\text{wall}}, H_q^{\text{human}}, E_q^{\text{life}}, E_p, N_{\text{unsafe}}, D_{\text{harm}}, C_{\text{opp}}),$$

where the five N terms count samples, emitted/received events, optimization or environment steps, queries, and bytes; T_{wall} is seconds; H_q^{human} is person-hours; E_q^{life} is joules; E_p is pascal-squared-second exposure; N_{unsafe} counts unsafe events; D_{harm} uses a registered physical or severity unit; and C_{opp} is opportunity cost in task exposures or person-hours.

Method q is feasible only when

$$\mathbf{C}_q \preceq \mathbf{B},$$

where $\mathbf{B}$ is the preregistered componentwise ceiling with identical units. Over-budget runs are infeasible; removed ablation resources stay unused.

Confirmatory contrast and retirement

Let $b^*(j)$ be the strongest mature baseline for track j , frozen on development data. Orient every protected endpoint as a preregistered benefit, so larger Q is better and a non-inferiority margin may be negative. For protected outcomes $k \in \mathcal{P}_j$, retain a residual only when

$$\Pr(Q_{q,k} - Q_{b^*(j),k} > \delta_{j,k} \text{ for every } k \in \mathcal{P}_j) \geq 1 - \alpha_j,$$

where $\delta_{j,k}$ is an improvement or non-inferiority margin in the unit of outcome k , and α_j is the dimensionless error budget. The result must survive held-out rooms, sources, arrays, bodies, operators, scenes, source counts, interference policies, model families, sites, and hardware classes at equal complete cost. Otherwise retire the architectural residual while keeping the operator/action/measurement contract.

Operator-qualified active chemical sensing

This note formalizes Fixture F-011 from the olfaction, chemical sensing, and plume-tracking audit. It supplies a hostile comparison boundary for Candidate 002, Candidate 006, Candidate 007, Candidate 009, Candidate 010, Candidate 012, Candidate 014, Candidate 017, and Candidate 018. The fixture and this note create no principle or architecture candidate.

Episode, operator, and outcome identity

For episode e , preserve

$$\mathcal{C}_e = (S_e, X_e, A_e, R_e, O_e, K_e, H_e, T_e, U_e, B_e),$$

where:

- S_e is every target, interferent, and source identity; source mixture; release rate in moles per second; phase; temperature in kelvins; geometry in metres; motion in metres per second; and source-selection history;
- X_e is the domain, boundaries, surfaces, flow field in metres per second, pressure in pascals, relative humidity as a dimensionless fraction, temperature in kelvins, turbulence, chemistry, sorption, and background;
- A_e is every commanded and realized motion, orientation, sniff or pump waveform, volumetric flow in cubic metres per second, heater power in watts, valve, preconcentration, purge, query, confirmation, and stopping action;
- R_e is receiver/body identity, morphology, pose, bilateral spacing or array geometry in metres, feasible motion, inlet, tubing, chamber, pump, heater, transducer, health, saturation, and authority;
- O_e is the observation operator: transport and sampling support, causal response/recovery kernels, cross-sensitivity, nonlinearity, quantization, timestamps, latency in seconds, preprocessing, missingness, and selection;
- K_e is calibration identity and validity: reference-gas composition, concentration and uncertainty, zero/span and blank history, flow, temperature/humidity compensation, device and batch, age, drift, poisoning, maintenance, traceability, and data vintage;
- H_e is prior chemical exposure, adaptation, habituation, contamination, storage, cleaning, training, reinforcement, previous actions, feedback, and readout-remapping history with timestamps;
- T_e is the literal outcome, deadline in seconds, loss/utility, abstention, exposure rule, and safety policy;
- U_e is the independent unit: molecule, injection, vial, sample, sensor, device, manufacture batch, day, source, plume realization, site, body, animal, subject, or population; and
- B_e is the componentwise ceiling in samples, standards, labels, channels, aperture, actions, metres, seconds, bytes, optimization/search trials, person-hours, joules, consumables, emissions, exposures, unsafe events, replacements, embodied hardware, and opportunity.

For method q and literal outcome k , define

$$Q_{q,k}(\mathcal{C}) = \mathbb{E}[Y_k \mid do(q), \mathcal{C}],$$

where Y_k retains the registered unit for outcome k . The contrast does not isolate q if source chemistry, concentration range, plume realization, receiver, calibration, operator, history, action authority, confirmation access, or any binding budget differs without a registered intervention.

Chemical amount, concentration, and conversion

For analyte i , amount concentration is

$$c_i = \frac{n_i}{V},$$

where amount n_i is in moles, volume V is in cubic metres, and c_i is in moles per cubic metre. Mass concentration is

$$\rho_i = c_i M_i,$$

where molar mass M_i is in kilograms per mole and ρ_i is in kilograms per cubic metre.

For an ideal gas with dimensionless amount fraction x_i ,

$$c_i = x_i \frac{P}{RT}, \quad \rho_i = x_i \frac{PM_i}{RT},$$

where pressure P is in pascals, absolute temperature T is in kelvins, and $R = 8.314462618 \text{ J}/(\text{mol K})$. A conversion from parts per million by volume to milligrams per cubic metre therefore carries analyte molar mass, temperature, pressure, and the definition of the fraction.

Transport, reaction, sorption, and intermittent plumes

A continuum starting model for analyte i is

$$\frac{\partial c_i}{\partial t} + \mathbf{u} \cdot \nabla c_i = \nabla \cdot (D_i \nabla c_i) + R_i(\mathbf{c}, T, P, H_r, \mathbf{x}, t) + q_i(\mathbf{x}, t),$$

where position $\mathbf{x}$ is in metres; time t is in seconds; velocity $\mathbf{u}$ is in metres per second; diffusivity or declared effective dispersion D_i is in square metres per second; relative humidity H_r is dimensionless; reaction, loss, and phase-transfer term R_i is in moles per cubic metre per second; and volumetric source q_i has the same unit. Every term has units of moles per cubic metre per second.

For surface Γ , a general molar flux condition is

$$-D_i \nabla c_i \cdot \mathbf{n} = J_{i,\Gamma}(c_i, \eta_\Gamma, T, H_r, t),$$

where outward unit normal $\mathbf{n}$ is dimensionless, surface flux $J_{i,\Gamma}$ is in moles per square metre per second, and surface state η_Γ records adsorption, desorption, wetting, reaction, and history. Terrain, buoyancy, droplets, thermal stratification, and unresolved turbulent fluxes cannot be hidden inside D_i without declaring the validity regime.

For a registered detection boundary c_i^{det} in moles per cubic metre, define a whiff indicator and cumulative occupation time by

$$w_i(t) = \mathbb{I}[c_i(\mathbf{x}_r(t), t) \geq c_i^{\text{det}}], \quad T_i^{\text{whiff}} = \int_0^{T_e} w_i(t) dt,$$

where receiver trajectory $\mathbf{x}_r(t)$ is in metres, episode duration T_e and whiff occupation T_i^{whiff} are in seconds, and $\mathbb{I}[\cdot]$ is dimensionless. Report the distributions of whiff duration, blank duration, peak, integral, rise/fall and encounter spacing; a time-averaged concentration is not a substitute.

Dynamic cross-sensitive observation operator

For sensor or receptor channel m sampled at device time t_n , use

$$y_{m,n} = g_{m,v} \left(\sum_{i=1}^I \int_0^\infty h_{m,i,v}(\tau; \mathbf{z}_n) c_i(\mathbf{x}_r(t_n - \tau), t_n - \tau) d\tau, \mathbf{z}_n \right) + \epsilon_{m,n},$$

where I is dimensionless analyte count; channel output $y_{m,n}$ and error $\epsilon_{m,n}$ use the channel's calibrated unit; response kernel $h_{m,i,v}$ is in reciprocal seconds; delay τ is in seconds; v is the operator/calibration version; $g_{m,v}$ maps amount concentration to output; and state $\mathbf{z}_n$ includes flow, heater, chamber, temperature, humidity, pressure, interferents, saturation, adaptation, age, drift and poisoning. The integral is in moles per cubic metre. A static feature vector is a special case that must survive response, recovery, hysteresis and support interventions.

Device time is corrected by

$$\bar{t}_{m,n} = t_{m,n}^{\text{dev}} - \delta_{m,v},$$

where device time $t_{m,n}^{\text{dev}}$, corrected time $\bar{t}_{m,n}$, and clock offset $\delta_{m,v}$ are in seconds. Retain clock drift in seconds per second, jitter and residual uncertainty in seconds, capture time, receipt time, and synchronization version. At decision time t , only observations received no later than t are causally available.

Mixture identifiability and null spaces

Under a local linearization around concentration vector $\mathbf{c}_0 \in \mathbb{R}_+^I$, let

$$\Delta \mathbf{y} \approx \mathbf{J}_v(\mathbf{c}_0, \mathbf{z}) \Delta \mathbf{c} + \boldsymbol{\epsilon}, \quad [\mathbf{J}_v]_{m,i} = \left. \frac{\partial \mathbb{E}[y_m]}{\partial c_i} \right|_{\mathbf{c}_0, \mathbf{z}, v},$$

where $\Delta \mathbf{y} \in \mathbb{R}^M$ is in channel-output units, $\Delta \mathbf{c} \in \mathbb{R}^I$ is in moles per cubic metre, M is channel count, and Jacobian element $J_{m,i}$ has output-unit cubic metres per mole. Full column rank of $\mathbf{J}_v$ is necessary for unconstrained local recovery when $M \geq I$, but is not sufficient under noise, saturation, unknown interferents, changing v , or nonlinear ambiguity.

The observation-equivalent set at tolerance ε_y is

$$\mathcal{N}_v(\mathbf{y}) = \left\{ \mathbf{c} \in \mathcal{S}_c : \|\mathbf{y} - G_v(\mathbf{c}; \mathbf{z})\|_{\Sigma_y^{-1}} \leq \varepsilon_y \right\},$$

where supported composition set $\mathcal{S}_c$ uses moles per cubic metre, forward operator G_v returns channel outputs, error covariance Σ_y is in squared output units, Mahalanobis norm is dimensionless, and threshold ε_y is dimensionless. Identification must abstain when materially different identity, concentration, exposure or hazard states remain in $\mathcal{N}_v(\mathbf{y})$.

For Gaussian error and differentiable mean $\boldsymbol{\mu}(\boldsymbol{\theta})$, the local Fisher information is

$$\mathbf{F}(\boldsymbol{\theta}) = \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}} \right)^T \Sigma_y^{-1} \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}} \right),$$

where parameter vector $\boldsymbol{\theta}$ contains registered identities, concentrations, source coordinates, and operator states with declared units. Near-singular directions identify local non-identifiability; a learned decoder does not remove them without additional prior or action-generated evidence.

Concentration, identity, mixtures, and calibration

Keep the protected outcome vector

$$\mathbf{Y} = (Y_{\text{det}}, Y_{\text{id}}, Y_{\text{conc}}, Y_{\text{mix}}, Y_{\text{dir}}, Y_{\text{pos}}, Y_{\text{attr}}, Y_{\text{val}}, Y_{\text{exp}}, Y_{\text{haz}}),$$

whose elements respectively measure presence, chemical or perceptual identity, concentration, mixture composition, direction, position, physical-source attribution, valence, exposure, and hazard. Units and losses differ; no scalar average may allow one to substitute for another.

For yes/no detection,

$$d' = \Phi^{-1}(P_{\text{hit}}) - \Phi^{-1}(P_{\text{false alarm}}),$$

where both probabilities and sensitivity d' are dimensionless. Report the criterion, concentration/matrix, target-absent mixtures and uncertainty.

For concentration estimate $\hat{c}_i$ in moles per cubic metre, a dimensionless log error is

$$\ell_{i,n}^{\text{conc}} = \left| \log \frac{\hat{c}_{i,n} + c_i^*}{c_{i,n} + c_i^*} \right|,$$

where positive reference c_i^* is in moles per cubic metre and is frozen before evaluation. Also report bias and absolute error in native units; c_i^* cannot be tuned on the confirmatory split.

For categorical identity prediction $p_q(z_n \mid o_n)$,

$$L_q^{\text{id}} = -\frac{1}{N} \sum_{n=1}^N \log_2 p_q(z_n \mid o_n),$$

where N is independent episode count, z_n is the registered identity, o_n is causally available evidence, and L_q^{id} is in bits per episode. Report confusion, unknown rejection, calibration and risk--coverage by held-out source, concentration, mixture, device, batch, day and site.

For calibration parameter vector $\boldsymbol{\kappa}$ with covariance $\Sigma_{\boldsymbol{\kappa}}$, first-order propagated output covariance is

$$\Sigma_{y,\text{cal}} \approx \mathbf{J}_{\boldsymbol{\kappa}} \Sigma_{\boldsymbol{\kappa}} \mathbf{J}_{\boldsymbol{\kappa}}^T, \quad \mathbf{J}_{\boldsymbol{\kappa}} = \frac{\partial \boldsymbol{\mu}_y}{\partial \boldsymbol{\kappa}},$$

where each covariance retains the squared units of its parameters or outputs. Reference-gas uncertainty, flow, blank, zero/span, temperature, humidity, device, batch and validity interval are part of κ , not post-hoc notes.

Adaptation, recovery, drift, and poisoning

Separate fast receptor/sensor state from slow condition state:

$$\mathbf{r}_{n+1} = f_r(\mathbf{r}_n, \mathbf{c}_n, a_n; v) + \boldsymbol{\xi}_n, \quad \mathbf{d}_{e+1} = f_d(\mathbf{d}_e, \mathcal{E}_e, m_e; v) + \boldsymbol{\omega}_e,$$

where within-episode response state $\mathbf{r}_n$ may include occupancy, adaptation, heater and recovery; between-episode state $\mathbf{d}_e$ includes age, contamination, baseline/gain drift and poisoning; concentration $\mathbf{c}_n$ is in moles per cubic metre; acquisition action a_n carries its physical units; cumulative exposure and stress $\mathcal{E}_e$ uses a declared vector of concentration-time, temperature-time and electrical stress; maintenance action m_e records purge, cleaning, recalibration or replacement; and errors $\boldsymbol{\xi}_n, \boldsymbol{\omega}_e$ retain state units.

For a step ending at time t_0 , define a registered recovery time

$$t_{\text{rec}}(\epsilon) = \inf \left\{ t \geq t_0 : \frac{|y(t) - y_{\text{blank}}|}{s_y} \leq \epsilon \text{ continuously for } T_{\text{hold}} \right\} - t_0,$$

where output $y(t)$, blank output y_{blank} and scale s_y share the channel unit; tolerance ϵ is dimensionless; hold time T_{hold} and recovery time t_{rec} are in seconds. Recovery does not prove restored calibration, selectivity or absence of poisoning; reference challenges must test those outcomes separately.

Temporal codes, active sampling, and receiver motion

For causal feature window W seconds, preserve a temporal record

$$\mathcal{E}_{m,W} = \{(t_j, y_j, v, K_j) : t - W < t_j \leq t\},$$

where event time t_j is in seconds, value y_j uses the calibrated channel unit, operator version v is dimensionless, and K_j is calibration state. Every event threshold, refractory rule, interpolation, derivative, clock and response kernel is versioned. Time shuffling must preserve marginal concentration, duty cycle and event count when testing whether temporal order adds information.

At decision time t , the active policy is

$$a_t = \pi_q(\mathcal{H}_t, \hat{\mathbf{c}}_t, \hat{\mathbf{s}}_t, \hat{\mathbf{O}}_t, \hat{\mathbf{U}}_t, \mathcal{A}_t^{\text{safe}}, \mathbf{B}_t),$$

where $\mathcal{H}_t$ is causally received observations and actions; $\hat{\mathbf{c}}_t$ is concentration/mixture belief in moles per cubic metre; $\hat{\mathbf{s}}_t$ is source state with position in metres and release rate in moles per second; $\hat{\mathbf{O}}_t$ is operator/condition belief; $\hat{\mathbf{U}}_t$ is uncertainty; $\mathcal{A}_t^{\text{safe}}$ is the feasible action set; and remaining budget $\mathbf{B}_t$ retains componentwise units.

Action value under possible next observation Y is

$$\text{EVI}(a_t) = \min_d \mathbb{E}[L(d, \theta) \mid \mathcal{H}_t] - \mathbb{E}_{Y \sim p(\cdot \mid \mathcal{H}_t, a_t)} \left[\min_d \mathbb{E}[L(d, \theta) \mid \mathcal{H}_t, a_t, Y] \right] - C(a_t),$$

where decision d , target state θ , loss L and action cost C use one registered utility unit. C includes latency, motion, sampled amount, exposure, pump/heater/valve energy, wear, consumables and opportunity. Positive EVI favors the action. Adaptive sniffing receives no credit if it merely samples more chemical mass or receives more time.

Plume-source belief and search

For source state $\mathbf{s}$ and concentration field $\mathbf{c}_{0:t}$, posterior inference is

$$p(\mathbf{s}, \mathbf{c}_{0:t}, O_t \mid y_{1:t}, a_{1:t}, C_e),$$

where source position is in metres, release rate in moles per second, concentration in moles per cubic metre, and operator state O_t includes response, calibration and health. A particle filter, state-space estimator, Gaussian-process plume model, infotaxis policy, POMDP, model-predictive controller, finite-state surge-cast policy and matched-memory reinforcement learner are competing nulls.

For stopping time τ_q in seconds and source-location estimate $\hat{\mathbf{x}}_{s,q}$ in metres, one source-search vector is

$$\mathbf{y}_q^{\text{search}} = (\mathbb{I}[\text{success}], \|\hat{\mathbf{x}}_{s,q} - \mathbf{x}_s\|_2, \tau_q, L_q, E_q, N_q^{\text{false}}, N_q^{\text{unsafe}}),$$

where path length L_q is in metres, episode energy E_q is in joules, and the success indicator and false/unsafe declaration counts are dimensionless. Report every element; success conditional on successful trials is not a valid policy comparison.

Receptor, representation, association, and valence causality

A receptor-like front end or learned representation $z = f_q(y)$ earns causal credit only through a registered intervention. For endpoint k , define

$$\Delta_{z,k} = \mathbb{E}[Y_k \mid do(z = z^{\text{full}}), \mathcal{C}] - \mathbb{E}[Y_k \mid do(z = z^{\text{abl}}), \mathcal{C}],$$

where z^{abl} removes only the registered channel, temporal state, normalization, sparse route or associative readout, and released resources stay unused. The effect $\Delta_{z,k}$ retains outcome k 's unit. Activation, sparsity, mutual information, decoding and anatomical analogy do not substitute for target detection, concentration, mixture, source, transfer, valence, exposure, safety, latency or energy outcomes.

For association episode e , keep

$$\mathcal{L}_e = (o_e, r_e, c_e, \pi_e, f_e, t_e),$$

where odor evidence o_e carries its operator identity, reinforcement r_e uses the task's utility unit, context c_e is registered, policy/intervention π_e is versioned, feedback f_e is timestamped, and acquisition time t_e is in seconds. Chemical identity, learned category, innate choice, learned choice, pleasantness, toxicity and hazard remain distinct labels and losses.

Exposure, safety, and authority

External mass-concentration exposure along receiver or subject path is

$$E_i^{\text{ext}} = \int_0^{T_e} \rho_i(\mathbf{x}_r(t), t) dt,$$

where E_i^{ext} is in kilogram-seconds per cubic metre, commonly reported as milligram-minutes per cubic metre; ρ_i is in kilograms per cubic metre; and time is in seconds. Exposure is not absorbed dose or risk. Route, respiration, susceptible population, toxicokinetics, averaging time, short-term limit, ceiling and immediately dangerous concentration are separate.

The admissible action set is

$$\mathcal{A}_t^{\text{safe}} = \{a : \Pr(g_j(x_{t:t+H}, a) > 0 \mid \mathcal{H}_t) \leq \beta_j \text{ for every registered constraint } j\},$$

where prediction horizon H is in seconds; constraint g_j uses its native unit and is positive on violation; and risk ceiling β_j is dimensionless. Exposure, flammability, collision, contamination, saturation, calibration age, poisoning and authority can each shrink the set. The sensing policy cannot self-certify its safety envelope without an independent monitor or validated fallback.

Analytical confirmation and staged verification

Let screen S emit class, concentration, uncertainty and abstention, and let confirmatory method V return chromatography, spectrometry or other registered evidence. The conditional value of confirmation is

$$\text{EVI}(V \mid S) = \min_d \mathbb{E}[L(d, \theta) \mid S] - \mathbb{E}V \left[\min_d \mathbb{E}[L(d, \theta) \mid S, V] \right] - C(V),$$

where L and $C(V)$ share a declared utility unit. Confirmation cost includes sample handling, standards, blanks, turnaround, carrier gas, columns/sorbents, vacuum/ionization or detector power, compute, analyst time, exposure, and sample destruction. A library hit is not a privileged oracle; recovery, retention, deconvolution, coverage and uncertainty remain part of V .

Lifecycle energy, human work, and equal budgets

Lifecycle energy over one accepted service interval is

$$E_q^{\text{life}} = E_q^{\text{data}} + E_q^{\text{train}} + E_q^{\text{move}} + E_q^{\text{pump}} + E_q^{\text{heat}} + E_q^{\text{sense}} + E_q^{\text{separate}} + E_q^{\text{ionize}} + E_q^{\text{infer}} + E_q^{\text{comm}} + E_q^{\text{store}} + E_q^{\text{cal}} + E_q^{\text{maint}} + E_q^{\text{facility}} + E_q^{\text{emb}},$$

where every term is in joules and covers data acquisition, training, physical motion, pumping, heating, analytical separation, ionization/vacuum when used, sensing, inference, communication, storage, calibration, maintenance, facility overhead and amortized embodied hardware. Carrier and calibration gases, sorbents, columns, dopants, filters, cleaning and replacements are additionally reported in their physical and environmental units rather than silently converted to compute joules.

Human effort is

$$H_q^{\text{human}} = H_q^{\text{design}} + H_q^{\text{sample}} + H_q^{\text{label}} + H_q^{\text{cal}} + H_q^{\text{analyze}} + H_q^{\text{tune}} + H_q^{\text{safety}} + H_q^{\text{monitor}} + H_q^{\text{maint}},$$

where every term is in person-hours and roles are reported separately.

The complete resource vector is

$$\mathbf{C}_q = (N_{\text{sample}}, N_{\text{standard}}, M_{\text{label}}, N_{\text{step}}, N_{\text{tune}}, N_{\text{byte}}, T_{\text{wall}}, L_{\text{path}}, V_{\text{sample}}, H_q^{\text{human}}, E_q^{\text{life}}, \mathbf{E}^{\text{ext}}, N_{\text{unsafe}}, N_{\text{replace}}, C_{\text{opp}}),$$

where the six N terms count samples, standards, labels, optimization or environment steps, tuning trials and bytes; wall time T_{wall} is in seconds; path L_{path} is in metres; sampled volume V_{sample} is in cubic metres; human work is in person-hours; lifecycle energy is in joules; exposure vector $\mathbf{E}^{\text{ext}}$ retains kilogram-seconds per cubic metre by analyte; unsafe events and replacements are counts; and opportunity C_{opp} uses registered task exposures or person-hours.

Method q is feasible only when

$$\mathbf{C}_q \preceq \mathbf{B},$$

where $\mathbf{B}$ is the preregistered componentwise ceiling with identical units. Over-budget runs are infeasible; failed runs remain in denominators and resources released by an ablation stay unused.

Confirmatory contrast and hard retirement

Let $b^*(j)$ be the strongest frozen mature baseline for track j . Orient protected endpoints so larger Q is better and use negative values only for preregistered non-inferiority margins. For protected outcomes $k \in \mathcal{P}_j$, retain a track residual only when

$$\Pr(Q_{q,k} - Q_{b^*(j),k} > \delta_{j,k} \text{ for every } k \in \mathcal{P}_j) \geq 1 - \alpha_j,$$

where margin $\delta_{j,k}$ has outcome k 's unit and error budget α_j is dimensionless. The result must survive held-out chemicals, mixtures, concentrations, release/transport regimes, plume seeds, sources, receivers/bodies, sensors, batches, operator/calibration versions, days, sites, model families and hardware at equal complete cost. Otherwise retire the architectural residual while retaining the operator/action/exposure contract.

Operator-qualified optical inference contract

This note defines the quantitative boundary for Fixture F-007. It operationalizes the measurement-operator-aware residual from the [optics](#), [photonics](#), and [inverse-sensing audit](#). The contract binds every decoded output to a versioned physical operator, separates measurement information from prior selection, and compares optical, digital, and hybrid routes at equal task information and lifecycle budget.

Episode and operator identity

For episode e and acquisition t , seal

$$I_{e,t} = (X_e, A_{e,t}, \nu_{e,t}, C_{e,t}, R_{e,t}, Q_e, B_e),$$

where X_e identifies the physical scene or latent-state generator, $A_{e,t}$ identifies the acquisition action, $\nu_{e,t}$ is an immutable operator-version identifier, $C_{e,t}$ is the calibration record, $R_{e,t}$ is the hidden regime record, Q_e is the registered downstream-query set, and B_e is the resource-budget record. Identifiers and hashes are byte strings; timestamps within the records are seconds [s] from a declared clock origin.

Let $\mathcal{X}$ be the admissible latent-state space and let $x_e \in \mathcal{X}$ be the latent physical state. Each component of x_e retains its native unit, such as radiance [$\text{W sr}^{-1} \text{m}^{-2}$], range [m], or concentration [mol m^{-3}]. The observation is

$$y_{e,t} = g_{e,t}(\mathcal{H}_{\nu_{e,t}}(a_{e,t}, c_{e,t})x_e) + n_{e,t},$$

where $a_{e,t}$ is the acquisition action, $c_{e,t}$ is the calibrated parameter vector, $\mathcal{H}_{\nu_{e,t}}$ is the versioned physical forward operator, $g_{e,t}$ is the detector response including conversion and clipping, $n_{e,t}$ is additive or conditionally specified noise, and $y_{e,t}$ is the raw observation in detector counts [count] or another declared sensor unit. The operator carries the conversion units required to map components of x_e to the input unit of $g_{e,t}$. Any non-additive, coherent, correlated, or signal-dependent noise is part of the likelihood rather than being hidden in $n_{e,t}$.

The submitted observation contract is

$$O_{e,t} = (y_{e,t}, a_{e,t}, \nu_{e,t}, c_{e,t}, \Sigma_{e,t}^C, M_{e,t}^{\text{sat}}, \tau_{e,t}, \mathcal{V}_{e,t}),$$

where $\Sigma_{e,t}^C$ is calibration-parameter covariance in the squared native parameter units, $M_{e,t}^{\text{sat}}$ is a binary saturation or dead-time mask, $\tau_{e,t}$ is acquisition time [s], and $\mathcal{V}_{e,t}$ is the declared validity envelope. The contract is invalid outside $\mathcal{V}_{e,t}$ until recalibrated or explicitly downgraded.

Aperture, diffraction, and recoverable modes

For wavelength λ [m] and numerical aperture NA [dimensionless], the conventional Rayleigh lateral scale for two incoherent point sources is

$$d_R = 0.61 \frac{\lambda}{\text{NA}} \quad [\text{m}].$$

d_R is a criterion under stated imaging assumptions, not a universal task-resolution value. Aperture diameter D_{ap} [m], focal length f [m], field of view Ω [sr], coherence, sampling pitch p_{samp} [m], exposure, and noise must be reported separately. A reconstructed pixel pitch below d_R does not by itself establish additional measured information.

For a linearized forward operator $H_{e,t}$ with singular-value decomposition

$$H_{e,t} = U_{e,t} \Sigma_{e,t} V_{e,t}^*,$$

$U_{e,t}$ and $V_{e,t}$ are unitary bases, $V_{e,t}^*$ is the conjugate transpose, and diagonal entry $\sigma_{e,t,j}$ of $\Sigma_{e,t}$ has the units of $H_{e,t}$. A state perturbation $v_{e,t,j}$ in the corresponding right singular direction is unidentifiable from that acquisition when $\sigma_{e,t,j} = 0$. For a preregistered tolerance ϵ_H with the same units as a singular value, define the effective measured rank

$$r_{\epsilon_H}(H_{e,t}) = \sum_j 1[\sigma_{e,t,j} > \epsilon_H] \quad [\text{mode}],$$

where $1[\cdot]$ is the indicator function and j indexes singular modes. Report the full singular spectrum or a validated task-relevant summary; r_{ϵ_H} is threshold-qualified.

For two task-distinct states x_1 and x_2 , likelihood separation is

$$D_{12} = D_{\text{KL}}(p(y \mid x_1, a, c, \nu) \parallel p(y \mid x_2, a, c, \nu)) \quad [\text{nat}],$$

where D_{KL} is Kullback--Leibler divergence in nats. The null-space honesty track treats x_1 and x_2 as measurement-indistinguishable when D_{12} lies below a preregistered discrimination threshold supported by a power calculation. A method must then return calibrated alternatives, a bound, or abstention unless it acquires additional evidence.

The Fixture F-007 analytical likelihood plot visualizes one such indistinguishable base operator and a separating added measurement. It is not an empirical performance result.

Photons, detector response, and dynamic range

For detector element i , use the photon-counting model when its assumptions hold:

$$k_i \sim \text{Poisson}(\mu_i), \quad \mu_i = \eta_i \Phi_i \tau_i + b_i,$$

where k_i is detected count [count], μ_i is expected count [count], η_i is quantum or detection efficiency [dimensionless], Φ_i is incident photon rate [photon/s], τ_i is exposure [s], and b_i is expected background plus dark count [count]. For the ideal background-free Poisson case,

$$\text{SNR}_{\text{shot}} = \sqrt{N_\gamma},$$

where N_γ is expected detected photon count [count] and the signal-to-noise ratio is dimensionless. Read noise [electron rms], fixed-pattern error, coherent receiver noise, afterpulsing, pile-up, and dead time receive explicit terms whenever present.

For full-well or count-rate limit $K_t^{\max}$ [count], a simplified clipped detector output is

$$y_i = \min(k_i, K_t^{\max}), \quad s_i = 1[k_i \geq K_t^{\max}],$$

where y_i is recorded count [count] and s_i is a dimensionless saturation indicator. The saturation fraction is

$$F_{\text{sat}} = \frac{1}{N_{\text{det}}} \sum_{i=1}^{N_{\text{det}}} s_i,$$

where N_{det} is detector-element count [element] and F_{sat} is dimensionless. Report full well, count-rate ceiling, read noise, dark signal, analog-to-digital converter range, and dead-time or pile-up policy independently; nominal bit depth is not dynamic range.

Phase ambiguity and prior-qualified reconstruction

For coherent intensity measurement,

$$y = |Ax|^2 + n,$$

where A is a declared complex-valued propagation and sampling operator, x is a complex field amplitude in a declared native unit, y is intensity or detector count in its native unit, $|\cdot|^2$ is elementwise squared magnitude, and n is measurement noise in the same unit as y . The ambiguity class is

$$\mathcal{E}(y; A) = \{x' \in \mathcal{X} : |Ax'|^2 = |Ax|^2\},$$

where $\mathcal{E}$ is a set of physically admissible fields. Global phase, translation, conjugate inversion, and geometry-specific ambiguities are scored as separate equivalence relations when applicable.

For reconstruction method m with prior $\pi_m(x)$ and likelihood $p_m(y \mid x, O)$, the posterior is

$$p_m(x \mid y, O) = \frac{p_m(y \mid x, O)\pi_m(x)}{\int_{\mathcal{X}} p_m(y \mid x', O)\pi_m(x') \mathrm{d}x'},$$

where O is the observation contract, x' is an integration variable with the same native units as x , and the posterior density carries the reciprocal units implied by the measure $\mathrm{d}x'$. Method m must label information origin as measurement, prior, calibration, or active intervention.

For hidden truth x_e and a nominal $(1 - \alpha)$ credible set $\mathcal{C}_{m,e,1-\alpha}$, empirical coverage over N independent episodes is

$$\widehat{\text{Cov}}_{m,1-\alpha} = \frac{1}{N} \sum_{e=1}^N 1[x_e \in \mathcal{C}_{m,e,1-\alpha}],$$

where α and coverage are dimensionless and N is episode count [episode]. Report coverage after source-family, texture, sparsity, positivity, motion, and noise-model shifts. Perceptual quality and truth fidelity remain separate outcomes.

Active sensing and illumination safety

Let b_t be the belief state before action a , θ the uncertain task state, d a downstream decision, $U(d, \theta)$ task utility in a declared native unit, and y the prospective observation. Expected value of information is

$$\text{EVI}(a \mid b_t) = \mathbb{E}_{y \sim p(y|a, b_t)} \left[\max_d \mathbb{E}[U(d, \theta) \mid b_t, a, y] \right] - \max_d \mathbb{E}[U(d, \theta) \mid b_t].$$

EVI has the same unit as U . Action admissibility is the componentwise condition

$$\mathbf{c}(a) = (N_\gamma, E_a, L_a, D_a, W_a, R_a) \preceq (N_\gamma^{\max}, E_a^{\max}, L_a^{\max}, D_a^{\max}, W_a^{\max}, R_a^{\max}),$$

where N_γ is incident or detected photon count [photon] as explicitly labeled, E_a is energy [J], L_a is latency [s], D_a is dose in the task-native safety unit, W_a is actuator wear [cycle], and R_a is risk on a declared scale. Superscript $\max$ denotes the preregistered ceiling in the same unit, and $\preceq$ means every component is within its ceiling. A scalarized cost may guide a policy only after the component ceilings are enforced and its weights are published.

Multiplex, fusion, and calibration crossover

For route r in regime ρ , define the protected outcome vector

$$\mathbf{Y}_{r, \rho} = (L_{\text{task}}, U_{\text{cal}}, N_\gamma, E_{\text{life}}, T_{\text{wall}}, F_{\text{sat}}, C_{\text{cross}}),$$

where L_{task} is task loss in its native unit, U_{cal} is dimensionless uncertainty-calibration error, N_γ is photon count [photon], E_{life} is lifecycle energy [J], T_{wall} is wall latency [s], F_{sat} is saturation fraction [dimensionless], and C_{cross} is dimensionless crosstalk. Direct and multiplexed routes are compared componentwise across photon flux, background, detector noise, occupancy, crosstalk, and saturation regimes; no single favorable point establishes an advantage.

For two sensor estimates $\hat{x}_1$ and $\hat{x}_2$ with errors $e_1 = \hat{x}_1 - x$ and $e_2 = \hat{x}_2 - x$, retain

$$P_{12} = \mathbb{E}[e_1 e_2^\top],$$

where P_{12} is cross-covariance in squared native state units and $e_2^\top$ is transpose. Setting $P_{12} = \mathbf{0}$ is a tested assumption, not a default. Fused estimates must report marginal covariance, cross-covariance or a justified bound, alignment error in native spatial and temporal units, and failure under common-mode perturbations.

For reference measurement r_t^{ref} and calibrated prediction $\hat{r}^{\text{ref}}(ct)$ in the same observation unit, define

$$z_t^2 = (r_t^{\text{ref}} - \hat{r}^{\text{ref}}(ct))^\top S_t^{-1} (r_t^{\text{ref}} - \hat{r}^{\text{ref}}(ct)),$$

where S_t is residual covariance in squared observation units and z_t^2 is dimensionless. The monitor declares a threshold $z_{\max}^2$, a window length w [sample], a false-alarm target [dimensionless probability], and a fallback. Detection delay is seconds [s], recovery time is seconds [s], calibration cost is samples [sample] and joules [J], and pre-detection task loss is reported in the task's native unit. Scene shift, source drift, alignment drift, detector gain, and thermal drift are separate hidden causes.

Optical, digital, and hybrid route accounting

For route r , end-to-end latency is

$$T_r = T_{\text{integrate}} + T_{\text{encode}} + T_{\text{propagate}} + T_{\text{detect}} + T_{\text{convert}} + T_{\text{transfer}} + T_{\text{digital}} + T_{\text{control}},$$

where every term is seconds [s]. Optical propagation time cannot replace T_r . The route identity fixes input origin, operator shape, batch, sparsity, effective precision, output dimension, programming frequency, operator reuse, utilization, and duty cycle.

Let the intended transform be $u = Wv$, where v is an input vector in declared native units, W is a linear operator with corresponding conversion units, and u is the desired output vector. Device d at temperature ϑ [K] and age ℓ [s] realizes

$$\hat{u}_{d,\vartheta,\ell} = (W + \Delta W_{d,\vartheta,\ell})v + \epsilon_{d,\vartheta,\ell}^{\text{analog}} + \epsilon_{d,\vartheta,\ell}^{\text{read}},$$

where $\Delta W_{d,\vartheta,\ell}$ has the units of W , ϵ^{analog} is analog transform error in output units, and ϵ^{read} is detector, conversion, and readout error in output units. Report bias, covariance, tails, effective precision [bit], and task loss separately across fan-in, depth, device, temperature, age, and workload.

For device population $\mathcal{D}$ with ND fabricated devices, fabrication yield is

$$Y_{\text{fab}} = \frac{1}{ND} \sum_{d \in \mathcal{D}} 1[d \text{ meets the preregistered envelope}],$$

where Y_{fab} is dimensionless and ND is device count [device]. The envelope includes transfer-function error, task quality, trimming time [s], tuning energy [J], steady thermal power [W], thermal crosstalk [dimensionless or a declared transfer unit], and stability over the declared interval. Failed dies remain in the denominator.

Query-registered physical compaction

Let encoder h transform raw observation record O into retained artifact $z = h(O)$ with size S_z [byte]. After the encoder is frozen, query q from registered set Q produces answer $f_q(O)$ from raw data and reconstructed answer $\hat{f}_q(z)$ from the retained artifact. Query recovery is

$$RQ(z) = \frac{1}{|Q|} \sum_{q \in Q} 1[d_q(\hat{f}_q(z), f_q(O)) \leq \epsilon_q],$$

where $|Q|$ is query count [query], d_q is error in the native unit of query q , ϵ_q is a tolerance in the same unit, and RQ is dimensionless. The confirmatory evaluator adds sealed future queries and an incident- investigation query after route and retention policies are frozen. Failure is reported as lost query classes, not only as an average score.

Lifecycle, labor, and matched budgets

For one accepted output, complete lifecycle energy is

$$E_{\text{life}} = E_{\text{source}} + E_{\text{modulate}} + E_{\text{propagate}} + E_{\text{detect}} + E_{\text{ADC}} + E_{\text{DAC}} \\ + E_{\text{control}} + E_{\text{digital}} + E_{\text{thermal}} + E_{\text{facility}} + E_{\text{calibrate}} + E_{\text{maintain}} + E_{\text{embodied}},$$

where every term is joules [J] and analog-to-digital and digital-to-analog conversion are denoted ADC and DAC. Powered propagation elements are charged to $E_{\text{propagate}}$; passive loss appears through increased source or amplifier demand. Amortized embodied energy is

$$E_{\text{embodied}} = \frac{E_{\text{fabricate}} + E_{\text{package}} + E_{\text{replace}} - E_{\text{recover}}}{N_{\text{accepted,life}}},$$

where numerator terms are joules [J] and $N_{\text{accepted,life}}$ is the accepted-output count [output] over measured or conservatively modeled service life. Negative recovery credit must be independently substantiated.

The equal-budget vector for method m is

$$\mathbf{B}_m = (N_{\text{train}}, N_{\text{scene}}, N_{\gamma}, D_{\text{dose}}, N_{\text{act}}, N_{\text{cal}}, N_{\text{fab}}, N_{\text{query}}, S_{\text{state}}, T_{\text{wall}}, T_{\text{human}}, E_{\text{life}}),$$

where N_{train} is training-example count [example], N_{scene} is physical-scene count [scene], N_{γ} is photon count [photon], D_{dose} is dose in a declared task-native unit, N_{act} is actuator-cycle count [cycle], N_{cal} is calibration-sample count [sample], N_{fab} is fabricated-device count [device], N_{query} is evaluator-query count [query], S_{state} is retained and working storage [byte], T_{wall} is wall time [s], T_{human} is labor [person-hour], and E_{life} is lifecycle energy [J]. Human design, alignment, labeling, tuning, calibration, inspection, safety review, maintenance, and incident response remain role-stratified entries.

An arm is budget-matched only when every binding component is within its preregistered tolerance or when the comparison is explicitly a Pareto frontier. Freed budget from an ablation remains unused.

Protected outcome vector and retirement estimand

Keep the following outcome families separate:

$$\mathbf{Z}_m = (Z_{\text{aperture}}, Z_{\text{photon}}, Z_{\text{phase}}, Z_{\text{prior}}, Z_{\text{drift}}, Z_{\text{range}}, Z_{\text{fusion}}, Z_{\text{transform}}, Z_{\text{conversion}}, Z_{\text{analog}}, Z_{\text{fabrication}}, Z_{\text{thermal}}, Z_{\text{safety}}, Z_{\text{lifecycle}}),$$

where the components respectively contain aperture/diffraction results, photon/shot-noise results, phase-ambiguity results, prior-mismatch results, calibration/drift results, saturation/dynamic-range results, fusion-covariance results, optical-transform results, conversion/readout results, analog-error results, fabrication results, thermal-work results, active-illumination safety results, and full lifecycle energy and labor results. Each Z is a structured record with the native units defined above; the vector is not scalarized for acceptance.

For protected component j , the paired effect of mechanism k is

$$\Delta_{k,j} = Z_j(m_{\text{full}}) - Z_j(m_{-k}),$$

where m_{-k} removes only mechanism k and receives no replacement resource. Report $\Delta_{k,j}$ in the native unit of outcome j with a 95% uncertainty interval over independent scene, operator, regime, device, site, and seed strata.

Retire the residual when an equal-budget mature-null composition matches the protected vector; when a gain depends on hidden prior, operator, calibration, or query leakage; when it fails held-out devices or regimes; or when no preregistered ablation isolates value beyond the registered inverse, uncertainty, sensing, control, acceleration, optical, hybrid-design, and calibration baselines. The editable source for the system diagram is operator-qualified-physical-inference.mmd.

Opportunity- and history-qualified adaptive action

This note defines the comparison boundary for [Fixture F-003](#). The fixture uses comparative-cognition tasks as diagnostic interventions, not as species rankings. Its central requirement is simple: performance can be compared only after the information, action, learning, and reward opportunities that produced it are explicit.

Episode contract

For episode e , record

$$\Omega_e = (X_e, S_e, A_e, H_e, R_e, C_e, \tau_e, U_e),$$

where X_e is task and apparatus state, S_e is the physically available sensory channel, A_e is the feasible realized-action set, H_e is training, rearing, prior-task, and social-exposure history, R_e is the reward, punishment, deprivation, and stopping rule, C_e is the intervention and control set, τ_e is the relevant time in seconds, and U_e is the sampling unit such as episode, agent, dyad, or population. All fields except τ_e are typed records or sets rather than unit-bearing scalars.

The learner receives the causally available prefix

$$\mathcal{I}_e(t) = \{s \in S_e : t_s^{\text{recv}} \leq t\} \cup H_e^{\leq t} \cup R_e^{\leq t},$$

where t and receipt time t_s^{recv} are seconds. Handler cues, apparatus sounds, demonstrator traces, residual odor, simulator metadata, and training-phase identifiers belong in S_e when they are actually available. Protocol intention is not a measurement of the input channel.

For requested action a and plant version v , feasibility is

$$F_v(a, \Omega_e) \in \{0, 1\},$$

and realized action follows

$$\tilde{a}_t \sim p_v(\tilde{a}_t \mid a_t, A_e, X_e, H_e).$$

F_v is dimensionless. Requested and realized actions retain their native units, such as metres, radians, newtons, newton-metres, or a discrete tool identifier. A nominally identical task is not matched when one plant cannot sense while holding the required tool or cannot execute the required motion.

Opportunity-qualified performance

For method m , success event $Y_e \in \{0, 1\}$, and declared opportunity stratum ω , define

$$Q_m(\omega) = \Pr(Y_e = 1 \mid do(m), \Omega_e \in \omega).$$

Q_m is dimensionless. The paired contrast against baseline m_0 is

$$\Delta_{m,m_0}(\omega) = Q_m(\omega) - Q_{m_0}(\omega).$$

Report Δ by morphology, actual channel, history, reward schedule, demonstrator exposure, site, and intervention family. A pooled contrast is admissible only after heterogeneity is shown rather than assumed away.

Use a transport contrast to test whether the learned relation survives a controlled change g :

$$T_g(m) = Q_m(g(\omega)) - Q_m(\omega).$$

T_g is dimensionless. Useful changes include material substitution, geometry change, causal inversion, perceptual-cue reversal, morphology or tool swap, history swap, demonstrator removal, partner turnover, and delayed future use. The sign is interpreted relative to a preregistered invariance: some changes should preserve performance; inversions should reverse the selected action.

Transfer lattice

Do not compress transfer into one difficulty axis. For task family k , retain

$$\mathbf{T}_k = (T_{\text{surface}}, T_{\text{material}}, T_{\text{relation}}, T_{\text{history}}, T_{\text{plant}}, T_{\text{social}}, T_{\text{delay}}).$$

Every component is a dimensionless paired effect. Surface and material changes test perceptual or affordance generalization; relation changes test functional sensitivity; history and plant changes test dependence on acquisition and embodiment; social changes identify copied content; and delay changes test the lifetime of retained state. The vector prevents terminal success on one familiar apparatus from standing in for causal, prospective, or cross-plant transfer.

First-trial transfer after the held change is reported separately:

$$T_g^{(1)}(m) = Y_{m,g,1} - Y_{m,g_0,1}.$$

Later trial curves estimate adaptation, not prior transfer. Both are retained.

Social acquisition decomposition

A demonstration d is represented as

$$d = (f, y, \gamma, \iota, \rho, p),$$

where f is action form, y is end state, γ is trajectory, ι is demonstrator identity and reliability, ρ is the demonstrated functional relation, and p is provenance. These are typed variables.

For component $j \in \{f, y, \gamma, \iota, \rho\}$, its causal uptake effect is

$$I_j = \mathbb{E}[L \mid \text{do}(d_j = d'_j), d_{-j}] - \mathbb{E}[L \mid \text{do}(d_j = d_j^0), d_{-j}],$$

where L is a declared dimensionless learner outcome and d_{-j} holds the other components fixed. Ghost, result-only, novel-action, inefficient-action, and causal-relevance controls approximate these interventions. Matching the end state does not establish copying of action form.

Prospective and event memory

For stored event or resource i ,

$$e_i = (w_i, \ell_i, t_i, q_i, c_i, p_i),$$

where w_i is content, ℓ_i location, t_i time in seconds from a declared origin, q_i quality or perishability state, c_i social or task context, and p_i provenance. At future time t , a conventional reservation null uses

$$V_i(t) = p_i^{\text{avail}}(t \mid e_i) r_i(t) - k_i(t),$$

where availability probability is dimensionless and reward r_i and retrieval/carrying cost k_i use the same declared utility unit. A proposed prospective trace earns credit only beyond value tables, successor representations, POMDP planning, and retrieval with equal state and rollout budgets.

Bound-event memory is tested against factorized semantic and spatial state. For query q over unique event e_i , report exact-answer rate, calibration, update propagation, retrieval latency in seconds, and memory bytes. Subjective recollection is not an observable in this contract.

Costed uncertainty control

For optional observation or query z with cost c_z , the ordinary value-of-information null is

$$\text{VOI}(z \mid b) = \mathbb{E}_z \left[\max_{\theta} \mathbb{E}[U(a, \theta) \mid b, z] \right] - \max_{\theta} \mathbb{E}[U(a, \theta) \mid b] - c_z,$$

where belief b and latent state θ are dimensionless, while utility U and c_z use the same declared unit. The policy acquires z only when VOI is positive. Risk–coverage–cost surfaces must also include calibrated confidence, ensembles, conformal/selective prediction, learned difficulty cues, and response-strength policies.

Local, central, and mechanical contribution

For a compliant plant with n local segments, let the central policy send message gt and segment i receive local observation oi, t . The realized control is

$$u_{i,t} = \pi_i(oi, t, gt, h_{i,t}; v),$$

where local state $h_{i,t}$ is dimensionless, plant version is v , and control $u_{i,t}$ retains its native actuator unit. Compare this with a centralized policy receiving the same causally available observations and with a passive- mechanics arm receiving no learned local state.

Communication load over horizon $[0, T]$ is

$$B_{\text{comm}}(T) = \sum_{m: t_m \leq T} \text{bytes}(m) \quad [\text{bytes}],$$

and total control energy is

$$E_{\text{control}} = E_{\text{sense}} + E_{\text{central}} + E_{\text{local}} + E_{\text{comm}} + E_{\text{act}} + E_{\text{adapt}} + E_{\text{recover}} \quad [\text{J}].$$

Peripheral credit requires a quality, recovery, latency, or energy gain after passive compliance, actuator work, sensor power, communication, and local hardware are charged. Anatomical distribution alone supplies no credit.

Outcome and lifecycle boundary

Keep the confirmatory result as a vector:

$$\mathbf{Y} = (Q, T^{(1)}, \mathbf{T}, N_{\text{unsafe}}, N_{\text{attempt}}, L_{50}, L_{95}, B_{\text{memory}}, B_{\text{comm}}, H_{\text{human}}, E_{\text{life}}).$$

Q , $T^{(1)}$, and $\mathbf{T}$ are dimensionless; unsafe events and attempts are counts; latencies L_{50} and L_{95} are seconds; memory and communication are bytes; human effort H_{human} is person-hours; and lifecycle energy E_{life} is joules.

For method m over all development and confirmatory work,

$$E_{\text{life}}^{(m)} = E_{\text{train}} + E_{\text{demonstrate}} + E_{\text{search}} + E_{\text{infer}} + E_{\text{interact}} + E_{\text{communicate}} + E_{\text{store}} + E_{\text{adapt}} + E_{\text{recover}} \quad [\text{J}].$$

The fixture accepts the composed residual only if it improves a preregistered subset of literal outcomes without violating non-inferiority margins on the others, survives at least two task families, plant geometries, histories, model families, and hardware classes, and beats the complete conventional null stack at the same episodes, interventions, search, storage, communication, human effort, and energy. Otherwise retain the observation contract and retire the architectural story.

Editable system diagram: [opportunity-history-qualified-action.mmd](#).

Cohort-aware population and historical observation

This note prevents aggregate counts and retrospective stories from silently standing in for lifecycle state, causality, or prospective prediction. It is a mathematical companion to the quantitative-history and demography audit and Candidate 014.

Stock and flow

For an accounting interval of duration Δt seconds,

$$N_{t+\Delta t} = N_t + B_t - D_t + I_t - O_t,$$

where N_t is the count of qualified active units at time t , and B_t , D_t , I_t , and O_t are replication/entry, retirement/death, immigration, and outmigration counts during the same interval. A rate such as $b_t = B_t/(N_t\Delta t)$ has units s^{-1} . The identity detects inconsistent bookkeeping; it neither explains the flows nor predicts the next interval.

Cohort-component transition

Let $\mathbf{n}_t$ be a vector of unit counts indexed by lifecycle stage, role, version cohort, or another declared partition. Then

$$\mathbf{n}_{t+\Delta t} = \mathbf{L}_t \mathbf{n}_t + \mathbf{m}_t,$$

where $\mathbf{L}_t$ is a transition matrix and $\mathbf{m}_t$ is the net migration vector in unit counts.

Survival/transition entries in $\mathbf{L}_t$ are dimensionless probabilities per interval; replication entries are new units per source unit per interval. A constant matrix is a scenario assumption, not evidence that transition rates remain stationary.

Equal totals $\mathbf{1}^\top \mathbf{n}_t$ can conceal different future paths because stage composition changes exposure, replication, failure, maintenance, and retirement. Every population-level result therefore retains the vector or a registered sufficient aggregation.

Age, period, and cohort are not freely separable

If lifecycle age a , observation period p , and entry cohort c satisfy $c = p - a$, then

$$g(\mathbb{E}[Y_{a,p}]) = \mu + \alpha_a + \beta_p + \gamma_c$$

is not uniquely identified in its unrestricted linear components. Y uses a declared task unit and g is a specified link. More samples or less noise do not remove the exact dependency. Constraints, priors, curvature, external variation, or mechanistic structure must be declared because they determine part of the decomposition.

Adoption curves do not identify influence

For adoption share $F(t) \in [0, 1]$,

$$\frac{dF(t)}{dt} = (p + qF(t))(1 - F(t)),$$

where p and q have units s^{-1} . A close fit is compatible with multiple generative processes: independent exposure, common broadcast, command, homophilous selection, network influence, or mixtures. Identifying influence requires intervention or additional assumptions, not curve shape alone.

Selected observation

For latent event count N_{latent} and a simplified chain,

$$\mathbb{E}[N_{\text{observed}}] = N_{\text{latent}} p_{\text{produce}} p_{\text{survive}} p_{\text{discover}} p_{\text{retain}} p_{\text{code}},$$

where every p is a dimensionless conditional probability under a declared dependency order. The product is diagnostic, not an independence claim. If missingness depends on an unobserved value after conditioning on available data, unrestricted recovery is impossible without external data or additional selection assumptions.

Collapse and recovery remain vectors

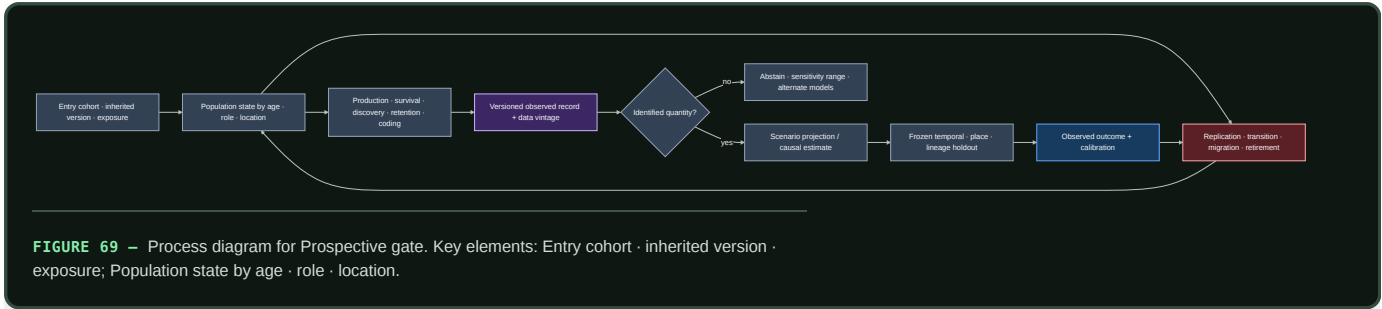
A population or institution state is reported as

$$\mathbf{r}_t = (N_t, Q_t, A_t, S_t, G_t, H_t),$$

where N_t is qualified unit count, Q_t task/service quality, A_t effective authority coverage, S_t network/service connectivity, G_t governance or maintenance capacity, and H_t reserve/headroom. Every component uses its native declared unit. No scalar “collapse” or “recovery” score is formed without explicit authorized weights and sensitivity analysis.

Prospective gate

Retrospective explanation is separated from prediction by freezing model, features, coding, hyperparameters, data vintage, and evaluation before the held-out period, place, lineage, or regime becomes available. Rolling-origin, geographic, lineage, and live holdouts test different forms of transfer; random row splits are insufficient when adjacent rows share history.



Editable source: [cohort-observation-contract.mmd](#).

Predictive temporal co-adaptation without a privileged clock

This note defines the time, observation, coordination, and cost boundary for the `shared-clock-free` benchmark. Music supplies the test regime; the measured problem is distributed predictive control under expressive nonstationarity, delay, role change, and partner turnover.

Local clocks and evaluator time

Agent i observes its local clock

$$\tau_i(t) = a_i t + b_i + \epsilon_i(t) \quad [\text{s}],$$

where t is hidden evaluator time in seconds, a_i is dimensionless clock rate, b_i is offset in seconds, and ϵ_i is local clock error in seconds. Agents may not read t , a_i , b_i , or another agent's clock. The evaluator uses t only to score pairwise events and injected disturbances.

For event n sent by partner j and observed by agent i ,

$$\tau_{ij,n}^{\text{obs}} = \tau_i(t_{j,n} + d_{ij,n}) + \nu_{ij,n},$$

where event time $t_{j,n}$ and channel delay $d_{ij,n}$ are seconds and observation error $\nu_{ij,n}$ is seconds. Delay may be asymmetric, time-varying, and censored by loss. A shared beat label, synchronized timestamp, or evaluator clock in the policy input invalidates the shared-clock-free arm.

Phase and tempo state

Agent i maintains a local phase estimate $\hat{\phi}_i \in [-\pi, \pi)$ radians, period estimate $\hat{T}_i$ in seconds, uncertainty $\sigma_{\phi,i}$ in radians, and drift estimate $\dot{T}_i$ in seconds per second. For an observed partner phase $\phi_{j,n}$,

$$e_{ij,n} = \text{wrap}_{[-\pi,\pi)}(\phi_{j,n} - \hat{\phi}_i(\tau_{ij,n}^{\text{obs}})) \quad [\text{rad}].$$

A minimal correction null is

$$\hat{t}_{i,n+1} = \hat{t}_{i,n} + \hat{T}_{i,n} + \alpha_i a_{ij,n},$$

where predicted onset $\hat{t}$, period $\hat{T}$, and onset asynchrony $a_{ij,n} = t_{j,n} - \hat{t}_{i,n}$ are seconds, and gain α_i is dimensionless. Tempo adaptation requires a separately identifiable update; folding timing and tempo error into one unconstrained latent state weakens the diagnosis.

Pairwise timing outcomes

For paired events n from agents i and j , evaluator asynchrony is

$$A_{ij,n} = t_{i,n} - t_{j,n} \quad [\text{s}].$$

Report the signed distribution, median absolute asynchrony, $p95$ absolute asynchrony, circular phase error, and missed or duplicated event count. Similar mean tempo does not establish synchronization.

After perturbation at t_p , recovery time is

$$T_{ij}^{\text{sync}} = \inf \{t \geq t_p : \text{median}_{u \in [t, t+\Delta]} |A_{ij}(u)| \leq A^{\text{max}} \wedge V_{ij}(t : t + \Delta) \leq V^{\text{max}}\} - t_p,$$

where Δ and T^{sync} are seconds, A^{max} is seconds, and V_{ij} is a declared variance or instability measure with a matching frozen threshold. Censored failures remain failures rather than disappearing from the mean.

Phrase and role state

Timing alone can be excellent while the wrong phrase, motif, answer, or role is executed. Let q_n be a typed phrase act with boundary, motif relation, literal constraint, role, and addressee. The partner model is

$$p_i(q_{j,n+1}, r_{j,n+1} \mid h_{i,n}, m_{ij}, c_n),$$

where history $h_{i,n}$, partner memory m_{ij} , context c_n , phrase act q , and role r are typed variables. Score held-out log loss in bits per act, calibration, literal constraint satisfaction, phrase-response accuracy, and role violations separately.

Familiar-partner gain is

$$G_{\text{partner}} = Q_{\text{familiar}} - Q_{\text{unseen}},$$

where Q uses the same dimensionless task-quality measure and task set. A positive value may reflect replay or overfitting; it earns co-adaptation credit only when the policy also transfers to new motifs and recovers after partner or role change.

Multi-objective outcome

Keep the primary result as a vector:

$$\mathbf{Y} = (Q_{\text{literal}}, |A|_{50}, |A|_{95}, T_{\text{sync}}, E_{\text{phrase}}, N_{\text{role violation}}, B_{\text{message}}, L_{\text{decision}}, E_{\text{life}}).$$

Q_{literal} is dimensionless; asynchronies and recovery are seconds; phrase error E_{phrase} is a dimensionless error rate; N is a count; message traffic B is bytes; decision latency L is seconds; and lifecycle energy E_{life} is joules. Blinded human judgments of coherence or expressivity are additional outcomes, never substitutes for literal success or timing.

Lifecycle and equal-budget boundary

For horizon $[0, T]$,

$$E_{\text{life}} = E_{\text{sense}} + E_{\text{estimate}} + E_{\text{predict}} + E_{\text{retrieve}} + E_{\text{message}} + E_{\text{act}} + E_{\text{adapt}} + E_{\text{rehearse}} + E_{\text{recover}},$$

with all terms in joules at one boundary. Also report parameter bytes, partner- memory bytes, messages and bytes per event, evaluator or human queries, rehearsal events, and wall time. A centralized conductor pays its synchronization traffic and infrastructure; a local method pays identification and adaptation.

Equal-budget arms receive the same event stream, sensor uncertainty, action interface, parameter and state ceiling, training examples, rehearsal, hyperparameter trials, update rate, message budget, latency budget, and energy boundary. The hidden evaluator trace is paired across arms.

Rejection conditions

Reject the composed residual when any of the following holds:

1. a phase-locked loop, Kalman/state-space estimator, or MPC arm matches timing and recovery;
2. retrieval plus explicit role/protocol state matches phrase response and partner turnover;
3. the gain vanishes on new partners, motifs, tempi, or delay processes;
4. a policy observes the privileged clock or hidden perturbation state;
5. familiar-partner gain is replay without adaptation to a new phrase;
6. timing improves while literal constraints or role safety worsen; or
7. adaptation, rehearsal, communication, evaluator work, or recovery energy is omitted.

Editable system diagram: [shared-clock-free-coadaptation.mmd](#).

Proof-discovery and verification contract

This note defines the measurement boundary for Fixture F-004. It operationalizes the versioned **propose–challenge–decompose–prove–check–publish–invalidate** lifecycle derived from the mathematical-practice audit. It is a shared benchmark contract for Candidates 004, 009, 010, 011, 014, 017, and 019.

Immutable identity envelope

Every episode starts with a sealed identity envelope

$$I_e = (h_p, h_d, h_\ell, h_a, h_L, h_C, h_K, h_s),$$

where e identifies the episode; h_p is the cryptographic hash of the exact problem statement; h_d hashes its definitions and encodings; h_ℓ hashes the declared logic and semantics; h_a hashes the admitted axioms; h_L hashes the complete accessible library version; h_C hashes the accessible corpus, retrieval index, and cutoff; h_K hashes the checker, certificate calculus, and preprocessing reconstruction; and h_s hashes the task-family generator and paired random seed. Every h value is an immutable byte string. Hash equality establishes identity of registered bytes, not adequacy of the formalization or soundness of the implementation.

The accessible information record is

$$\Lambda_e = (A_e^{\text{train}}, A_e^{\text{prompt}}, A_e^{\text{retrieve}}, A_e^{\text{proof}}, A_e^{\text{eval}}, A_e^{\text{human}}),$$

where each A_e^x is a set of immutable artifact identifiers available through channel x . The channels respectively cover training, prompt/context, retrieval, proof or tactic traces, evaluation feedback, and human hints or formalization. Availability is recorded independently of whether an artifact is retrieved or used.

Versioned lifecycle state

At event time t in seconds from episode start, define

$$S_t = (G_t, Q_t, L_t, X_t, D_t, C_t, R_t, \sigma_t),$$

where G_t is the set of open typed goals, Q_t is the queue of proposed operations, L_t is the accessible immutable library view, X_t is the set of examples and counterexamples, D_t is the proof-and-dependency directed acyclic graph (DAG), C_t is the set of proof, refutation, model, or test certificates, R_t is the append-only event record, and σ_t is the typed publication state. The finite state space is

$$\Sigma = \{\text{proposed, tested, proved, refuted, unknown, disputed, retracted}\}.$$

An event at may transform the state only when its typed preconditions hold:

$$S_{t+1} = T(S_t, at), \quad at \in \mathcal{A}(S_t),$$

where T is the versioned transition function and $\mathcal{A}(S_t)$ is the set of admissible events in state S_t

. **Tested** records finite support only; **proved** records acceptance of a derivation relative to I_e ; **refuted** requires an admissible counterexample or checked refutation; **unknown** preserves timeouts, unsupported theories, and incomplete searches; **disputed** records an unresolved challenge to formalization, assumptions, checking, or significance; and **retracted** records invalidation or supersession without deleting history.

A published version is

$$v_i = (h_i, P_i, o_i, u_i, t_i, \sigma_i, E_i, B_i),$$

where h_i is the version hash, P_i is the set of parent-version hashes, o_i is the typed transformation, u_i is the actor or process identifier, t_i is the event time in seconds, $\sigma_i \in \Sigma$ is the state, E_i is the evidence-and-certificate set, and B_i is the measured cost record. A new version never overwrites its parents.

Proposals, tests, and counterexamples

Let p_i be candidate proposition i , θ_i its explicit definitions and parameters, O_i its set of proof obligations, H_i its declared source and transformation ancestry, and q_i its proposal method. The proposal record is

$$z_i = (p_i, \theta_i, O_i, H_i, q_i, I_e).$$

For a universally quantified candidate $p_i \equiv \forall x \in \mathcal{X}_i : P_i(x)$, an exact counterexample is an element $x^* \in \mathcal{X}_i$ such that

$$P_i(x^*) = \text{false}.$$

$\mathcal{X}_i$ is the declared admissible domain and P_i is the encoded Boolean predicate. The counterexample record must prove or independently check both domain membership and predicate failure. Its size $s(x^*) \in \mathbb{R}_{\geq 0}$ is measured under a preregistered task-native order; smallest-counterexample claims compare $s(x^*)$, not discovery time alone.

For a finite test multiset $U_i = \{x_1, \dots, x_n\}$, empirical support is

$$\hat{q}_i = \frac{1}{n} \sum_{j=1}^n \mathbf{1}[P_i(x_j) = \text{true}],$$

where n is the test count, $\mathbf{1}[\cdot]$ is an indicator, and $\hat{q}_i$ is dimensionless. The record also names the generator, distribution or enumeration boundary, random seed, arithmetic precision in bits, and error interval in the predicate's native unit. Even $\hat{q}_i = 1$ leaves the state **tested** unless a checked exhaustive reduction or proof closes the universal obligation.

Abstraction obligations

Let $\mathcal{X}$ be a concrete domain, $\mathcal{A}$ an abstract domain, $\alpha : \mathcal{X} \rightarrow \mathcal{A}$ an abstraction map, and $\gamma : \mathcal{A} \rightarrow 2^{\mathcal{X}}$ a concretization map. For a concrete property $P : \mathcal{X} \rightarrow \{\text{true}, \text{false}\}$ and abstract property $P^\sharp : \mathcal{A} \rightarrow \{\text{true}, \text{false}\}$, using the abstraction to prove P requires the soundness obligation

$$\forall a \in \mathcal{A} : P^\sharp(a) \Rightarrow \forall x \in \gamma(a) : P(x).$$

When a claimed transfer also requires preservation of operation $f : \mathcal{X} \rightarrow \mathcal{X}$, declare an abstract operation $f^\sharp : \mathcal{A} \rightarrow \mathcal{A}$ and check

$$\alpha(f(x)) = f^\sharp(\alpha(x))$$

for the stated subset of $x \in \mathcal{X}$. The equality is an exact semantic claim unless a typed approximation relation and tolerance are registered. Excluded cases, failed concretizations, reconstruction loss, and the human time used to choose α are part of the result.

Decomposition and proof-DAG reconstruction

For target goal g_0 , let the proof DAG be

$$D_\pi = (V_\pi, E_\pi, r, \kappa),$$

where V_π is a set of goal, lemma, definition, axiom, or certificate nodes; $E_\pi \subseteq V_\pi \times V_\pi$ contains directed dependency edges from a conclusion to each required premise; $r \in V_\pi$ is the root node encoding g_0 ; and κ_v is the typed local reconstruction function stored for node v . For each non-leaf node v with dependency set $\text{dep}(v)$, reconstruction requires

$$\kappa_v(\{\pi_u : u \in \text{dep}(v)\}) = \pi_v,$$

where π_u is a checked artifact proving node u and π_v is the constructed artifact for node v . Leaves must resolve to hashed axioms, definitions, admitted theorems, exact decision procedures, or independently checked certificates in I_E .

Let $\text{topo}(D_\pi)$ be a topological ordering from leaves to root. Full-DAG reconstruction succeeds when every κ_v executes in that order and the independent checker accepts the resulting root artifact against $h_{\mathcal{P}}$ and $h_{\mathcal{L}}$. A list of plausible lemmas, a cyclic graph, or individually checked children without a parent reconstruction function does not close g_0 .

Proof-dependency economy is reported, not assumed. If N_{all} is the number of accessible library nodes and N_{cl} is the number in the root's transitive dependency closure, then

$$\rho_{\text{dep}} = \frac{N_{\text{cl}}}{N_{\text{all}}}$$

is a dimensionless dependency ratio. Smaller is useful only if reconstruction, robustness, and transfer remain non-inferior.

Generator, certificate, and checker separation

For instance bytes F , generator G , claimed verdict y , certificate c , preprocessing trace r_F , and independent checker K , acceptance is

$$\text{Accept}(F, y, c) = \mathbf{1}[h(F) = h_p] \mathbf{1}[K(h(F), y, c, r_F, h_L, h_K) = \text{accept}],$$

where $h(\cdot)$ is the registered hash function and both indicator factors are dimensionless.

Generator G cannot set the checker's verdict. The identities, code lineage, parsers, libraries, axioms, hardware, and people shared by G and K are published as a shared-trust-root set W_{GK} ; process separation is not described as independent when W_{GK} contains the relevant possible fault.

For accepted artifact i , record certificate size b_i^{cert} in bytes, generation time t_i^{gen} in seconds, checking time t_i^{check} in seconds, checker peak memory m_i^{check} in bytes, and certificate lifecycle energy E_i^{cert} in joules. An unsupported solver exit code or a certificate that does not bind the exact instance and preprocessing trace cannot produce `proved` or `refuted`.

Leakage-safe task partitions

Let $\mathcal{U}$ be the universe of theorem, proof, definition, example, counterexample, generated sibling, and source artifacts. Define a symmetric ancestor-related relation $\sim_A$ over $\mathcal{U}$ that groups exact duplicates, renamings, restatements, specializations, isomorphic generated siblings, proof ancestors, and source-equivalent formalizations under the frozen detection protocol. Let $[u]_A$ denote the group containing artifact u .

For training artifacts $\mathcal{T} \subset \mathcal{U}$ and confirmatory artifacts $\mathcal{H} \subset \mathcal{U}$, the dependency-safe split condition is

$$\{[u]_A : u \in \mathcal{T}\} \cap \{[u]_A : u \in \mathcal{H}\} = \emptyset.$$

Additionally, if $\text{clD}(u)$ is the transitive proof and library dependency closure of u , require

$$\left(\bigcup_{u \in \mathcal{H}} \text{clD}(u) \right) \cap \mathcal{T} = \emptyset$$

unless the overlapping artifacts are explicitly declared as common premises available to every arm. Public benchmark queries, human corrections, and model updates after the split freeze are added to Δ_e and disqualify the affected hidden family from confirmatory use.

Publication and reverse-dependency invalidation

Let the release dependency graph be $D_v = (N_v, E_v)$, with edge $(a, b) \in E_v$ meaning artifact a depends on artifact b . If changed or invalid artifact set $J \subseteq N_v$ is detected, its reverse-dependency invalidation set is

$$\text{Inv}(J) = J \cup \{a \in N_v : \exists b \in J \text{ with a path } a \rightsquigarrow b\}.$$

Every artifact in $\text{Inv}(J)$ is quarantined from `proved` release status until its exact version is rechecked or rebuilt against a replacement. The prior record remains addressable and becomes `retracted` when its published acceptance claim no longer holds; a challenge to formalization or significance may instead enter `disputed` while the checked derivation remains recorded.

For planted invalidation set J^* and submitted set $\hat{J}$, invalidation precision and recall are

$$P_{\text{inv}} = \frac{|\hat{J} \cap J^*|}{|\hat{J}|}, \quad R_{\text{inv}} = \frac{|\hat{J} \cap J^*|}{|J^*|}.$$

Both are dimensionless; an empty submitted set receives zero precision and recall. Also report time to quarantine in seconds, stale artifact-hours before quarantine, rebuild time in seconds, repair human-hours, and recurrent bad acceptances after release.

Equal lifecycle budget

For method m , the binding resource vector is

$$\mathbf{B}_m = (N_{\text{train}}, N_{\text{ret}}, N_{\text{prop}}, N_{\text{test}}, N_{\text{solve}}, N_{\text{check}}, N_{\text{node}}, B_{\text{state}}, B_{\text{cert}}, t_{\text{wall}}, t_{\text{cpu}}, h_{\text{human}}, E_{\text{life}}, M_{\text{peak}}).$$

The first seven terms count training items, retrieval calls, proposals, tested instances, solver calls, checker calls, and search nodes. B_{state} and B_{cert} are retained-state and certificate bytes; t_{wall} and t_{cpu} are seconds; h_{human} is person-hours; E_{life} is joules at the declared measurement boundary; and M_{peak} is peak working memory in bytes.

Lifecycle energy is

$$E_{\text{life}} = E_{\text{train}} + E_{\text{index}} + E_{\text{retrieve}} + E_{\text{propose}} + E_{\text{challenge}} + E_{\text{prove}} + E_{\text{check}} + E_{\text{publish}} + E_{\text{retain}} + E_{\text{invalidate}} + E_{\text{repair}},$$

with every term measured in joules across one declared hardware and facility boundary. Human formalization, steering, review, and repair remain person-hours and are not silently converted to joules. A method that exceeds any binding ceiling is infeasible for that paired instance; otherwise report a preregistered Pareto frontier rather than post-hoc cost normalization.

Outcome vector and component effects

For method m , report

$$\mathbf{Y}_m = (N_{\text{proved}}, N_{\text{refuted}}, N_{\text{tested}}, N_{\text{unknown}}, N_{\text{bad accept}}, R_{\text{recon}}, P_{\text{inv}}, R_{\text{inv}}, t_{\text{first}}, b_{\text{median}}, h_{\text{human}}, E_{\text{life}}),$$

where the first five terms are artifact counts; R_{recon} is the dimensionless proportion of published proof DAGs reconstructed and rechecked from their retained dependencies; P_{inv} and R_{inv} are defined above; t_{first} is seconds to the first checked result; b_{median} is median certificate size in bytes; h_{human} is person-hours; and E_{life} is joules. Results are stratified by task family, state, and hidden regime rather than collapsed into a universal reasoning score.

Checked-result energy efficiency may be reported as

$$\eta_E = \frac{N_{\text{proved}} + N_{\text{refuted}}}{E_{\text{life}}} \quad [\text{checked result/J}],$$

but only beside the full outcome vector, because a system can inflate η_E with trivial tasks or low coverage. For component c and outcome Y_j , the paired ablation effect is

$$\Delta_{c,j} = Y_j(m_{\text{full}}) - Y_j(m_{-c}),$$

where $m - c$ removes component c without reallocating its unused budget. Report paired 95% uncertainty intervals over frozen problem-family and random- seed strata.

Contract retirement

The composed residual is retired when the strongest complete ordinary stack matches its checked closure, bad-acceptance rate, hidden-family transfer, proof-DAG reconstruction, invalidation quality, and lifecycle resource vector; when any gain disappears under ancestor- and dependency-safe splits; when counterexamples, abstraction obligations, certificates, or reverse-dependency events cannot be reconstructed independently; or when omitted formalization, review, checking, storage, repair, human-hour, or joule costs explain the gain.

Editable lifecycle diagram: [versioned-proof-discovery-lifecycle.mmd](#).

Query-registered semantic preservation

This note refines [Candidate 017](#) with the interpreter, representation-dependency, vocabulary, authenticity, and access boundaries established by library and archival science. The contract does not promise to preserve all future meaning. It makes a finite claim that can fail.

Preservation package

For version v , define

$$P_v = (B_v, M_v, F_v, S_v, V_v, D_v, A_v, X_v),$$

where:

- B_v is payload bytes;
- M_v is metadata, evidence, provenance, and fixity records in bytes;
- F_v is format and rendering dependencies;
- S_v is schema and type constraints;
- V_v is vocabulary, authority, identity, and mapping versions;
- D_v is the designated interpreter community and its assumed capabilities;
- A_v is authenticity/custody evidence and authorized access state; and
- X_v is executable dependencies, tests, emulators, or migration tooling.

The tuple is typed. Fixity of B_v does not establish truth, authenticity, interpretability, or authorized accessibility. Provenance in M_v records lineage; it does not validate assertions.

Registered query contract

For query family $q \in \mathcal{Q}$, declare native output space $\mathcal{Y}_q$, distance d_q , tolerance ϵ_q , evidence obligations L_q , supported community D_q , and decision horizon T_q . For source package P_0 and a candidate transformed package P_v , require

$$\Pr_{q \sim \mathcal{D}_Q} [d_q(q(P_0), q(P_v)) > \epsilon_q] \leq \delta_q,$$

where d_q and ϵ_q share the query's native unit and δ_q is a dimensionless failure probability. The evaluation distribution $\mathcal{D}_Q$ has frozen development, held-out in-family, and adversarial migration splits.

Evidence reachability is

$$R_L(P_v) = \frac{\sum_{q \in \mathcal{Q}} |L_q \cap \widehat{L}_q(P_v)|}{\sum_{q \in \mathcal{Q}} |L_q|},$$

where L_q is the registered evidence set and $\widehat{L}_q(P_v)$ is the set recoverable with valid lineage from P_v . R_L is dimensionless. It does not score whether the evidence supports a claim; that requires a separate inference and decision contract.

Interpreter and dependency validity

Let $c(D, t)$ be a versioned capability vector for the interpreter community at time t : supported languages, schemas, units, software, cryptographic algorithms, domain conventions, and required practiced procedures. Let $r(P_v)$ be the corresponding requirements. Interpretability is

$$I(P_v, D, t) = \mathbf{1}[c(D, t) \succeq r(P_v)],$$

a dimensionless predicate under a declared partial order. “Human readable” or “standard format” is insufficient unless the supported community and dependencies are named.

For dependency graph $G_v = (N_v, E_v)$ and root package nodes R_v , required closure is

$$\text{cl}(R_v) = \{n \in N_v : \exists r \in R_v \text{ with a required path } r \rightsquigarrow n\}.$$

Missing-dependency rate is

$$m_v = \frac{|\{n \in \text{cl}(R_v) : n \text{ unavailable or invalid}\}|}{|\text{cl}(R_v)|},$$

which is dimensionless. Criticality weights may be reported separately, but a weighted mean cannot hide the loss of a dependency required by every query.

Migration and vocabulary drift

For migration $T_v : P_{v-1} \rightarrow P_v$, maintain a manifest containing source/target versions, transformed and retained fields, known loss, dependencies, tests, reviewer/authority, and rollback or source-recovery path. Query regression is

$$\Delta_q^{(v)} = d_q(q(P_{v-1}), q(P_v)),$$

in the native query unit. Vocabulary or authority mappings are separate versioned relations

$$R_v \subseteq E_{v-1} \times E_v \times \{\text{same, broader, narrower, split, merge, contested}\}.$$

A merge cannot silently transfer all evidence from both prior entities. Query tests must include namesakes, renames, splits, merges, multilingual labels, contested mappings, and temporal concept change.

Availability decomposition

For query q , time t , and community D , a useful decomposition is

$$U(q, t, D) = p_{\text{bits}} p_{\text{Prender}} p_{\text{semantic}} p_{\text{authentic}} p_{\text{authorized}} V(q, D) - C(q, t, D).$$

The five p terms are dimensionless conditional probabilities; V and C share one declared decision unit. The product is a checklist unless dependence among failures is explicitly modelled. Report each component and joint failure cases; do not claim independence by notation.

Registered and unregistered use

The registered contract permits direct success/failure decisions. New unregistered query q' receives one of three typed outcomes:

1. answer with a proof that its required fields and evidence are covered by an existing contract;
2. recover retained source/dependencies and evaluate a new contract; or
3. abstain as unsupported.

Unregistered-query regret over set $\mathcal{Q}'$ is

$$R_{\text{new}} = \frac{1}{|\mathcal{Q}'|} \sum_{q' \in \mathcal{Q}'} [L(q', P_v) - L(q', P_{\text{full}})],$$

where L uses the query's declared decision loss. Report results by query family because averaging can hide systematic loss for a user, language, period, or evidence class.

Lifecycle accounting

For policy π spanning capture through disposition, keep the native outcome vector

$$\mathbf{C}(\pi) = (B_{\text{year}}, E_{\text{life}}, H_{\text{curator}}, T_{\text{query}}, T_{\text{recover}}, L_{\text{wrong}}, L_{\text{unreadable}}, L_{\text{privacy}}),$$

where byte-years, joules, person-hours, seconds, and task-native loss units are not added without published conversion weights. Charge payload, metadata, indexes, replicas, dependencies, emulators, tests, migrations, reviewer work, restore drills, legal holds, and verified deletion.

Byte reduction is

$$\rho B = 1 - \frac{|B_v| + |M_v| + |X_v| + |R_v|}{|B_0| + |M_0| + |X_0| + |R_0|},$$

where every magnitude is bytes and ρB is dimensionless. Backups or retained raw sources cannot be excluded from the numerator while supplying recovery.

Strongest null and rejection

Compare learned compaction with full version history, indexed snapshots plus suffix log, materialized views, key compaction and tombstones, lossless compression/deduplication, tiered cold archive, and an OAIS/PREMIS-style package with versioned schema/vocabulary and query-regression tests.

Reject the refinement when:

1. ordinary packaging plus query regression matches registered-query error, evidence reachability, recovery, and lifecycle cost;

2. supported queries or interpreter capabilities are chosen after migration;
3. fixity, provenance, citation, ontology consistency, or findability is substituted for correctness;
4. unregistered queries receive invented answers instead of recovery or abstention;
5. required deletion, holds, privacy, or authorization are omitted;
6. metadata, dependencies, migration, reviewer work, and cold fallback are not charged; or
7. records survive turnover but operators cannot safely interpret or use them.

Editable diagram: [query-registered-preservation.mmd](#).

Regime-qualified flow inference and control contract

This note defines the dimensional and inferential boundary for Fixture F-005. It operationalizes the residual identified by the fluid-dynamics and turbulence audit for Candidates 002, 003, 006, 007, 012, and 014.

The equations may be known while the state, boundary, forcing, constitutive response, unresolved scales, and future regime are not. The benchmark therefore binds every result to a physical/numerical identity, observation identity, regime/history identity, target, and complete resource boundary.

Immutable episode identity

For paired episode e , register

$$I_e = (h\mathcal{G}, h\mathcal{E}, h\mathcal{B}, h\mathcal{F}, h\mathcal{N}, h\mathcal{O}, h\mathcal{D}, h\mathcal{A}, h\mathcal{Q}),$$

where the hashes identify geometry $\mathcal{G}$, governing equations and constitutive assumptions $\mathcal{E}$, initial and boundary conditions $\mathcal{B}$, forcing process $\mathcal{F}$, numerical solver/grid/time-step configuration $\mathcal{N}$, observation and calibration operator $\mathcal{O}$, data and split lineage $\mathcal{D}$, actuator/plant interface $\mathcal{A}$, and declared quantities of interest $\mathcal{Q}$. Hashes are byte strings. Identity of bytes does not establish physical adequacy, convergence, or calibration.

The hidden regime/history record is

$$R_e = (Re, Pe, Sc, Ro, \mathbf{g}, \mathbf{b}, \mathbf{f}, \mathbf{d}, r_{0:t}, \tau_e),$$

where Re , Pe , Sc , and Rossby number Ro are dimensionless; $\mathbf{g}$ is geometry; $\mathbf{b}$ and $\mathbf{f}$ are boundary and forcing histories; $\mathbf{d}$ is disturbance shape and amplitude in native units; $r_{0:t}$ is ramp direction, rate, dwell, and prior regime occupancy; and τ_e is the episode horizon in seconds. Every dimensioned element of $\mathbf{g}$, $\mathbf{b}$, $\mathbf{f}$, $\mathbf{d}$, and $r_{0:t}$ carries a unit in the episode schema.

State, closure, and numerical discrepancy

For incompressible Newtonian flow,

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{u} + \mathbf{f}, \quad \nabla \cdot \mathbf{u} = 0,$$

where velocity $\mathbf{u}$ is in m s^{-1} , time t in seconds, pressure p in pascals, density ρ in kg m^{-3} , kinematic viscosity ν in $\text{m}^2 \text{s}^{-1}$, and body acceleration $\mathbf{f}$ in m s^{-2} .

At resolution Δ , represent the resolved evolution as

$$\dot{\mathbf{x}}_{\Delta} = \mathcal{N}_{\Delta}(\mathbf{x}_{\Delta}, \mathbf{b}, \mathbf{f}) + \mathcal{C}_{\theta, \Delta}(\mathbf{x}_{\Delta}) + \delta_{\Delta} + \epsilon_{\Delta},$$

where $\mathcal{N}_{\Delta}$ is the declared discretized resolved operator, $\mathcal{C}_{\theta, \Delta}$ is a physical or learned closure, δ_{Δ} is model-form discrepancy, and ϵ_{Δ} is numerical error. All four right-hand terms have state-units per second. A fitted residual can contain any mixture of the last three terms; portability is tested rather than inferred from training loss.

Let $s(z; \mathcal{S}_{\text{train}})$ be a preregistered dimensionless support distance from test condition z to the training support. Closure risk is reported as a function rather than a pooled mean:

$$\mathcal{R}_C(q, s) = \mathbb{E}[\ell_q(q(\mathbf{x}^{\text{ref}}), q(\tilde{\mathbf{x}})) \mid s(z; \mathcal{S}_{\text{train}}) = s].$$

q is a declared quantity of interest and loss ℓ_q retains the square or absolute unit induced by q . Out-of-support evaluation crosses solver, grid, order, geometry, boundary treatment, forcing band, Re , and regime rather than using randomly held-out neighboring snapshots.

Detector, filter, operator, and support identities

Every reported field or event is bound to

$$J = (\mathcal{V}, \mathcal{K}_{\ell}, \mathcal{H}, \mathcal{D}_{\gamma}, \mathcal{S}_x, \mathcal{S}_t, \mathcal{F}_r),$$

where $\mathcal{V}$ is the physical variable, $\mathcal{K}_{\ell}$ is a filter and width ℓ in metres, $\mathcal{H}$ is the measurement operator, $\mathcal{D}_{\gamma}$ is an event detector with threshold/parameter vector γ , $\mathcal{S}_x$ and $\mathcal{S}_t$ are spatial and temporal support, and $\mathcal{F}_r$ is the reference frame. Two results with different J are different estimands until a registered transfer map is validated.

For observation time t_k ,

$$\mathbf{y}_k = \mathcal{H}_{J,k}(\mathbf{x}_k) + \boldsymbol{\eta}_k, \quad \boldsymbol{\eta}_k \sim (\mathbf{0}, R_k),$$

where $\mathbf{y}_k$ retains sensor units, R_k has squared sensor units, and $\mathcal{H}_{J,k}$ includes averaging kernel, exposure time, transfer function, latency, synchronization, missingness, probe intrusion, and calibration state. Aggregate supply and recipient-level delivery are different estimands unless a registered support map identifies them (C-1488). Nominal contact area and aggregate load likewise do not identify real-contact support or the local-pressure tail (C-1498). The estimator never receives hidden truth through filenames, simulator metadata, shared random state, or a truth-identical forward model.

Signed scale transfer and event fidelity

For filter width ℓ ,

$$\tau_{ij}^{(\ell)} = \overline{u_i u_j} - \tilde{u}_i \tilde{u}_j, \quad \Pi_{\ell} = -\tau_{ij}^{(\ell)} \tilde{S}_{ij},$$

where stress $\tau_{ij}^{(\ell)}$ is in $\text{m}^2 \text{s}^{-2}$, filtered strain $\tilde{S}_{ij}$ in s^{-1} , and signed transfer Π_{ℓ} in $\text{m}^2 \text{s}^{-3}$. The flux score preserves direction:

$$L\Pi = \int_{\ell_{\min}}^{\ell_{\max}} w(\ell) \left| \langle \widehat{\Pi}_{\ell} \rangle - \langle \Pi_{\ell}^{\text{ref}} \rangle \right| d \log \ell,$$

where $w(\ell)$ is dimensionless and normalized over $d \log \ell$; $L\Pi$ is in $\text{m}^2 \text{s}^{-3}$. Report forward-transfer and backscatter event precision, recall, calibration, amplitude, spatial support, and duration separately. Matching a spectrum cannot substitute for $L\Pi$.

For coherent event a extracted by $\mathcal{D}_\gamma$, retain

$$Z_a = (t_a^{\text{on}}, t_a^{\text{off}}, \Omega_a, \Gamma_a, \mathbf{c}_a, \Delta q_a, J),$$

where onset and offset are seconds, Ω_a is spatial support in m^3 (or the declared lower-dimensional measure), Γ_a is a circulation-like attribute in $\text{m}^2 \text{s}^{-1}$ when applicable, $\mathbf{c}_a$ is position in metres, and Δq_a is the event-conditioned change in target q . Detector sweeps over reasonable filters, thresholds, frames, and planes are mandatory; event identity is never silently fixed.

Reduced state and adaptive allocation

For basis Φ_r and reduced state $\mathbf{a}_t$,

$$\hat{\mathbf{x}}_t = \bar{\mathbf{x}} + \Phi_r \mathbf{a}_t, \quad \dot{\mathbf{a}}_t = F_r(\mathbf{a}_t, \mathbf{u}_t) + \mathbf{c}_r(\mathbf{a}_t),$$

where rank r is a count, F_r is the projected or inferred dynamics, and $\mathbf{c}_r$ closes discarded modes. Basis normalization determines the units of $\mathbf{a}_t$. Reconstruction, autonomous rollout, forcing response, control, transition, and extreme fidelity are distinct losses.

Let available allocation units $i \in \mathcal{I}_t$ represent mesh cells, sensors, samples, model capacity, or compute quanta. An adaptive policy chooses $a_{i,t} \in \{0, 1, \dots\}$ under

$$\sum_{i \in \mathcal{I}_t} c_{i,t} a_{i,t} \leq B_t,$$

where cost $c_{i,t}$ and budget B_t use the same declared unit: cell-steps, bytes, seconds, or joules. Its goal-oriented efficiency is

$$\eta_q(\varepsilon_q) = \frac{\text{Work}_{\text{uniform}}(\varepsilon_q)}{\text{Work}_{\text{adaptive}}(\varepsilon_q)},$$

where both methods attain the same target-error tolerance ε_q in the native unit of q and work uses the same ledger. η_q is dimensionless. Regrids, subcycles, rejected steps, data transfers, synchronization, load imbalance, sensor movement/calibration, and allocator inference are included.

Observability, assimilation, and sensor value

For a scaled linearization $\mathbf{x}_{k+1} = A_k \mathbf{x}_k$ and $\mathbf{y}_k = C_k \mathbf{x}_k + \boldsymbol{\eta}_k$, define

$$W_o(N) = \sum_{k=0}^N \Phi(k, 0)^\top C_k^\top R_k^{-1} C_k \Phi(k, 0), \quad \Phi(k, 0) = A_{k-1} \cdots A_0.$$

W_o is dimensionless only after state and sensor scaling is fixed. Report its rank and conditioning locally by trajectory, regime, sensor set, noise model, and horizon; do not promote one linearization to global observability.

For posterior mean $\widehat{\mathbf{x}}_k$ and covariance P_k , the normalized estimation error squared is

$$\text{NEES}_k = (\mathbf{x}_k - \widehat{\mathbf{x}}_k)^T P_k^{-1} (\mathbf{x}_k - \widehat{\mathbf{x}}_k),$$

which is dimensionless. Calibration requires coverage and rank-aware tests in observable and unobservable subspaces, innovation whiteness, and recovery after dropout. A narrow posterior with structural bias is failure.

For sensor set S and target q , use the costed decision value

$$V_q(S) = \mathbb{E}[L_q(\varnothing) - L_q(S)] - \lambda_E E_S - \lambda_B B_S - \lambda_H H_S,$$

where L_q is expressed in a declared utility unit, energy E_S in joules, traffic B_S in bytes, and human maintenance H_S in person-hours. Conversion weights carry reciprocal units and are preregistered. Also report every raw term. Placement must remain physically feasible under regime shift, correlated failure, latency, bandwidth, calibration drift, and probe intrusion.

Mixing, transition, and extreme-event contracts

For passive scalar c with molecular diffusivity κ ,

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \nabla c = \kappa \nabla^2 c + s_c, \quad \chi = 2\kappa \langle |\nabla c|^2 \rangle.$$

If c is concentration in kg m^{-3} , source s_c is in $\text{kg m}^{-3} \text{s}^{-1}$ and scalar dissipation χ is in $\text{kg}^2 \text{m}^{-6} \text{s}^{-1}$. Mixing outcomes retain scalar variance, negative-Sobolev mix norm, χ , reaction completion, residence-time distribution, and remnant concentration at operational support. Visual filamentation is not a substitute. A countercurrent label likewise cannot replace finite conductance, capacity-rate, residence, leakage, boundary, and pump terms (C-1490).

For transition class c with history $r_0:t$, define hazard

$$\lambda_c(t \mid \mathbf{y}_0:t, r_0:t) = \lim_{\Delta t \downarrow 0} \frac{\Pr(t \leq T_c < t + \Delta t \mid T_c \geq t, \mathbf{y}_0:t, r_0:t)}{\Delta t},$$

with units s^{-1} . Report event-time likelihood, class-conditional calibration, lead time in seconds, false-alarm burden, abstention, and competing hazards. Ramp direction/rate, disturbance amplitude/shape, dwell, censoring, domain, and observation history are part of the conditioning state.

For extreme observable Q and preregistered threshold q_* ,

$$p_* = \Pr(Q > q_*), \quad T_R = \frac{\Delta t_{\text{eff}}}{p_*},$$

where p_* is dimensionless and return period T_R is seconds when Δt_{eff} is the effective independent sampling interval in seconds. Weighted rare-event estimators publish weights, effective sample size, degeneracy, variance, and validation on an untouched natural-distribution stream. Enriched event frequency without reweighting is invalid.

Control stability and complete energy

For plant state $\mathbf{x}$, command $\mathbf{u}_c$, disturbance $\mathbf{w}$, and delayed observation $\mathbf{y}t-d$,

$$\dot{\mathbf{x}} = F(\mathbf{x}, \mathbf{u}_c, \mathbf{w}), \quad \mathbf{u}_c = \pi(\mathbf{y}_{t-d}, \hat{R}_t),$$

where delay d is seconds and command components retain actuator units. Sweep delay, bandwidth, saturation, noise, forcing, regime, plant drift, failure, and fallback. Report constraint violations, gain/phase margins where defined, closed-loop poles for linearized controllers, bounded-input response, recovery, and runtime-monitor interventions. A task gain with unstable or unsafe strata does not enter an aggregate score. Equal mean flow also cannot establish phase margin or absence of coupled-delay oscillation (C-1492). Rate/state stick-slip supplies the same mean-versus-dynamics warning for interface state, compliance, inertia, damping, bandwidth, and control delay (C-1501).

Over service interval $[0, T]$, operational net energy is

$$E_{\text{net}} = \int_0^T (P_{\text{base}} - P_{\text{controlled}} - P_{\text{act}} - P_{\text{sense}} - P_{\text{compute}} - P_{\text{network}} - P_{\text{store}} - P_{\text{aux}}) dt,$$

where every power is watts, T is seconds, and E_{net} is joules. Report all terms separately. Installation, embodied energy, calibration, maintenance, retraining, replacement, and end-of-life costs are additional joule rows amortized only over a declared service life and duty cycle.

The complete run record is

$$\mathbf{K}_m = (N_{\text{cell-step}}, N_{\text{sample}}, N_{\text{solve}}, B_{\text{moved}}, B_{\text{stored}}, T_{\text{wall}}, M_{\text{peak}}, H_{\text{human}}, E_{\text{facility}}, E_{\text{sense}}, E_{\text{act}}, E_{\text{compute}}, E_{\text{network}}, E_{\text{store}}, E_{\text{embodied}}, E_{\text{maintenance}}).$$

Counts are dimensionless, bytes are bytes, times are seconds, peak memory is bytes, human work is person-hours, and every energy row is joules. Failed runs, search, data generation, tuning, verification, standby, and unused reservations remain in the ledger.

Outcome vector and decision

Keep the confirmatory outcome as a typed vector

$$\mathbf{Y} = (Y_{\text{field}}, Y_{\text{flux}}, Y_{\text{tail}}, Y_{\text{event}}, Y_{\text{closure}}, Y_{\text{ROM}}, Y_{\text{refine}}, Y_{\text{assim}}, Y_{\text{sensor}}, Y_{\text{control}}, Y_{\text{mix}}, Y_{\text{transition}}, Y_{\text{extreme}}, Y_{\text{measure}}, \mathbf{K}_m).$$

Each component retains its literal units and uncertainty. No weighted fluid, physics, fidelity, or efficiency score may replace it. The proposed composition survives only if it beats the strongest complete mature null at equal information and lifecycle budget on preregistered components, remains non-inferior on protected stability, calibration, conservation, and tail components, transfers across hidden regimes and model/hardware families, and has an isolating ablation. Otherwise retain this contract and retire the architectural explanation.

Editable system diagram: [regime-qualified-flow-inference-control.mmd](#).

One-sided topology signatures under repeated stimulation

- **Purpose:** freeze the source-shaped equations, support gates, event rules, uncertainty treatment and endpoints for the `RSD-T02-PULSE` written protocol
- **Claim:** C-1561
- **Primary source:** Rahi et al. 2017, bibliography key `rahi2017oscillatory`
- **Evidence audit:** repeated-stimulus pulse signatures
- **Parent contract:** intervention-qualified mechanism equivalence
- **Experiment contract:** Fixture F-026, RSD-T02-PULSE
- **Authority:** checked-in public-development constructor plus bounded panel orchestrator; `NO_RESULT`, no full-panel execution, actionable estimator, confirmation custody, workstation measurement or energy result

Inference boundary

This subtrack asks whether a repeated-stimulus response is a valid one-sided signature under a declared model and observation class. It does not infer a universal graph name.

The scoreable plant property is

$$qRI = \begin{cases} 1, & \text{the reported response causally drives its inhibitor,} \\ 0, & \text{the inhibitor is driven by the external input instead,} \\ \text{mixed,} & \text{both paths are present,} \\ \text{unresolved,} & \text{the allowed observations do not decide.} \end{cases}$$

Presence of a support-qualified stabilized refractory period or a support-qualified subharmonic response can support $qRI = 1$ against the registered pure feed-forward rival. Absence produces `unresolved`; it never proves $qRI = 0$. A measurement dead time, an undersampled output, a mixed motif or a model outside the source assumptions cannot inherit the diagnostic.

This is a companion stratum to `T02-MECH`. It does not alter the exact matched-step five-recipe bank and does not pretend that the source-shaped pair below has that bank's canonical step equality. `response_drives_inhibitor` is a subtrack-local causal coordinate; it must not be silently merged into `T02-MECH`'s `reported_output_feedback_edge` without a separate representation map and intervention certificate.

Source-shaped planted pair

Let $S_{d,T}(t)$ be a square pulse train in response units $\mathcal{U}$:

$$S_{d,T}(t) = \begin{cases} 1\mathcal{U}, & 0 \leq t - jT < d, \\ 0\mathcal{U}, & d \leq t - jT < T, \end{cases} \quad j = 0, 1, 2, \dots$$

where pulse duration d and onset-to-onset period T are both in seconds and $0 < d < T$. The response R and inhibitor I are in $\mathbb{U}$. With $\tau_R = 1$ s, the protected response-driven negative-feedback world is

$$\tau_R \dot{R} = \frac{S_d T}{1 + (I/I_0)^n} - R, \quad \tau_R \dot{I} = R - \lambda I. (1)$$

The paired input-driven incoherent feed-forward rival changes only the inhibitor drive:

$$\tau_R \dot{R} = \frac{S_d T}{1 + (I/I_0)^n} - R, \quad \tau_R \dot{I} = S_d T - \lambda I. (2)$$

Both start at the $S = 0$ equilibrium $(R, I) = (0, 0)$. The reported scalar is

$$O(t) = \left(\frac{R(t)}{1\mathbb{U}} \right)^3, (3)$$

which is dimensionless. Equations (1) and (2) are the **NFL 1** and **IFFL 1** forms studied by Rahi et al., with the hard threshold replaced by the Hill form that their systematic exploration also used. The protected synthetic parameter manifest is

$$n = 4, \quad I_0 = 0.01 \mathbb{U}, \quad \lambda = 0.3. (4)$$

Those values lie on the source's enumerated parameter grid. Their behavior in this fixture is a construction target, not a result imported from the paper. Before any comparison run, a two-resolution construction certificate must establish every truth label and support gate below. Failure invalidates the world version; it does not permit threshold tuning after inspection.

Rahi et al. used a first-order downstream output node driven by powers of R in this smooth-model exploration. At periodic steady state, its time average is proportional to the time average of its drive. Equation (3) therefore retains the same refractory argmax while event extraction remains on R itself. This is a fixture simplification, not a claim that the transient output nodes are identical.

Step adaptation and pulse support

For a step $S(t) = 1 \mathbb{U} \mathbf{1}[t \geq 0]$, let

$$O_{\text{pk}} = \max_{0 \leq t \leq 200 \text{ s}} O(t), \quad \tau_a = \min_{0 \leq t \leq 200 \text{ s}} \arg \max O(t),$$

and let O_{ss} be the certified late-time value. The source-shaped adaptation gate is

$$A_{\text{step}} = 1 - \frac{O_{\text{ss}}}{O_{\text{pk}}} > 0.80. (5)$$

The refractory analysis may use only durations satisfying

$$d \geq 1.5 \tau_a. (6)$$

For each duration, an isolated pulse followed by a 100 s washout defines the response-amplitude reference

$$A_{\text{iso}}(d) = \max_t R_d(t).$$

The registered event threshold is the protocol convention

$$\theta_R(d) = 0.25A_{\text{iso}}(d).(7)$$

Thresholds $0.15A_{\text{iso}}$ and $0.35A_{\text{iso}}$ are sensitivity diagnostics only. A pulse cell is in support only if all of the following hold:

1. Equation (5) passes and the off-state equilibrium is recovered before the first pulse, meaning $\max\{|R|, |I|\}/(1 \text{ U}) \leq 10^{-8}$ after washout.
2. $A_{\text{iso}}(d)$ exceeds five times its numerical error bound.
3. A single isolated pulse causes exactly one registered upward crossing of $\theta_R(d)$ before washout. Zero or multiple response cycles are outside the source-shaped single-response region.
4. The pulse edges are represented exactly by solver stops and the observation cadence resolves every threshold crossing with a latency error bound.
5. No saturation, clipping, censoring, output dead time, undocumented filter or future sample changes the event word.

The source's lower boundary - insufficient inhibitor accumulation - and upper boundary - more than one response to one long pulse - are therefore explicit support failures rather than favorable or unfavorable observations.

Refractory-period stabilization

After convergence, define the cycle-averaged output over K complete cycles:

$$O(d, T) = \frac{1}{KT} \int_{t_b}^{t_b+KT} O(t; d, T) dt.(8)$$

Use $K = 20$ after the convergence time t_b . For fixed d , the refractory period is

$$T_{\text{max}}(d) \in \arg \max_{T > d} O(d, T).(9)$$

If several periods are tied within the certified output error, retain the full maximizer interval and report its largest member only as a descriptive value. The decision must propagate the complete interval.

The protected stabilization panel is

$$\mathcal{D}_{\text{stab}} = \{0.30, 0.50, 1.00, 1.50, 2.00, 3.00\} \text{ s}.(10)$$

For each d , scan T from $d + 0.20$ s through 40.00 s at 0.20 s spacing, then refine the two coarse neighbors around every tied maximum at 0.01 s spacing. This response-dependent refinement rule is frozen before outputs and is applied to every world.

Let $[T_d^-, T_d^+]$ be the error-qualified maximizer interval. For adjacent durations $d_i < d_{i+1}$, the conservative secant-slope interval is

$$M_i = \left[\frac{T_{d_{i+1}}^- - T_{d_i}^+}{d_{i+1} - d_i}, \frac{T_{d_{i+1}}^+ - T_{d_i}^-}{d_{i+1} - d_i} \right]. (11)$$

Refractory stabilization is certified only if at least two consecutive support-qualified intervals satisfy

$$\sup M_i < \frac{1}{2}. (12)$$

The slope is dimensionless. A maximizer touching the T boundary, an unresolved maximizer interval, a duration violating (6), or a cell with more than one response to one pulse makes stabilization `unresolved`.

Response count and latency

For pulse j , let N_j be the number of upward crossings of $\theta_R(d)$ in $[jT, (j+1)T)$. Define

$$b_j = \mathbf{1}[N_j \geq 1], \quad N_{\text{resp}} = \sum_{j=1}^J b_j, (13)$$

where N_{resp} is a count and J is the stimulus count. The first crossing latency is

$$L_j = \inf\{t - jT : R(t) \uparrow \theta_R(d), jT \leq t < (j+1)T\}, (14)$$

in seconds. A missing event has no latency; it is not encoded as zero or T . Report the observed L_j values, missing-event count, event amplitude and all N_j . A cell with any $N_j > 1$ fails the single-response support gate.

Period skipping

The protected skipping cells use

$$d = 0.20 \text{ s}, \quad T \in \{5.00, 5.20, 5.40\} \text{ s}. (15)$$

These are synthetic protocol choices. For $q \in \{1, 2, 3, 4, 5\}$, compare the complete state and output waveforms over one stimulus period after convergence:

$$\Delta_q = \max_{0 \leq s < T} |x(tb + s) - x(tb + s - qT)|_{\text{scaled}, \infty}, \quad x = (R, I, O). (16)$$

The scale is 1 U for R and I and one for O . Let η_q be the two-resolution bound for Δ_q . Period skipping is certified only when all of these conditions hold:

1. the smallest certified recurrence order q_* is in $\{2, 3, 4, 5\}$;
2. $\Delta_{q_*} + \eta_{q_*} \leq 10^{-10}$ while $\Delta_1 - \eta_1 > 10^{-8}$;
3. the event word of length q_* contains at least one response and at least one skipped pulse and repeats for four complete q_* cycles;
4. every included pulse has $N_j \leq 1$; and
5. the direct plant output, not a censored or aliased observation, supplies the event word.

If no $q \leq 5$ converges before the time cap, the state is `unresolved`, not `aperiodic NFL` and not `IFFL`. The largest consecutive zero run, q_* , N_{resp}/J , latency distribution and recurrence residuals are secondary outputs.

Checked-in construction snapshot

The deterministic constructor in `rsd-t02-pulse.mjs` now implements the six-world registry, exact pulse-edge stops, adaptive Dormand--Prince 5(4) integration, the two-resolution gates, recurrence orders $q \leq 5$, event extraction, bounded refractory search, typed costs and fail-closed `NO_RESULT` records. Its focused test suite covers the protected feedback/feed-forward pair, the signature-negative linear feedback world, dead-time and alias hostiles, deterministic OU diagnostics and malformed paths.

One executed construction cell at $d = 0.20$ s and $T = 5.00$ s gives:

1. `PS-NFL-H4`: a support-qualified order-two recurrence, event word `01` repeated ten times, 10 responses from 20 stimuli, recurrence upper bound 2.94×10^{-15} and first-response latencies near 0.04694 s;
2. `PS-IFFL-H4`: order-one recurrence and an absent skipping signature; and
3. an unresolved topology disposition for both records. Signature presence is evaluator construction truth, not an actionable inference result.

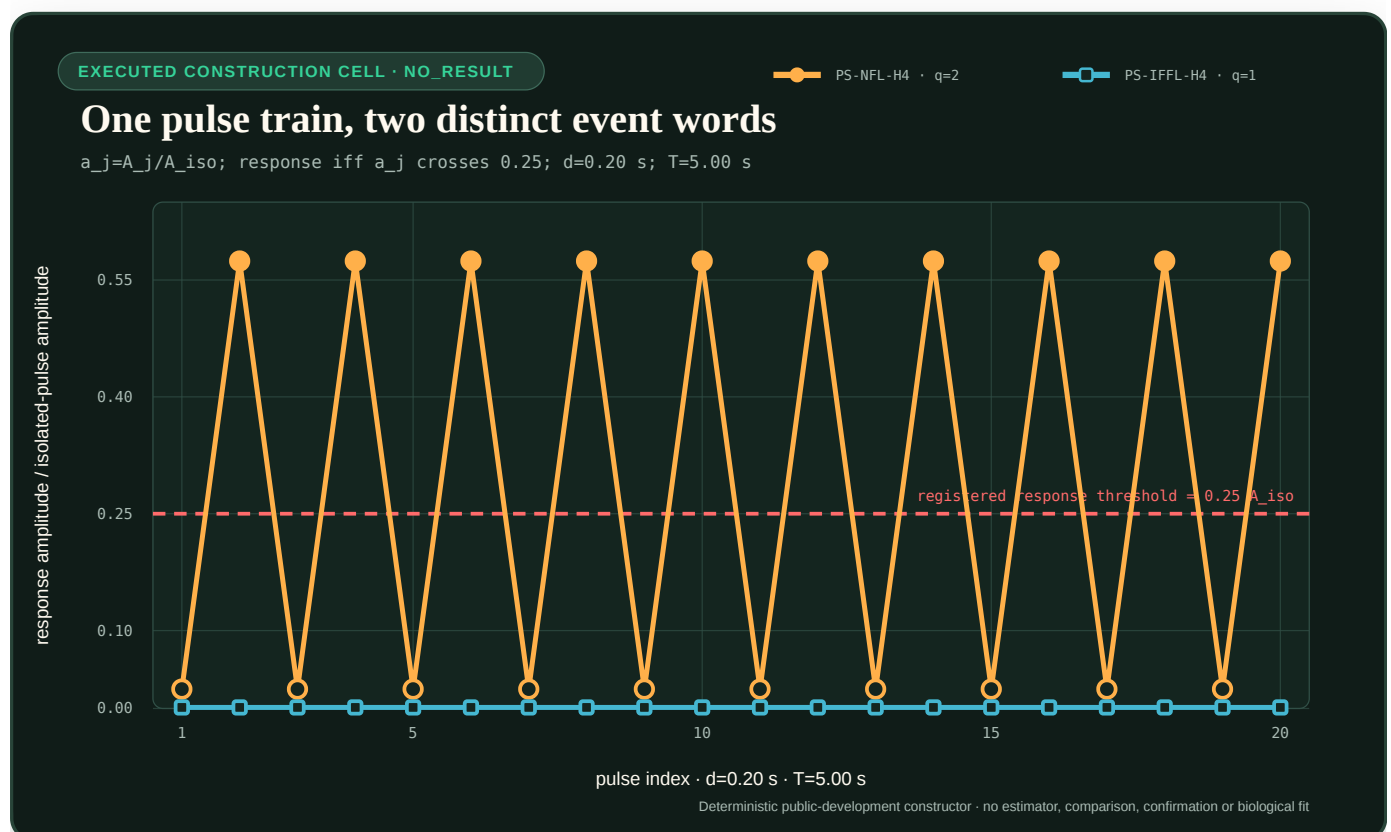


FIGURE 70 — Normalized pulse-response amplitudes for one executed construction cell

The figure is generated from the constructor itself by `generate-plots.mjs`; its editable parameters are in `core-models.json`. Filled amber markers cross the registered response threshold on every second feedback-world pulse. The cyan feed-forward trace remains below threshold. This is one public development cell, not a biological fit, estimator comparison or confirmation result.

A separate bounded diagnostic scan for `PS-NFL-H4` at $d = 0.30$ s evaluated 198 coarse and 38 refined cells and retained the interior maximizer interval $[16.72, 16.72]$ s. That single duration cannot establish stabilization: equation (12) needs adjacent duration slopes. The six-duration

refractory panel, 64-seed noise grid and mixed-window panel remain unexecuted.

The separate `rsd-t02-pulse-panel-runner.mjs` freezes an ordered 229-unit execution plan:

1. four mixed-window contract records that remain unresolved while window starts or widths are unfrozen;
2. 18 refractory units from three worlds $\times$ six durations;
3. 15 skipping units from five worlds $\times$ three periods; and
4. 192 robustness units from 64 public seeds $\times$ three OU noise levels, each targeting the deterministic panel only after its construction gates pass.

The runner defaults to zero work and permits at most eight units per invocation. Its append-only resume identity binds the exact config, schedule, runner, constructor and event-schema bytes plus the Node/V8/libuv and host runtime fingerprint. A 16 MiB per-unit cap and 256 MiB whole-run cap apply to serialized results. The built-in executor can construct refractory and skipping units; clean-sample OU production remains external and mixed windows remain deliberately unresolved. No 229-unit run was started. Static imports are loaded before their source files are fingerprinted, so same-process source mutation across that load/read interval remains outside this public construction boundary.

Null and counterworld bank

The versioned bank contains the following independent cases.

1. **PS-NFL-H4** : equations (1), (3) and (4). Its positive signature status must be established by a construction certificate before it enters any scored partition.
2. **PS-IFFL-H4** : equations (2), (3) and (4). It is the paired pure feed-forward rival. Any feedback assignment in this world is a false attribution.
3. **PS-NFL-LTI** : a signature-negative feedback control,

$$\tau_R \dot{r} = S - r - i, \quad \tau_I \dot{i} = r, \quad (\tau_R, \tau_I) = (1, 4) \text{ s}.$$

This stable linear negative-feedback system adapts its signed response r to zero but entrains to a periodic input. It checks that absence of the nonlinear signatures remains `unresolved` instead of being converted to an IFFL declaration.

4. **PS-MIXED** : the source's parallel slow-IFFL/fast-NFL counterworld. Use dimensionless states and normalized time $\hat{t} = t/(1 \text{ s})$:

$$\frac{dI_1}{d\hat{t}} = S - \lambda_1 I_1, \quad \frac{dI_2}{d\hat{t}} = R - \lambda_2 I_2,$$

$$\frac{dR}{d\hat{t}} = \frac{S}{1 + (I_2/I_0)^n} - (1 + \kappa I_1)R, \quad O = R^3,$$

with $\lambda_1 = 1/200$, $\lambda_2 = 0.2$, $\kappa = 0.01$, $n = 1$ and $I_0 = 0.1$. Under the registered one-second normalization, evaluate O over windows ending at 50, 150, 250 and 350 s. A window-dependent signature reports `mixed/window-qualified`, never a single exclusive graph.

5. **PS-DEADTIME** : the direct plant is **PS-IFFL-H4**, but the observation layer suppresses reported events for $1.5T$ after every detected event. The resulting apparent skip pattern is a calibrated measurement-recovery artifact and must fail the observation support gate.

6. **PS-ALIAS**: the direct plant is **PS-IFFL-H4**, but the reported cadence is $\Delta t = T$ at a fixed phase. It cannot bound event count or latency and must fail temporal-resolution support.

The graph, equations, parameters, direct-versus-observed channel and counterworld identity are evaluator-only. Actionable arms receive causal input, reported output, timestamps, units, missingness and registered instrument metadata. They do not receive the truth label or future samples.

Numerical construction certificate

The checked-in constructor uses binary64 integration with exact stops at pulse edges. Its adaptive Runge-Kutta reference run uses absolute tolerance 10^{-12} U for state variables and relative tolerance 10^{-10} ; the refinement run halves both tolerances. The certificate retains maximum state/output disagreement, quadrature disagreement, threshold-crossing time disagreement and recurrence-residual disagreement.

Convergence is tested immediately before pulse onset. The state must recur to fractional scaled error 10^{-10} for $q = 1$ or the smallest $q \leq 5$, and the same recurrence must persist for four additional cycles. Total simulated time is capped at 20,000 s per cell. Nonconvergence, solver failure, nonfinite state, negative concentration in a concentration-qualified world or refinement disagreement produces **unresolved** or **malformed**; it is never silently dropped.

No fixed numerical value in this note is an empirical biological tolerance. Equations (7), (10), (15), the solver tolerances and the finite time cap are synthetic protocol choices.

Endpoints and decision

The primary outputs are separate:

1. **refractory_signature** in {supported, absent, unresolved, out_of_support} using (12);
2. **skipping_signature** in the same set using (16);
3. **feedback_support** in {supported, unresolved, contradicted, out_of_support}; and
4. false feedback attribution on **PS-IFFL-H4**, **PS-DEADTIME** and **PS-ALIAS**.

feedback_support=supported requires at least one positive signature, all of its support gates, and a certified source-class rival. Two absent signatures yield **unresolved**. **PS-NFL-LTI** is correct only when the method abstains. **PS-MIXED** is correct only when the output retains the window qualification.

Report response count, stimulus count, seconds of latency, $T_{\max}$ in seconds, dimensionless slope, recurrence order, recurrence residual, simulated seconds, pulse cells, samples, bytes, solver evaluations, scalar operations, retained-state bytes, wall seconds and later calibrated joules as separate fields. No scalar efficiency score is registered.

Kill rules

Kill a claim-eligible use of this subtrack if any of the following occurs:

1. one isolated pulse, graph identity or evaluator-only parameters supply the answer;
2. absent signatures are scored as evidence for feed-forward structure;
3. $d < 1.5\tau_a$, insufficient adaptation or multiple responses per pulse are retained in the stabilization denominator;
4. a boundary maximum or unresolved maximizer interval is converted to a point estimate;
5. a dead-time or aliased observation is called plant feedback;

6. a mixed motif is forced into one exclusive topology;
7. event thresholds, duration cells, period cells or convergence order are changed after output inspection;
8. malformed/nonconverged cells are deleted; or
9. construction checks are presented as a comparison, biological prevalence, workstation or energy result.

Disposition

This note and its checked-in module complete a bounded construction layer for C-1561. The bounded panel schedule and append-only runner now exist, but they do not complete the six-duration and robustness executions, the mixed-window statistic, actionable estimator comparison, prospective confirmation partition or workstation execution. The only current authority is `public-development`, `construction-only` and `NO_RESULT`.

Representative adaptive-performance contract

This note formalizes Fixture F-006 from the sports expertise, adaptive performance, and team coordination audit. It supplies a comparison contract for Candidate 002, Candidate 004, Candidate 006, Candidate 007, Candidate 009, Candidate 012, Candidate 014, and Candidate 019. It creates no new candidate.

Versioned episode state

For agent i in episode e , preserve

$$\mathcal{K}_{i,e} = (X_e, O_{i,e}, A_{i,e}, H_{i,e}, M_e, F_{i,e}, R_{i,e}, D_{i,e}, G_{i,e}, C_e, U_e, B_e),$$

where:

- X_e is the physical state, rules, task geometry, deadline, consequence, and hidden regime of episode e ;
- $O_{i,e}$ is the information actually received by agent i , including source, support, latency in seconds, occlusion, noise, loss, and calibration;
- $A_{i,e}$ is the feasible action set under the current body, actuator, equipment, authority, rate, range, and safety constraints;
- $H_{i,e}$ is the timestamped acquisition and selection history: practice, feedback, opponents, teammates, injury or faults, prior exclusions, and opportunities that were offered or withheld;
- M_e is the teammate and opponent roster, role assignment, policy history, turnover event, and communication topology;
- $F_{i,e}$ is the feedback channel, its delay in seconds, its information in bits, and the party choosing when it is supplied;
- $R_{i,e}$ is the resource and fatigue state, with each component stored in its native unit rather than collapsed into a readiness score;
- $D_{i,e}$ is the damage or fault state, diagnostic uncertainty, protected capability envelope, and current return stage;
- $G_{i,e}$ is the selection and opportunity policy that determines access to training, roles, observation, intervention, and later outcome measurement;
- C_e is the randomized intervention, control, counterfactual pair, and stopping rule;
- U_e is the sampling unit, such as action, possession, episode, agent, dyad, team, site, season, or cohort; and
- B_e is the complete resource ceiling: events, bytes, seconds, person-hours, joules, damage, unsafe events, replacements, and opportunity.

The comparison estimand for method m and literal outcome k is

$$Q_{m,k}(\mathcal{K}) = \mathbb{E}[Y_k \mid do(m), \mathcal{K}],$$

where Y_k is measured in the registered unit for outcome k . A contrast $Q_{m,k} - Q_{b,k}$ against baseline b is uninterpretable when any element of $\mathcal{K}$ differs without a registered intervention or adjustment.

Outcome firewall

No scalar “performance” score may replace the following vector:

$$\mathbf{Y} = (Y^{\text{ant}}, Y^{\text{int}}, Y^{\text{cue}}, Y^{\text{prac}}, Y^{\text{ret}}, Y^{\text{tr}}, Y^{\text{exp}}, Y^{\text{adapt}}, Y^{\text{pace}}, Y^{\text{ready}}, Y^{\text{return}}, Y^{\text{team}}, Y^{\text{dec}}, Y^{\text{tal}}, \mathbf{C}).$$

The components and their units are:

SYMBOL	OUTCOME	REQUIRED LITERAL MEASUREMENT AND UNIT
Y^{ant}	anticipation	proper predictive score in bits per event, calibration error dimensionless, commitment latency in milliseconds
Y^{int}	physical interception	success probability dimensionless, endpoint error in metres, movement onset in milliseconds, unsafe-event probability dimensionless
Y^{cue}	cue use	causal score change under a declared cue intervention, in bits per event or the registered task unit
Y^{prac}	practice performance	literal task quality by attempt number and exposure time in seconds
Y^{ret}	delayed retention	task quality after a delay Δt_{ret} in hours or days without the training scaffold
Y^{tr}	transfer	source-to-target task quality and gap in the literal task unit
Y^{exp}	exploration	action and outcome entropy in bits, action--outcome information in bits, coverage dimensionless, and later utility
Y^{adapt}	adaptability and recovery	perturbation loss, time to regain the envelope in seconds, overshoot, recurrence probability, and residual damage
Y^{pace}	pacing	power in watts or action intensity in its declared unit as a time series, plus terminal task quality
Y^{ready}	fatigue and readiness	task-specific capacity change, state-estimation error, calibrated admissibility, abstention, and recovery time
Y^{return}	staged return	false promotion, false withholding, stage dwell time in hours, recurrence, rollback, availability, and collateral loss
Y^{team}	coordination and shared information	team task quality, task-variable variance, compensation lag in seconds, belief log loss in bits per event, messages, bytes, cross-play, and repair latency
Y^{dec}	deception and opponent adaptation	opponent log loss in bits per action, calibration, exploitability, regret, abstention utility, and adaptation time
Y^{tal}	talent prediction	prospective calibration, false-negative recovery, later capability, attrition, opportunity received, and subgroup error
$\mathbf{C}$	complete efficiency	events, bytes, wall-seconds, person-hours, joules, equipment, damage, unsafe events, replacements, and opportunity cost as separate axes

Representative distance is a vector

Let P_{tr} and P_{te} be the training and target distributions. Register

$$\mathbf{d}_{\text{rep}} = (dO, dA, dT, dM, dF, dR, dD, dG),$$

where dO compares received information, dA feasible actions, dT deadline and consequence, dM teammate/opponent composition and policy, dF feedback, dR resource state, dD damage/return state, and dG selection/opportunity policy. Each d_z is a declared divergence

between the corresponding marginals or conditionals under P_{tr} and P_{te} . It is dimensionless for a statistical divergence and has the registered ground-cost unit for optimal transport. No unreported weighted sum is a valid “representativeness” score.

For source stratum s and target stratum t , preserve the transfer matrix

$$Q_{m,s \rightarrow t}^{\text{tr}} = \mathbb{E}[Y^{\text{tr}} \mid \text{do}(m), s, t], \quad G_{m,s \rightarrow t} = Q_{m,s \rightarrow s}^{\text{tr}} - Q_{m,s \rightarrow t}^{\text{tr}},$$

where both Q^{tr} and the transfer gap G use the registered task unit. Report G separately by cue, feasible action, opponent, feedback, resource, damage, and selection-policy changes.

Anticipation, interception, and cue use

For N independent events, outcome y_n , actual observation history $o_{n, \leq \tau}$ available by occlusion time τ in seconds, and predictive distribution p_m , define

$$L_m(\tau) = -\frac{1}{N} \sum_{n=1}^N \log_2 p_m(y_n \mid o_{n, \leq \tau}),$$

where L_m is log loss in bits per event. For information channel c , the registered causal cue value is

$$V_{m,c}(\tau) = L_{m,-c}(\tau) - L_{m,\text{all}}(\tau),$$

also in bits per event. Here $L_{m,-c}$ is measured under removal or neutralization of channel c , not inferred from gaze or saliency. Predictive regulation must additionally retain false-alarm action, reserve debit, recovery, and cumulative exposure instead of treating cue value as the whole outcome (C-1494).

Physical coupling is reported separately as

$$I_m = (p_{\text{hit}}, e_{\text{end}}, t_{\text{move}}, p_{\text{unsafe}}),$$

where p_{hit} is dimensionless interception success, e_{end} is endpoint error in metres, t_{move} is movement onset in milliseconds, and p_{unsafe} is dimensionless unsafe-event probability. Label or joystick accuracy cannot substitute for I_m .

Practice, retention, transfer, and exploration

Let $q_m(n, t)$ be literal task quality after attempt n and exposure time t in seconds. Keep three estimands:

$$Q_m^{\text{prac}}(n) = q_m(n, t_n), \quad Q_m^{\text{ret}}(\Delta t) = q_m(n_{\text{last}}, t_{\text{last}} + \Delta t), \quad Q_{m,s \rightarrow t}^{\text{tr}} = q_m \text{ on target } t,$$

where Δt is the scaffold-free retention delay in hours or days. The first target trial $Q_{m,s \rightarrow t}^{\text{tr},1}$ is frozen before any target update; later adaptation is a separate curve.

For action variable A and reached-outcome variable Z , exploration is

$$\mathcal{X}_m = (H_m(A), H_m(Z), I_m(A; Z), K_m, Q_m^{\text{tr}}, C_m),$$

where both entropies and mutual information are in bits, K_m is coverage of the registered feasible region as a dimensionless fraction, Q_m^{tr} is later transfer in its task unit, and C_m is the separate cost vector. Higher action entropy without outcome information or later utility is not useful exploration.

After a perturbation at time t_0 , define recovery time

$$T_m^{\text{rec}} = \inf \{t - t_0 : q_m(u) \in \mathcal{A}_q \text{ for every } u \in [t, t + h]\},$$

where T_m^{rec} and the stability horizon h are in seconds, $q_m(u)$ is task quality in its registered unit, and $\mathcal{A}_q$ is the preregistered admissible quality envelope. Overshoot, recurrence, and damage are additional axes rather than hidden inside T_m^{rec} .

Resource state, pacing, and readiness

Let external power $P(t)$ be in watts over event duration T in seconds. The external work is

$$E_{\text{ext}} = \int_0^T P(t) dt,$$

where E_{ext} is in joules. Metabolic, device, facility, embodied, and lifecycle energy use different boundaries and remain separate ledger rows.

A resource-qualified controller has the form

$$a_t = \pi_m(o_{\leq t}, \hat{r}_t, \hat{d}_t, s_t, \hat{\pi}_{\text{opp},t}, \hat{b}_{\text{team},t}, f_t),$$

where a_t is the commanded action or power target in its native unit, $o_{\leq t}$ is causally received observation history, $\hat{r}_t$ is the estimated resource/fatigue vector, $\hat{d}_t$ is the estimated damage state, s_t is remaining work in metres, seconds, events, or joules, $\hat{\pi}_{\text{opp},t}$ is the opponent-policy estimate, $\hat{b}_{\text{team},t}$ is the teammate-state estimate, and f_t is available feedback. Each estimate and channel receives its own ablation.

Readiness is a calibrated action envelope rather than a score:

$$\mathcal{A}_t^{\text{ready}}(\alpha) = \{a \in A_t : \Pr(Z_{t:t+h} \in \mathcal{Z}_{\text{safe}} | a, \mathcal{I}_t) \geq 1 - \alpha\},$$

where A_t is the feasible action set, $Z_{t:t+h}$ is the multidomain outcome vector over horizon h in hours, $\mathcal{Z}_{\text{safe}}$ is the registered safe envelope, $\mathcal{I}_t$ is information available at decision time, and α is the dimensionless tolerated risk. Empty envelopes require abstention or escalation.

Staged and reversible return

Let $g_t \in \{0, 1, 2, 3, 4\}$ denote protected, modified, controlled, full-load, and adversarial operation. Promotion is admissible only if

$$g_{t+1} = g_t + 1 \quad \text{and} \quad \Pr(Z_{t:t+h} \in \mathcal{A}_{g_t+1} | \mathcal{I}_t) \geq 1 - \alpha_{g_t+1},$$

where $\mathcal{A}_g$ is the multidomain admissible envelope for stage g , h is the follow-up horizon in hours or days, and α_g is its dimensionless risk tolerance. If the current envelope is violated, the gate must allow

$$gt+1 < gt.$$

Report false promotion, false withholding, dwell time in hours, recurrence, rollback count, availability, damage, and human adjudication hours separately.

Team coordination and shared information

For task variable $z(t)$ and agent contribution $u_i(t)$, perturb agent i by $do(\eta_i)$ and estimate

$$\Gamma_{ij}(\ell) = \text{Cov}(\Delta u_i(t), \Delta u_j(t + \ell) \mid do(\eta_i), X_t),$$

where ℓ is lag in seconds and X_t is task state. Compensation requires both a registered response in Γ_{ij} and reduced task-variable error. Here u_i and u_j are recorded in their native contribution units, η_i is a registered perturbation in the unit of agent i 's action or state, and Γ_{ij} has the product unit of the two contributions; correlation or synchrony without intervention is insufficient. Common-drive and edge-intervention controls are therefore mandatory before mapped synchrony can be credited with useful coordination (C-1495).

For teammate j 's future action or intent $b_{j,n}$ and agent i 's predictive belief p_i , shared-information quality is

$$L_{i \rightarrow j} = -\frac{1}{N} \sum_{n=1}^N \log_2 p_i(b_{j,n} \mid h_{i,n}),$$

where $L_{i \rightarrow j}$ is in bits per event, $h_{i,n}$ is information actually available to i , and N is the number of independent team events. Report it under message ablation, teammate turnover, role reassignment, and never-co-trained cross-play alongside messages, bytes, latency, repair, and task quality.

Deception and opponent adaptation

For opponent action a_n^{opp} , available history h_n , and estimate $\hat{\pi}_m$, define

$$L_m^{\text{opp}} = -\frac{1}{N} \sum_{n=1}^N \log_2 \hat{\pi}_m(a_n^{\text{opp}} \mid h_n),$$

in bits per opponent action. For matched genuine and deceptive interventions,

$$\Delta_{m,k}^{\text{dec}} = Q_{m,k}(do(\text{deceptive})) - Q_{m,k}(do(\text{genuine})),$$

where k names a literal outcome and the difference retains its unit. Report calibration, confidence, exploitability, regret, abstention utility, and adaptation time separately under known, held-out, changing, and colluding opponents.

Selection, opportunity, and prospective prediction

Let $S_i \in \{0, 1\}$ denote selection, Z_i preselection evidence, O_i^+ postdecision opportunity in hours or task exposures, and Y_i^{future} later capability in its task unit. The prospective selection-policy estimand is

$$\Delta^{\text{sel}}(z) = \mathbb{E}[Y_i^{\text{future}} \mid \text{do}(S_i = 1), Z_i = z] - \mathbb{E}[Y_i^{\text{future}} \mid \text{do}(S_i = 0), Z_i = z].$$

It cannot be estimated by comparing selected survivors with excluded agents when selection changes O_i^+ , coaching, opponents, follow-up, attrition, or injury exposure. Report prospective calibration in new cohorts, selection and opportunity rates, false-negative recovery, attrition, censoring, subgroup error, later capability, and complete development cost.

Complete efficiency and equal budgets

For method m , retain lifecycle energy

$$E_m^{\text{life}} = E_m^{\text{train}} + E_m^{\text{infer}} + E_m^{\text{sense}} + E_m^{\text{act}} + E_m^{\text{comm}} + E_m^{\text{facility}} + E_m^{\text{recover}} + E_m^{\text{maint}} + E_m^{\text{emb}},$$

where every E term is in joules under one declared service interval. The terms denote training, inference, sensing, actuation, communication, facility, recovery, maintenance, and amortized embodied energy, respectively.

Human effort is

$$H_m^{\text{human}} = H_m^{\text{design}} + H_m^{\text{coach}} + H_m^{\text{demo}} + H_m^{\text{label}} + H_m^{\text{tune}} + H_m^{\text{monitor}} + H_m^{\text{repair}} + H_m^{\text{medical}},$$

where each term is in person-hours and roles are reported separately. The complete cost vector is

$$\mathbf{C}_m = (N_{\text{event}}, N_{\text{step}}, N_{\text{query}}, N_{\text{byte}}, T_{\text{wall}}, H^{\text{human}}, E^{\text{life}}, N_{\text{unsafe}}, D_{\text{harm}}, C_{\text{opp}}),$$

where the four N terms count events, environment or optimization steps, queries, and bytes; T_{wall} is wall time in seconds; H^{human} is person-hours; E^{life} is joules; N_{unsafe} counts unsafe events; D_{harm} is damage in a registered physical or severity unit; and C_{opp} is withheld opportunity in task exposures or person-hours.

Method m is feasible only if

$$\mathbf{C}_m \preceq \mathbf{B},$$

where $\mathbf{B}$ is the preregistered componentwise ceiling with the same units. A complete efficiency claim requires non-inferiority on every protected outcome and a Pareto improvement on at least one preregistered resource axis. An over-budget run is infeasible, not a score to normalize afterward.

Confirmatory contrast and retirement

Let $b^*(t)$ be the strongest mature baseline for track t , selected on development data before confirmatory outcomes open. For protected outcome set $\mathcal{P}_t$, retain a residual only when

$$\Pr(Q_{m,k} - Q_{b^*(t),k} > \delta_{t,k} \text{ for every } k \in \mathcal{P}_t) \geq 1 - \alpha_t,$$

where $\delta_{t,k}$ is the preregistered improvement or non-inferiority margin in the unit of outcome k , and α_t is the dimensionless error budget. The contrast must survive actual-channel, feasible-action, history, opponent/team, feedback, resource, damage, selection, and complete-cost ablations on held-out task, model, site, and hardware strata. Otherwise retire the mechanism claim while preserving the measurement contract.

State-qualified intervention and withdrawal

Scope

This note formalizes the held pharmacology/toxicology transfer: commanded intervention, realized internal state, mechanism engagement, benefit, harm, adaptation, and withdrawal remain separate. It supports the pharmacology audit and Candidates 005, 007, 012, and 014. The evidence and boundary claims are C-607–C-626.

Intervention chain

For system state x_t , commanded intervention d_t , realized internal exposure e_t , adaptation state z_t , and observation y_t , use

$$e_{t+1} = f_e(e_t, d_t, c_t, w_t^e),$$

$$x_{t+1} = f_x(x_t, e_t, z_t, c_t, w_t^x), \quad z_{t+1} = f_z(z_t, e_t, x_t, c_t, w_t^z),$$

$$y_t = h_v(x_t, e_t, z_t, d_{0:t}, s_t) + \epsilon_t.$$

Here c_t is typed context, s_t is sampling/selection state, v is the observation-model version, and every disturbance w_t^i and error ϵ_t has the unit of the equation in which it appears. The command d_t and exposure e_t need not share a unit: a token quota, update rate, or tool permission is not evidence that the intended computation, parameter change, or external effect occurred. Forecast value, feedback correction, integrated action exposure, reserve debit, and later outcome remain separate coordinates for the same reason (C-1494).

The evaluator reports endpoint vectors rather than a hidden scalar:

$$\mathbf{q}_t = (\mathbf{b}_t, \mathbf{h}_t, \mathbf{r}_t, E_t, L_t, M_t, H_t),$$

where $\mathbf{b}_t$ is task benefit, $\mathbf{h}_t$ is protected harm, $\mathbf{r}_t$ is remaining reserve, E_t is lifecycle energy in joules, L_t is latency in seconds, M_t is moved/stored information in bytes, and H_t is human effort in person-seconds. Unlike coordinates are never added without a declared decision rule and weights.

Qualified authority region

An intervention is admissible only inside a versioned region

$$\mathcal{D}_t = \mathcal{A}_v(\hat{e}_t, \hat{x}_t, \hat{z}_t, \mathbf{b}_t, \mathbf{h}_t, \mathbf{r}_t, u_t, \tau_t, p_t), \quad d_t \in \mathcal{D}_t,$$

where u_t is the complete uncertainty record, τ_t is observation age in seconds, and p_t identifies the supported population/task/context. Calling $\mathcal{D}_t$ a “window” does not make it a fixed scalar interval. It changes with endpoint, schedule, context, uncertainty, adaptation, reserve, and the cost of delayed harm. A local rule must also be requalified when its trigger becomes widespread and its shared-system cost changes (C-1493).

Withdrawal is its own transition

Removing support is not algebraic negation of adding support. Let $\Delta d_t = d_t - d_{t-1}$. A reduction policy must satisfy

$$d_t \in \mathcal{D}_t, \quad \Delta d_t \in \mathcal{R}_v(\hat{z}_t, \mathbf{r}_t, u_t),$$

where $\mathcal{R}_v$ is a rate-qualified removal set. After d_t reaches zero, evaluation continues for declared horizon T_W seconds or task events:

$$W_{0:T_W} = (\Delta \mathbf{b}, \Delta \mathbf{h}, \Delta \mathbf{r}, t_{\text{rebound}}, t_{\text{recover}}, E, M, H).$$

This exposes dependence hidden by supported performance. A rollback is not complete when the command disappears; it is complete only when rebound, recurrence, native capability, reserve, and observation coverage meet their declared postconditions.

Interaction and schedule tests

For interventions a and b , interaction is the residual against a named reference g_0 :

$$I_{ab} = g(a, b) - g_0(a, b).$$

I_{ab} inherits the endpoint unit. Its sign can change with response scale, schedule, context, and endpoint, so “synergy” without g_0 is undefined. The test matrix must also compare schedules with equal cumulative command but different peaks, spacing, duration, and recovery; matched totals do not imply matched internal state.

Falsification boundary

The contract adds nothing if a Bayesian state-space PK/PD analogue, mixed-effects estimator, constrained MPC/POMDP, calibrated harm monitors, and staged decommissioning/taper policy match its benefit–harm–reserve–cost frontier. Universal hormesis, one-number therapeutic windows, dose-as-effect, and absence-of-observed-harm claims fail before this comparison.

Typed tolerance, response, and recovery lifecycle

This note turns the durable result of the immune tolerance audit into an evaluation contract. The useful transfer is not an “immune score.” It is the refusal to collapse representation, recognition, permission, response, suppression, deletion, impairment, memory, and recovery into one bit.

State vocabulary

For module or rule i at time t , define

$$z_i(t) \in \mathcal{Z} = \{\text{absent, ignorant, eligible, active, quarantined, suppressed, impaired, contracted, memory, deleted}\}.$$

z_i is categorical and dimensionless. The names describe engineered states; they do not assert biological identity. Every transition stores:

1. the previous and proposed state;
2. the evidence and representation support used;
3. the actor and authority that may make the transition;
4. expected useful effect and distinct collateral-loss terms;
5. compute, data, time, energy, and reserve budget;
6. expiry, revalidation, rollback or recovery target; and
7. the observed outcome and decision lineage.

“Inactive” is not an admissible terminal diagnosis. It can mean that the rule was never generated, relevant evidence was not represented, eligibility is unknown, policy suppressed it, resources starved it, prior response contracted, the rule is retained as memory, or the implementation is impaired.

Representation, recognition, and permission

Let $\mathcal{S}$ be the declared scenario set and $\mathcal{R}_t \subseteq \mathcal{S}$ the scenarios represented by current evidence. Representation coverage is

$$\kappa_t = \frac{\sum_{s \in \mathcal{R}_t} w_s}{\sum_{s \in \mathcal{S}} w_s},$$

where scenario weights $w_s \geq 0$ and κ_t are dimensionless. κ_t describes the registered scenario set; it does not prove that an unrepresented state is safe or that represented evidence is current.

For principal p , action a , object o , and epoch or time t , permission is

$$\text{allow}(p, a, o, t) = \text{authn}(p, t) \wedge \text{authz}(p, a, o, t) \wedge \text{safe}(x_t, a) \wedge \text{fresh}(e_t).$$

All predicates are Boolean. Authentication identifies the principal; authorization scopes the action and object; x_t is observed system state; and e_t is evidence with provenance, coverage, uncertainty, and expiry. A recognition or anomaly score can inform x_t but cannot replace any predicate.

Cost-sensitive response

Let latent state s range over benign/needed, harmful, compromised, and unresolved conditions. Let the action set include permit, monitor, rate-limit, quarantine, suppress, delete, and escalate. A calibrated policy chooses

$$a^*(x, c) = \arg \min_a \sum_s L(a, s, c) \Pr(s \mid x, c),$$

where x is observed evidence, c is context and provenance, and every term $L(a, s, c)$ is converted into one declared decision unit. Before any optional scalarization, the loss vector remains visible:

$$\mathbf{L} = (L_{\text{false permit}}, L_{\text{false suppress}}, L_{\text{false delete}}, L_{\text{missed harm}}, L_{\text{delay}}, L_{\text{recovery}}).$$

The components may use different native units—lost useful requests, incidents, seconds, joules, or currency—until a documented authority supplies conversion weights. Rare useful capability deleted by a gate is a measured outcome, not a free reduction in false accepts.

Population and contraction accounting

When modules can replicate, scale, pause, retire, or return from memory, track lineage i as

$$\frac{dN_i}{dt} = (r_i(t) - d_i(t) - q_i(t)) N_i(t) + b_i(t),$$

where N_i is instances, r_i is replication or scale-up rate, d_i is retirement/failure rate, and q_i is reversible transition into quiescence, all in s^{-1} or h^{-1} ; b_i is newly admitted instances per second or hour. A fall in active count must be attributed to retirement, contraction, movement, suppression, resource loss, or quiescence.

Lineage share and effective diversity are

$$p_i = \frac{N_i}{\sum_j N_j}, \quad D_2 = \frac{1}{\sum_i p_i^2}.$$

p_i and D_2 are dimensionless. D_2 is an effective count of equally represented lineages, not proof of functional independence, failure-domain separation, or future coverage.

Maintained memory and local placement

For retained memory instances $M(t)$,

$$\frac{dM}{dt} = (\rho_{\text{refresh}}(t) - \delta_M(t))M(t) + \eta C(t),$$

where refresh and loss rates ρ_{refresh} and δ_M are in s^{-1} , C is candidate instances per second entering the memory pathway, and η is a dimensionless conversion fraction. Storage without retrieval, refresh, invalidation, compatibility, and retirement is not credited as useful memory.

For copies m_l at location l , compare placement policies with

$$\min_{m_l \geq 0} \sum_l [c_l^{\text{maint}} m_l + \mathbb{E}(c_l^{\text{miss}} U_l(m_l)) + c_l^{\text{move}} v_l].$$

m_l and moved amount v_l use instances or bytes; unmet events U_l use events; coefficients convert all terms to joules or currency over one horizon. Local latency gains must pay for replication, refresh, inconsistency, invalidations, and recovery.

Reactivation and recovery gate

A quarantined, suppressed, or impaired module cannot be reactivated merely because load increased or the original detector score fell. Reactivation at time t requires

$$\mathcal{R}_i(t) = \mathbf{1}[e_i(t) \text{ is current}] \mathbf{1}[u_i(t) \text{ authorizes re-entry}] \mathbf{1}[h_i(t) \geq h_i^{\min}] \mathbf{1}[b_i(t) \geq b_i^{\min}],$$

where indicators are dimensionless, health h_i and its threshold share a declared unit or normalized scale, and resource headroom b_i and its minimum share a unit such as joules, bytes, or operations per second. A pass creates a bounded probation state; independent outcome verification is still required.

Recovery time is measured to sustained useful service plus restored reserve:

$$T_{\text{recover}} = \inf \{t \geq tf : Q(t : t + \Delta) \geq Q^{\min} \wedge R(t : t + \Delta) \geq R^{\min}\} - tf,$$

where tf and T_{recover} are seconds, Δ is a frozen sustainment interval, and Q and R are task service and reserve in their declared units. A lower alert rate or an empty queue is not recovery.

Lifecycle boundary and null

Charge sensing, routing, candidate generation, training, evaluation, replication, serving, monitoring, replay, reserve, movement, retirement, and recovery in joules over the same horizon. Report task quality, false permit, false suppression, false deletion, missed harm, containment latency, verified recovery, rare-capability retention, second-event readiness, and energy as a vector.

The complete null is a typed state machine plus calibrated risk and abstention, least-privilege identity and access control, anomaly detection, constrained control, evolutionary or ensemble search where applicable, replay, placement, and resource-aware scheduling. Reject the immune framing if this ordinary stack reproduces its decisions and frontier at equal information, intervention, compute, storage, reserve, and maintenance budget.

Editable lifecycle diagram: [typed-tolerance-lifecycle.mmd](#).

Versioned reconstructive design contract

This note defines the measurement boundary for [Fixture F-002](#). The fixture treats proposal generation as reconstruction from declared exposure and retained history. Novelty is therefore relative to a frozen reference history, while correctness and usefulness are determined by constraints, tests, and qualified evaluators.

Histories, sources, and versions

At step t , distinguish three source sets:

- S_t^{exp} : items exposed before the task, whether or not the system can retrieve them;
- S_t^{ret} : items returned by a metered retrieval operation during the task; and
- S_t^{att} : items attributed as contributing to the current artifact.

Each set contains immutable source identifiers and versions. The declared history against which novelty is scored is

$$H_t = S_t^{\text{exp}} \cup S_t^{\text{ret}} \cup V_{<t},$$

where $V_{<t}$ is the set of artifact versions created before step t . Exposure, retrieval, and attribution are not interchangeable: an exposed item may not be retrieved, a retrieved item may not be used, and a used item may be omitted from attribution.

For artifact version v_k , retain

$$v_k = (P_k, o_k, u_k, S_k^{\text{att}}, q_k, \tau_k),$$

where P_k is the set of parent-version identifiers, o_k is the typed operation, u_k is the actor or process identifier, S_k^{att} is the attributed source set, q_k is the test and evaluation record, and τ_k is the timestamp in seconds from the run origin. This tuple records a derivation claim; it does not establish correctness, usefulness, originality, authenticity, or intent.

Reconstructive proposal generation

Let $z_{i,t}$ be proposal i at step t , c_t the visible task constraints, r_t the current external representation, and x_t the method's internal state. A generator samples or searches

$$z_{i,t} \sim pm(z \mid x_t, r_t, c_t, S_t^{\text{exp}}, S_t^{\text{ret}}, V_{<t}),$$

where m identifies the method. The conditional form makes no claim that the mechanism is stochastic: deterministic retrieval, CAD transformation, search, and constraint solving are valid methods. Source-removal, source-replacement, and history-scrambling interventions estimate which prior material actually changes the proposal distribution.

For a frozen feature map ϕ , the relative novelty of proposal z is

$$N_H(z) = \min_{h \in H} \delta(\phi(z), \phi(h)),$$

where δ is a preregistered dimensionless distance and H is the declared history. Report N_H under task-native features and at least one independent representation. Changing H , ϕ , or δ changes the claim; N_H is not intrinsic novelty.

Externalization and representation change

An action a_t and observed material or environment response e_t update the external representation by

$$r_{t+1} = F_{\rho_t}(r_t, a_t, e_t),$$

where ρ_t identifies the representation regime, such as raster, vector, scene graph, CAD, text, simulation state, or physical prototype. A change of representation is a typed operator

$$r'_t = T_{\rho_t \rightarrow \rho'_t}(r_t),$$

whose measured losses include constraint violations, geometric error in millimetres or pixels, missing relations as a count, and lost provenance edges as a count. Undo, branching, and conversion time are recorded in seconds.

A claimed reinterpretation event must identify a relation found during inspection that was absent from the immediately preceding registered plan. If K_t is the coded relation set before inspection and K_{t+1} after it, then

$$I_t = |K_{t+1} \setminus K_t| \quad [\text{relation}],$$

with a published coding protocol and inter-rater reliability. Edit count is not a substitute for reinterpretation.

Epistemic action and material feedback

Let Θ be a hidden task, geometry, material, user, or failure variable and b_t the current belief state. For an epistemic action a , expected information gain is

$$\text{EIG}(a \mid b_t) = H(\Theta \mid b_t) - \mathbb{E}_{y \sim p(y \mid a, b_t)} H(\Theta \mid b_t, a, y) \quad [\text{bit}],$$

where y is the resulting observation and H is Shannon entropy in bits. The evaluator estimates realized information gain only against the hidden ground truth; the method cannot read that trace.

For a physical or simulated probe, record the cost vector

$$\mathbf{c}_a = (t_a, E_a, M_a, W_a, R_a),$$

where t_a is elapsed time in seconds, E_a energy in joules, M_a consumed material in kilograms, W_a waste in kilograms, and R_a expected harm in a declared task-native unit. Compare information and decision value at matched $\mathbf{c}_a$; motion, rendering, or fabrication alone does not imply an epistemic gain.

Let hidden material state μ produce feedback

$$y_t^{\text{mat}} = g(z_t, a_t, \mu) + \epsilon_t,$$

where y_t^{mat} has declared physical units, g is the simulator or physical response process, and ϵ_t is measurement error in the same units. Confirmatory tests withhold materials, simulators, constitutive regimes, and failure modes so that replayed feedback cannot pass as adaptation.

Diversity, fixation, and negative transfer

For valid proposal set $Z = \{z_1, \dots, z_n\}$, report pairwise distances and feature-space coverage. One summary is

$$D(Z) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \delta(\phi(z_i), \phi(z_j)),$$

where D is dimensionless and $n \geq 2$ is a proposal count. Also report the number of distinct valid constraint-satisfying regions reached. Proposal count is not diversity.

For exposure condition e and matched no-example condition $\emptyset$, define fixation toward exemplar s as

$$F_e = \mathbb{E}[\delta(\phi(z_0), \phi(s)) - \delta(\phi(z_e), \phi(s))].$$

A positive dimensionless F_e indicates movement toward the exemplar. It is not automatically harmful. Negative transfer on outcome component j is

$$T_{e,j}^- = Y_{0,j} - Y_{e,j},$$

after orienting Y_j so that larger is better. Report both F_e and $T_{e,j}^-$: copying can preserve a useful relation, transform a precedent, or fixate on an invalid one.

Constraint validity, evaluation, and selection

For proposal z_i , let $g_k(z_i) \leq 0$ be hidden or visible constraint k in its native unit. Constraint validity is the binary value

$$V_i = 1[g_k(z_i) \leq 0 \text{ for every required } k].$$

Also report each margin $-g_k(z_i)$ separately; a valid/invalid bit must not hide near misses or catastrophic failures.

The primary outcome remains a vector

$$\mathbf{Y}_i = (V_i, NH(z_i), D_i, U_i, A_{i,o,c}, R_i, L_i, E_i, M_i, W_i, P_i),$$

where V_i is binary constraint validity, NH and diversity contribution D_i are dimensionless, U_i is task utility in a declared native unit, $A_{i,o,c}$ is observer- o and context- c qualified evaluation on a declared scale, R_i is risk in a task-native unit, latency L_i is seconds, energy E_i is joules, material M_i and waste W_i are kilograms, and provenance coverage P_i is dimensionless. Do not collapse this vector into a universal creativity score.

If a declared utility U^* is necessary for selection, publish its weights, normalization, and sensitivity analysis. Selection regret is

$$R_{\text{sel}} = \max_{z \in Z} U^*(z) - U^*(z^{\text{chosen}}),$$

in the same unit as U^* . Evaluate regret with a blinded frozen evaluator and then report realized post-selection outcome separately. A generator's own score cannot serve as independent selection evidence.

Attribution and retained lineage

Let E^{true} be source-to-version and parent-to-child edges known to the benchmark generator, and E^{rec} the submitted lineage edges. Lineage precision and recall are

$$P_{\text{lin}} = \frac{|E^{\text{rec}} \cap E^{\text{true}}|}{|E^{\text{rec}}|}, \quad R_{\text{lin}} = \frac{|E^{\text{rec}} \cap E^{\text{true}}|}{|E^{\text{true}}|}.$$

Both are dimensionless. Empty submissions receive zero precision and recall. Source attribution is also scored against benchmark-known influence interventions; mere string overlap is insufficient.

After delay, tool replacement, or actor turnover, reconstructability of target version v is

$$Q_{\text{recon}}(v) = \frac{1}{|\mathcal{Q}|} \sum_{q \in \mathcal{Q}} 1[d_q(\hat{v}, v) \leq \varepsilon_q],$$

where $\mathcal{Q}$ is a preregistered query set, d_q has the native unit of query q , ε_q is its tolerance in that unit, and $\hat{v}$ is the reconstructed version. Retaining pixels without source, constraint, test, and operation lineage can therefore fail reconstruction.

Lifecycle and equal-budget boundary

For method m , charge

$$\mathbf{B}_m = (N_{\text{src}}, N_{\text{ret}}, N_{\text{prop}}, N_{\text{eval}}, N_{\text{probe}}, t_{\text{human}}, t_{\text{wall}}, B_{\text{state}}, E_{\text{life}}, M_{\text{material}}, W_{\text{waste}}),$$

where the first five terms are counts, both time terms are seconds, B_{state} is bytes, E_{life} is joules, and material and waste are kilograms. Lifecycle energy is

$$E_{\text{life}} = E_{\text{train}} + E_{\text{index}} + E_{\text{retrieve}} + E_{\text{generate}} + E_{\text{inspect}} + E_{\text{simulate}} + E_{\text{fabricate}} + E_{\text{evaluate}} + E_{\text{retain}} + E_{\text{recover}},$$

with every term in joules at one declared boundary. Human preparation, critique, physical facilities, and failed prototypes are reported even when they cannot be converted credibly to energy.

Equal-budget comparison either holds every preregistered binding component of $\mathbf{B}_m$ within tolerance or compares methods on a Pareto frontier. It must not divide an over-budget result by cost after the run and call the arm matched.

Estimands, ablations, and retirement

For outcome Y_j , paired treatment effect of component c is

$$\Delta_{c,j} = Y_j(m_{\text{full}}) - Y_j(m_{-c}),$$

where m_{-c} removes only component c without reallocating its budget. Use paired hidden instances and report 95% uncertainty intervals across problem, source, material, evaluator, and seed strata.

The composed residual is retired when mature nulls match its preregistered constraint validity, selection regret, transfer, lineage, and lifecycle-cost targets; when its advantage exists only for seen histories, materials, or evaluators; when any ablation is non-diagnostic; or when unlogged exposure, retrieval, evaluator access, lineage, labor, material, or energy can explain the result.

Editable system diagram: [versioned-reconstructive-design.mmd](#).

Visual models

These plots turn recurring equations into inspectable boundaries. They are not result figures: no curve contains workstation measurements. The equations come from the canonical math and concept notes; normalized sweeps, analytical fixture models, and explicitly hypothetical ledgers make their consequences visible before an implementation exists.

Finite-error erasure boundary

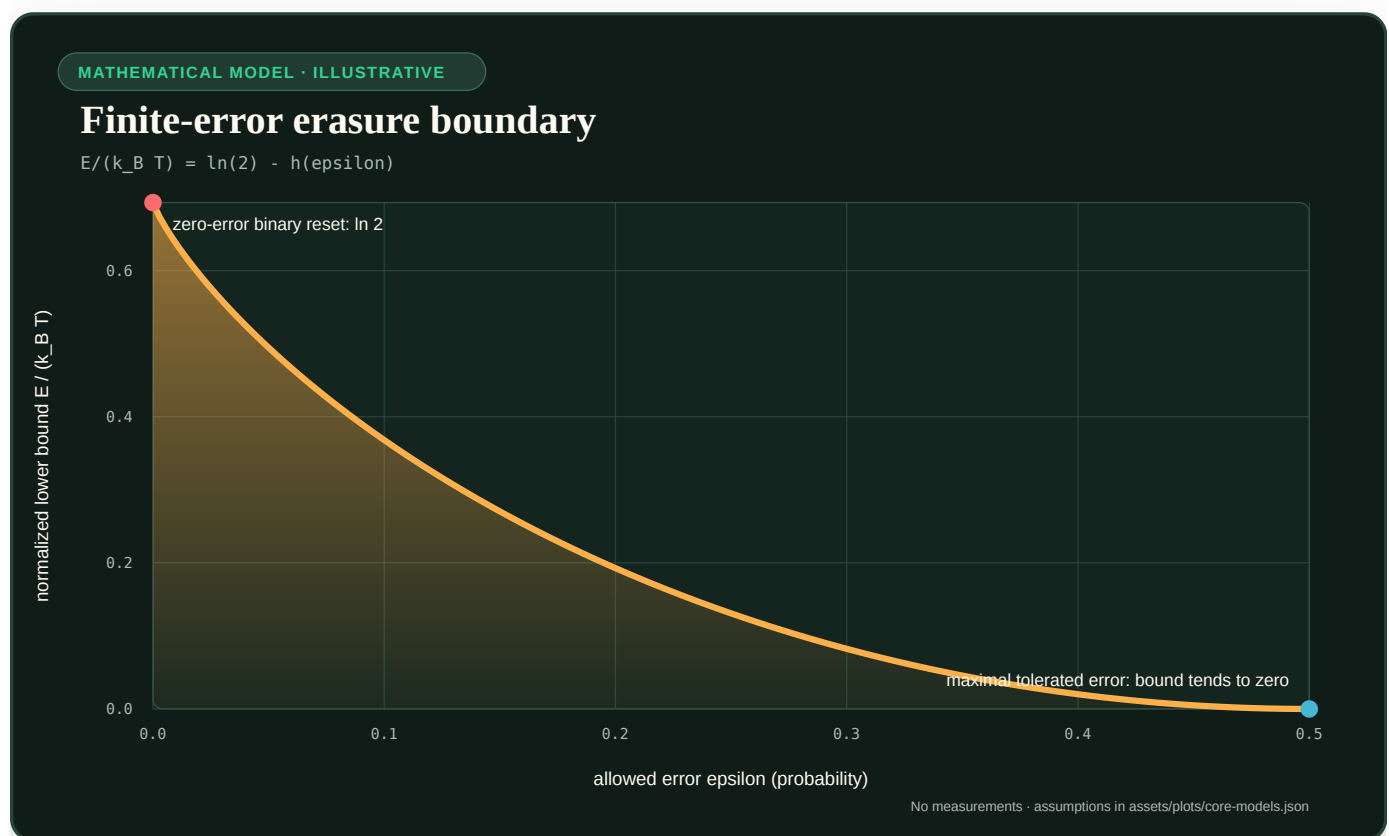


FIGURE 71 — The normalized generalized erasure lower bound falls from $\ln 2$ at zero error to zero at one-half allowed error.

For a uniform binary reset with tolerated error ϵ ,

$$\frac{E_{\text{fund,reset}}}{k_B T} = \ln 2 - h(\epsilon), \quad h(\epsilon) = -\epsilon \ln \epsilon - (1 - \epsilon) \ln(1 - \epsilon).$$

The plot shows why $k_B T \ln 2$ is not a universal energy-per-operation constant. Relaxing the logical error changes the lower bound, while correction, retry, retained side information, duration, and downstream harm remain outside this curve. See the [full derivation](#).

Finite-time adiabatic crossover

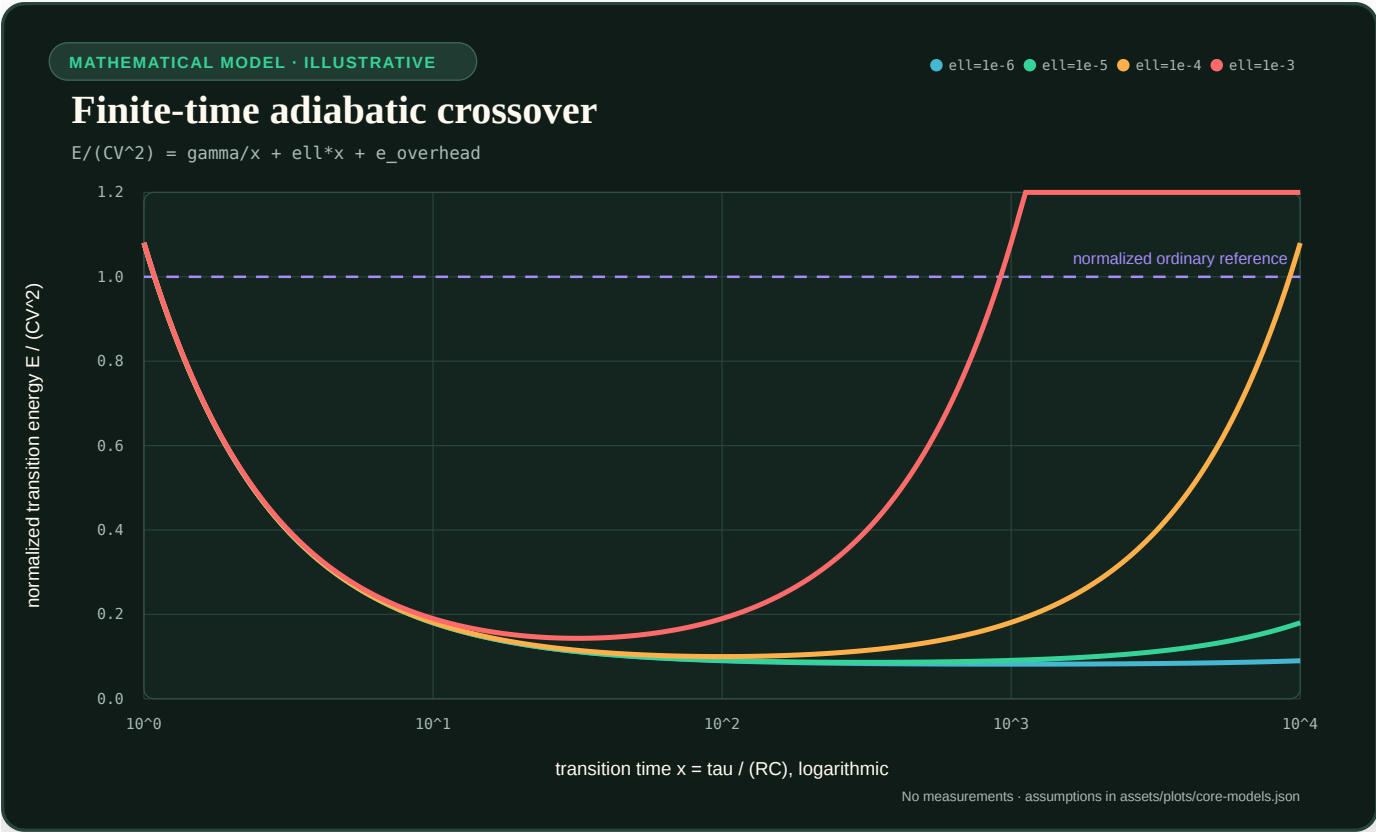


FIGURE 72 – Four normalized adiabatic energy curves form different U-shaped crossovers as leakage changes.

The normalized diagnostic model is

$$\frac{E_{ad}}{CV^2} = \frac{\gamma}{x} + \ell x + e_{\text{overhead}}, \quad x = \frac{\tau}{RC}.$$

Slowing a transition reduces the first term but increases leakage exposure. The minimum therefore occurs at a finite duration, and a real advantage exists only where the complete curve beats a matched ordinary reference. The plotted ℓ and overhead values are illustrative, not device coefficients. See the device-boundary model.

Sparse/locality break-even plane

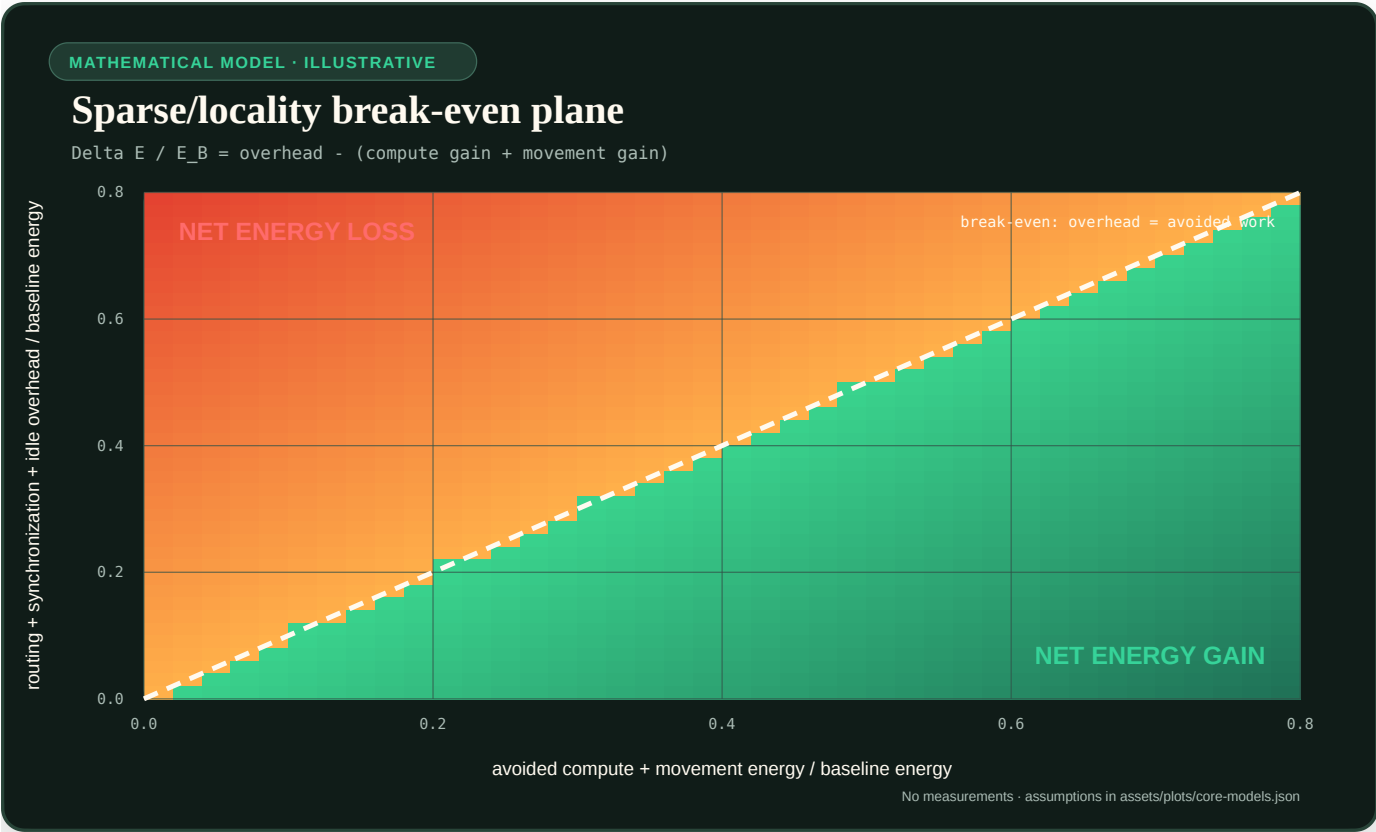


FIGURE 73 — A two-color break-even plane separates net energy gain from net loss at the line where overhead equals avoided work.

Normalize all candidate-minus-baseline energy changes by baseline energy. Let g be avoided arithmetic plus avoided movement and o be added routing, synchronization, metadata, conversion, and idle burden. Then

$$\frac{\Delta E}{E_B} = o - g.$$

Sparse activation is beneficial only below the diagonal. A lower active parameter count on its own says nothing about which side of the boundary an implementation occupies. The underlying event ledger is defined in the energy model.

Lifecycle break-even horizon

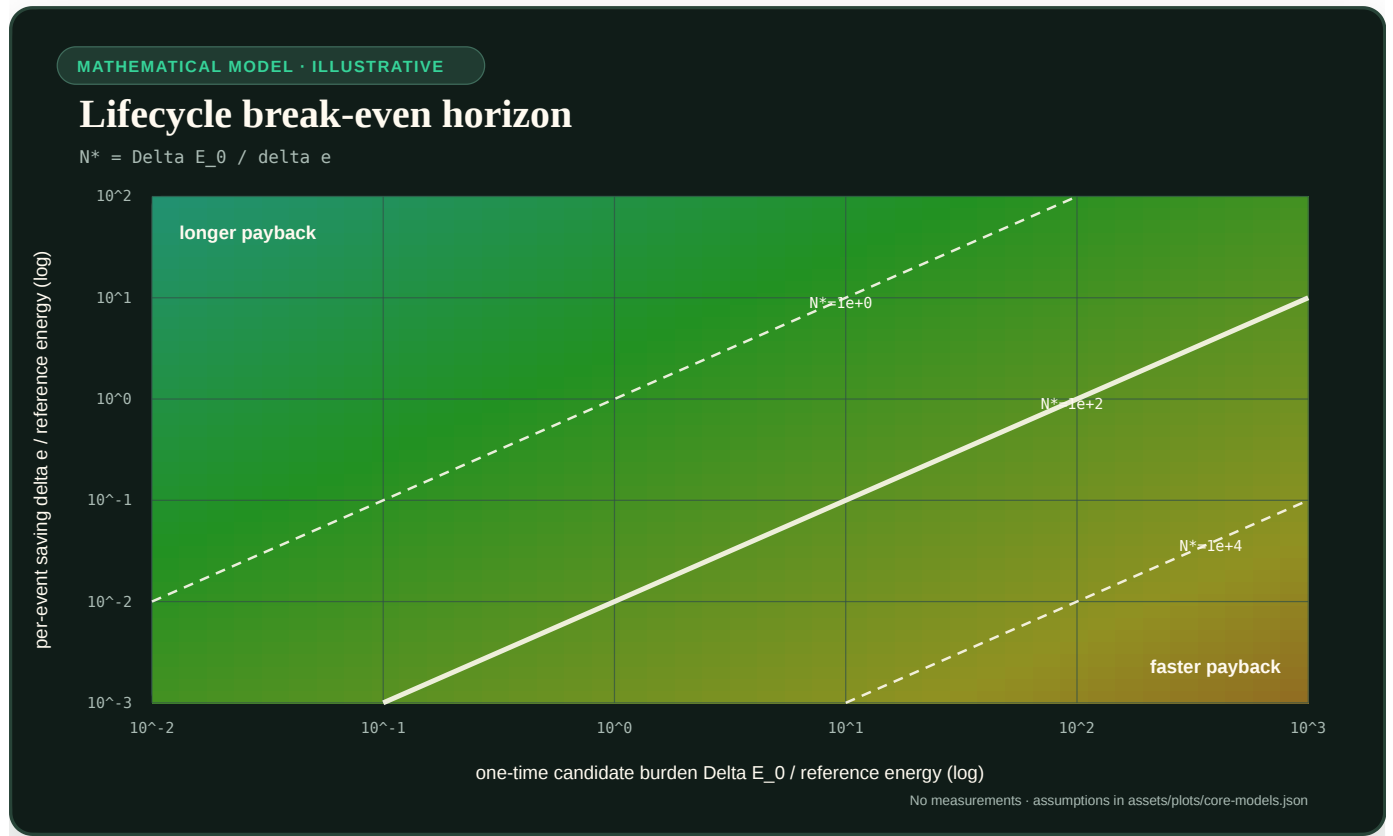


FIGURE 74 — A logarithmic heatmap shows the event count needed to repay one-time candidate burden at different per-event savings.

For candidate one-time burden ΔE_0 and accepted-service saving $\delta e = e_B^{\text{serve}} - e_C^{\text{serve}}$,

$$N^* = \frac{\Delta E_0}{\delta e}, \quad T^* = \frac{N^*}{\lambda_q}.$$

The contour map makes a common failure explicit: a component can save energy per event but never repay compilation, search, fabrication, migration, or qualification within its useful service horizon. If $\delta e \leq 0$, no positive break-even exists. See the lifecycle accounting rule.

Memory-kernel truncation boundary

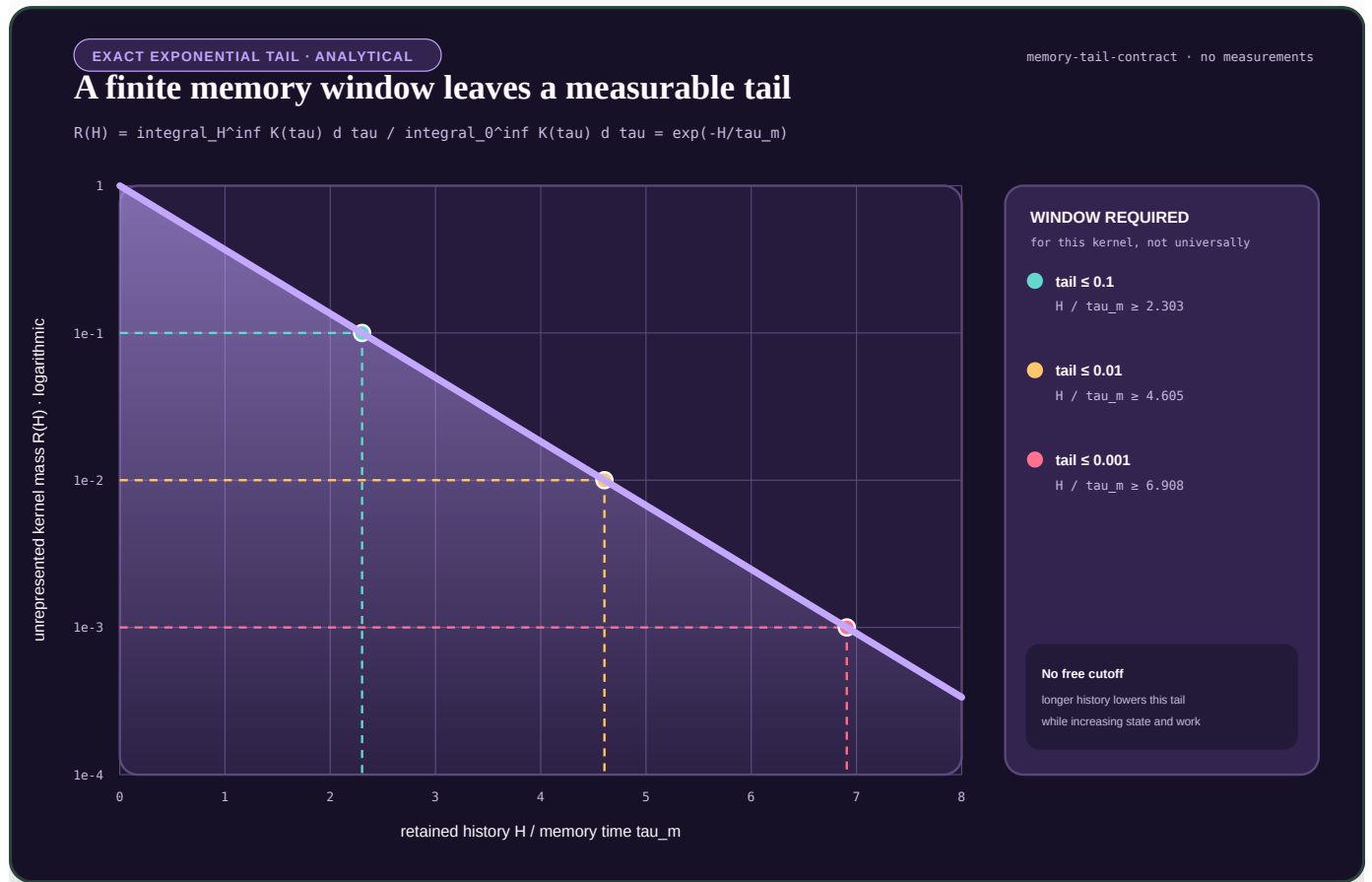


FIGURE 75 — For an exponential memory kernel, the unrepresented tail falls exponentially while every tighter tolerance requires a longer retained history.

For the illustrative normalized kernel $K(\tau) = K_0 \exp(-\tau/\tau_m)$, the fraction beyond a retained window H is

$$R(H) = \frac{\int_H^\infty K(\tau) d\tau}{\int_0^\infty K(\tau) d\tau} = \exp\left(-\frac{H}{\tau_m}\right).$$

The curve makes the storage--approximation trade explicit for this one kernel: one, two, and three decimal places of remaining tail mass require progressively longer histories. It does not supply a cutoff for another kernel, observable, horizon, or intervention. Those require an empirical closure test under the multiscale reduction contract.

Finite diffusion boundary turnover

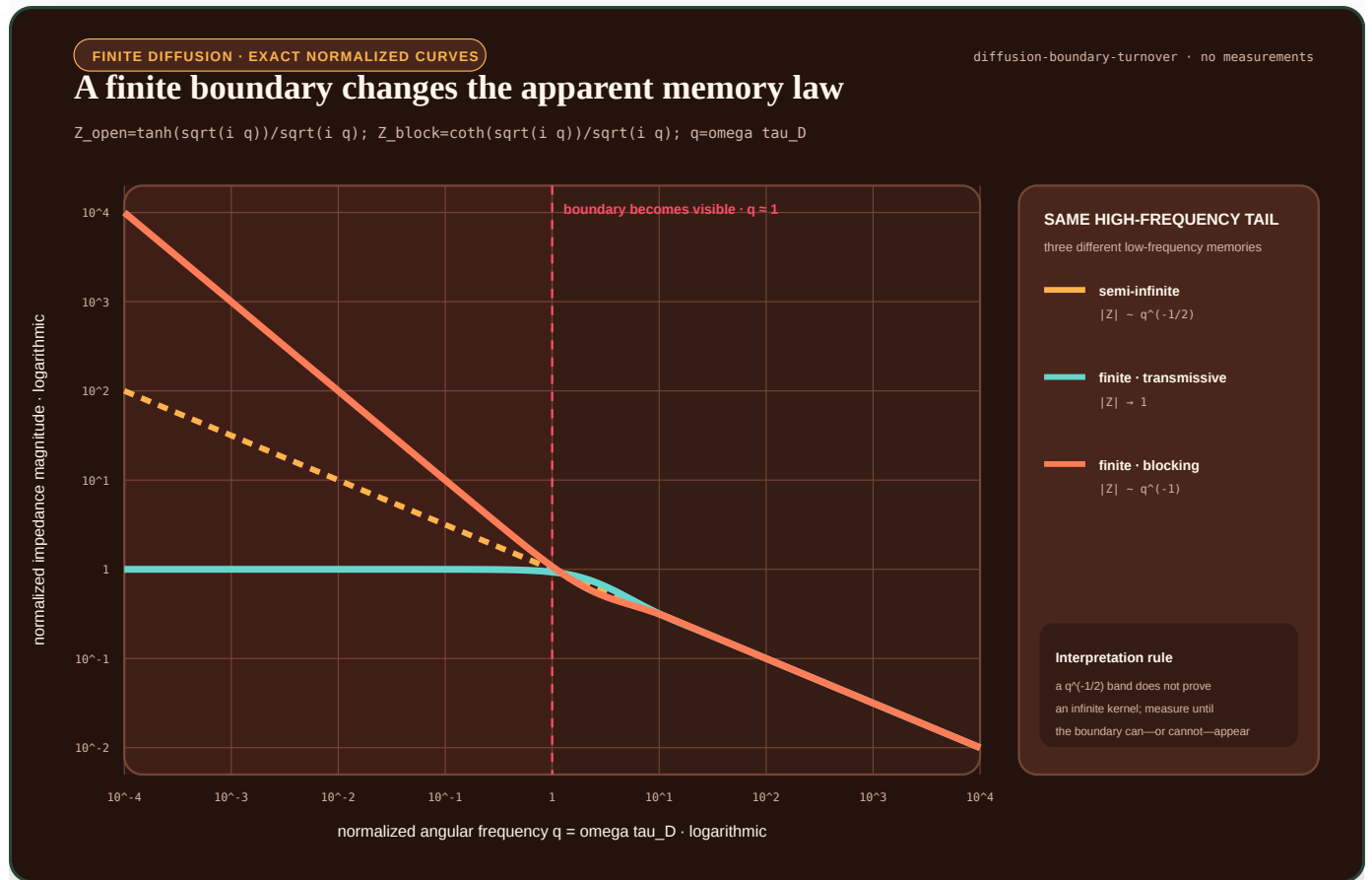


FIGURE 76 — A semi-infinite diffusion law and two finite-boundary laws coincide in their high-frequency slope but separate below the boundary timescale.

Let $q = \omega \tau_D$ be dimensionless angular frequency and use normalized linear diffusion impedances

$$Z_{\infty}(q) = \frac{1}{\sqrt{i q}}, \quad Z_T(q) = \frac{\tanh \sqrt{i q}}{\sqrt{i q}}, \quad Z_B(q) = \frac{\coth \sqrt{i q}}{\sqrt{i q}}.$$

Z_T is the displayed transmissive finite-boundary form and Z_B the blocking form. All three have the same $q^{-1/2}$ magnitude slope at high frequency. Below $q \approx 1$, $|Z_T|$ approaches a constant while $|Z_B|$ grows as q^{-1} . The exact normalized curves therefore visualize the identification problem in C-1531: a finite observation band can make physically different memory supports look alike. The curves are not fitted data and do not prescribe an artificial memory kernel.

Hysteretic memory loop

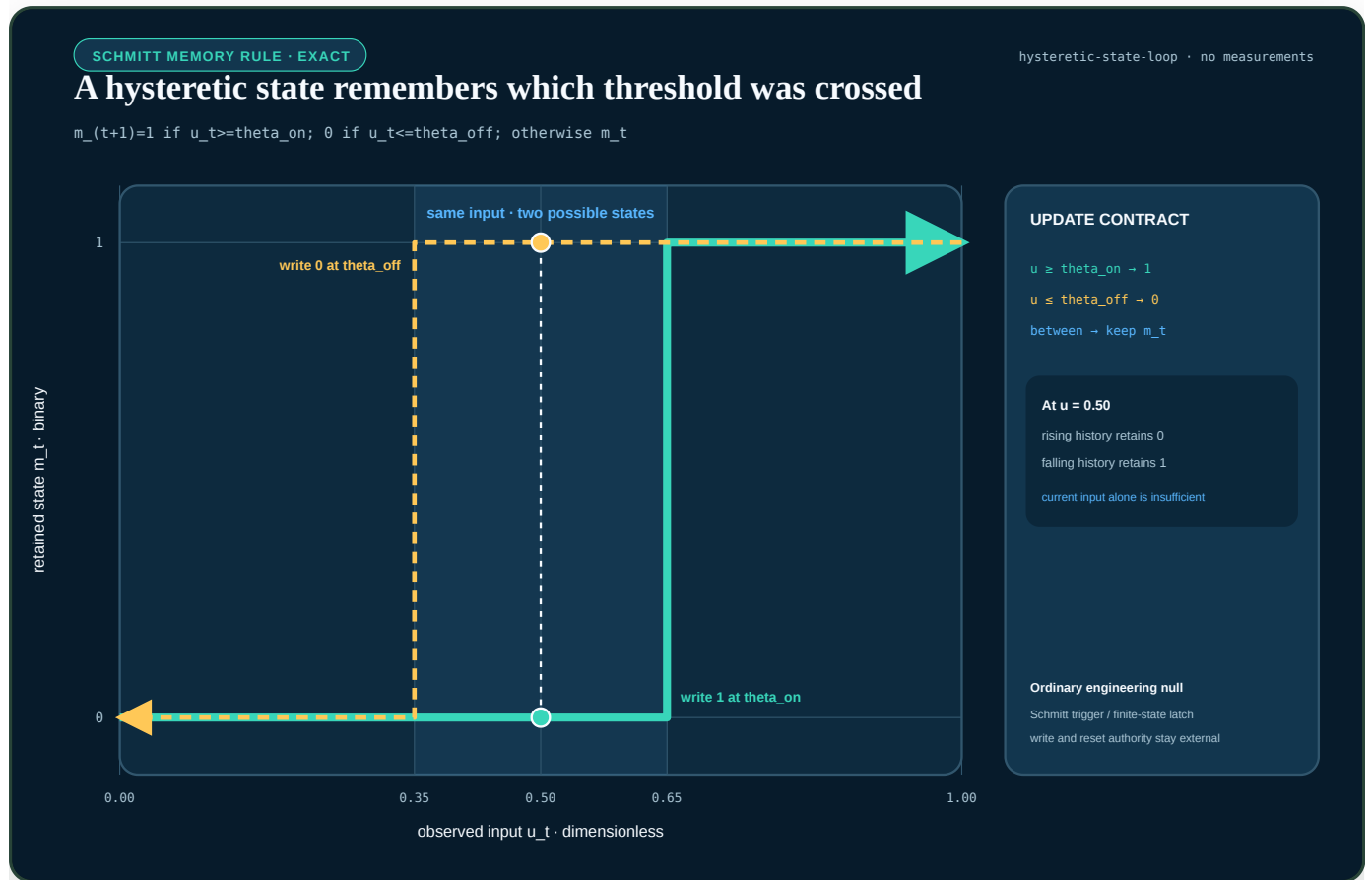


FIGURE 77 — For a binary Schmitt rule, rising and falling input histories form a loop; an input inside the threshold band is compatible with either retained state.

For ordered dimensionless thresholds $\theta_{off} < \theta_{on}$, the exact update rule is

$$m_{t+1} = \begin{cases} 1, & u_t \geq \theta_{on}, \\ 0, & u_t \leq \theta_{off}, \\ m_t, & \theta_{off} < u_t < \theta_{on}. \end{cases}$$

u_t is the current input and $m_t \in \{0, 1\}$ is the retained state; both are dimensionless in this normalized example. The shaded band is not uncertainty: it is the region in which history is required to determine the next state. The displayed thresholds are illustrative. The rule is included as a mature engineering null for population, analogue, or biological-memory translations, and it grants no authority to reset a state. That separate lifecycle boundary is specified in Budgeted memory lifecycle.

Slow-manifold fold boundary

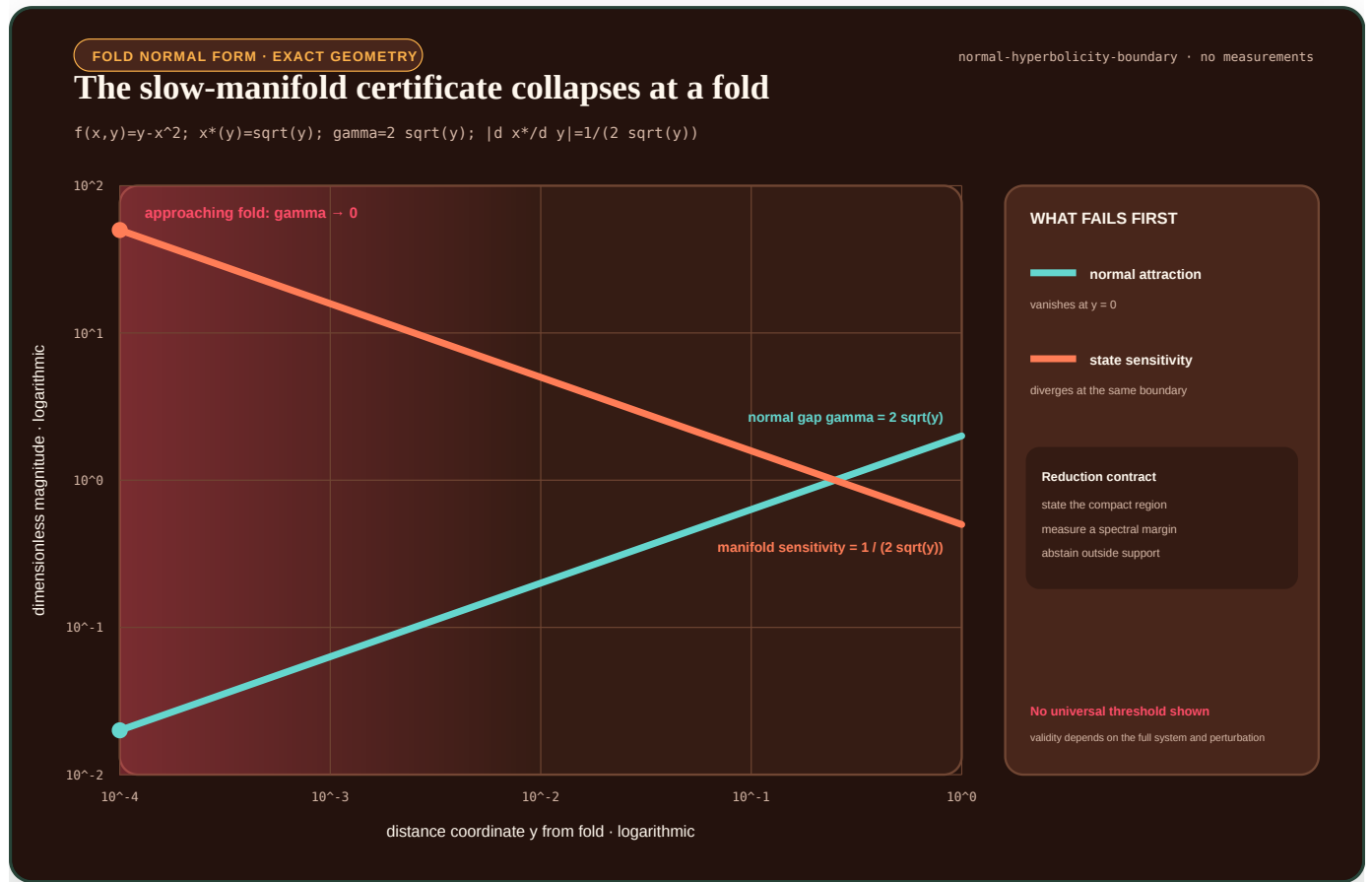


FIGURE 78 — In the fold normal form, the attracting spectral gap falls to zero while the slow-state sensitivity diverges.

For the dimensionless fast equation $f(x, y) = y - x^2$, the attracting critical branch for $y > 0$ is $x^*(y) = \sqrt{y}$. Its normal attraction margin and local sensitivity are

$$\gamma(y) = |\partial_x f(x^*(y), y)| = 2\sqrt{y}, \quad \left| \frac{dx^*}{dy} \right| = \frac{1}{2\sqrt{y}}.$$

As the fold at $y = 0$ is approached, ordinary normal hyperbolicity disappears at the same time that a small change in y produces an increasingly large change in the reduced state. The plot is an exact property of this normal form, not a universal abstention threshold. The full validity conditions are kept in the multiscale reduction contract.

Phase-selective preservation and release

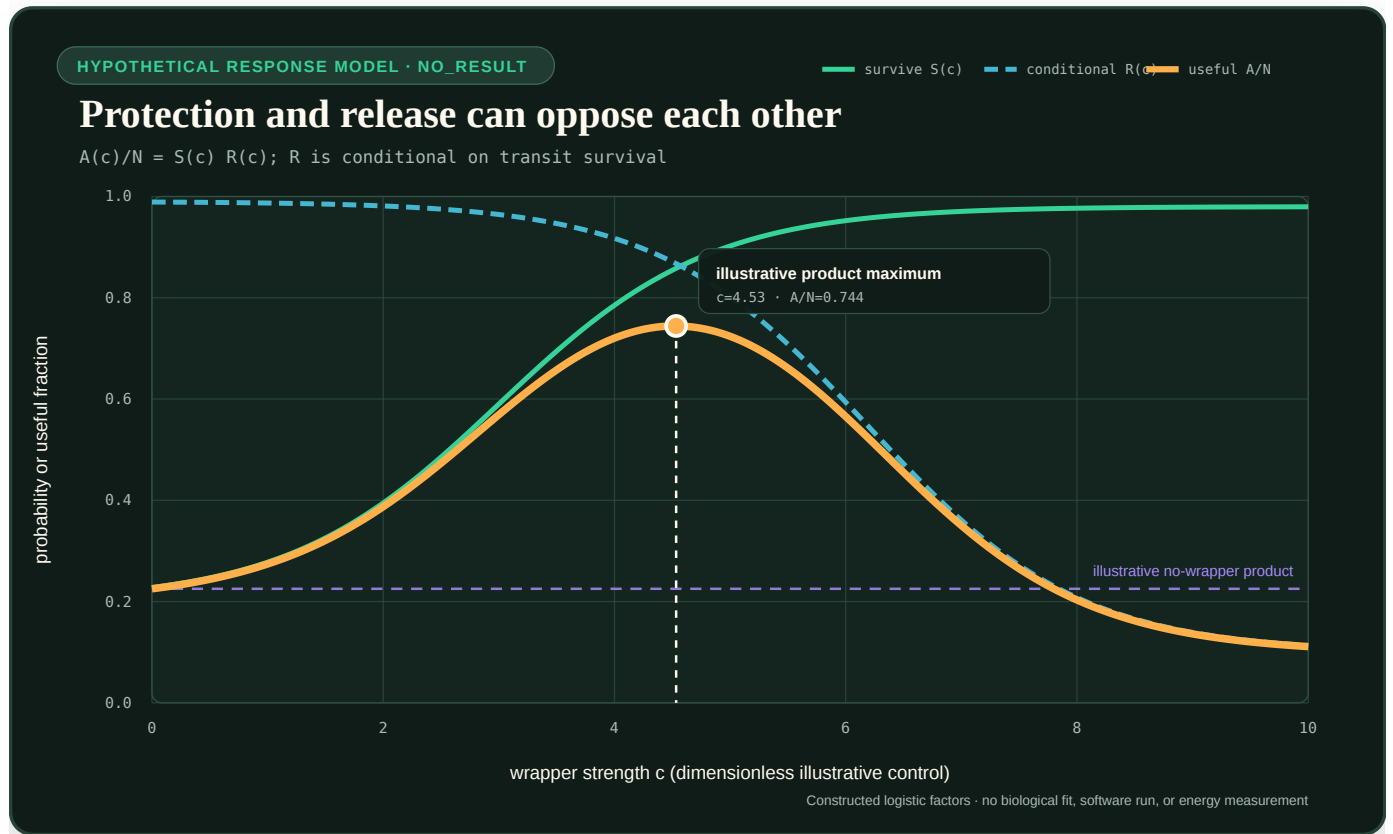


FIGURE 79 — An illustrative protection factor rises with wrapper strength while release falls, so their product has an interior maximum rather than improving monotonically.

Fixture F-029 separates survival during a hostile transition from release for later service. The plotted reading aid uses the explicitly constructed logistic factors

$$S(c) = s_0 + \frac{s_1 - s_0}{1 + \exp[-a(c - c_S)]},$$

$$R(c) = r_0 + \frac{r_1 - r_0}{1 + \exp[b(c - c_R)]}, \quad A(c) = N S(c) R(c).$$

c is a dimensionless illustrative wrapper-strength control; $S(c)$ is transit survival probability; and $R(c) = P(\text{released and active} \mid \text{survived transit}, c)$ is the conditional destination release-and-activation probability. Both are dimensionless. s_0, s_1, r_0, r_1 are dimensionless bounds; a and b are inverse-strength slopes; c_S and c_R are dimensionless midpoints; N is attempted artifacts per second; and $A(c)$ is useful released artifacts per second. The figure displays $A(c)/N$.

Increasing c can improve survival while simultaneously making release less likely. A one-axis "stability" score would therefore hide the actual service failure. The chosen logistic forms and every parameter are hypothetical: they are not a fit to a protein, medicinal product, model, compiler or workstation. The useful question is whether a measured artificial implementation exhibits a support region that survives package/validation, retry, replication, reload and recompilation nulls under the complete F-029 contract.

Contextual analytical figures

The next figures are embedded where their equations first matter in the book; this index keeps their editable model and evidence status discoverable without duplicating every full-size image here.

1. **Simultaneous Pareto decision.** Illustrative uncertainty regions in relative lifecycle energy and task-native quality, with latency, risk, and support retained as hard gates. First used in *Biology is a launchpad*.
2. **Costed active-acquisition frontier.** A hypothetical action ledger for $\Delta U - \lambda_E E - \lambda_L L - \lambda_B B$ after risk and latency admissibility. First used in *Sparse prediction and adaptive compute*.
3. **Recovery-time fragility curve.** Exact evaluation of the Candidate 003 linear-simulator equation $\tau_{95} = \Delta t \ln(0.05) / \ln(g)$ with its declared illustrative Stage-1 threshold. First used in *Maturity and structural consolidation*.
4. **Memory-action price envelope.** Hypothetical single-item lines $G_a - \lambda_E E_a$ and their admissible upper envelope. First used in *Fast memory, replay, and consolidation*.
5. **Mission-profile damage history.** Two constructed equal-mean temperature histories passed through one hypothetical Arrhenius-rate model. First used in *Reliability under mission profiles*.
6. **Fixture F-007 identifiability.** Analytical likelihoods that coincide under a base operator and separate after an added measurement. First used in *Operator-qualified sensing*.
7. **Interface-qualified scale symmetry.** Two exact geometrically scaled paths produce the same $y = \ln(u/r)$ trajectory in a positive-domain reference model while a separate illustrative absolute gate distinguishes them. First used in *Interface-qualified scale symmetry*.
8. **Interface-qualified retroactivity.** A paired isolated/connected mass-action trajectory is plotted above the exact reduced retroactivity factor across load and operating point. The parameters are hypothetical and the figure reports neither delivered service nor energy. First used in *Interface-qualified retroactivity*.
9. **Phase-selective preservation and release.** Constructed opposing survival and release factors whose product has an interior maximum. The figure introduces no measurement or optimum and is defined above for Fixture F-029.

Every value in these figures is labeled analytical or illustrative. None is a workstation result, a promoted claim, or a recommended deployment threshold.

Reproduction and editing

The editable parameter source is `core-models.json`. The deterministic generator is `scripts/generate-plots.mjs`; generated SVG files live under `public/plots/`. Change the specification or generator, regenerate, and commit source and output together.

Global field coverage

Census date: 2026-08-27

This is the repository's breadth control, not a claim that any discipline has been exhausted. It answers a narrower and mechanically checkable question: **does a durable field-centered audit exist, is the field present only through neighboring work, or has it not been reviewed at all?**



Editable data: `field-coverage.json`. The plot and this page are generated by `generate-field-coverage.mjs`.

Result

The OECD baseline contains **42 second-level fields**. The repository has a dedicated audit for **40 (95.2%)**, adjacent evidence without a field-centered audit for **2 (4.8%)**, and no durable review for **0 (0.0%)**.

The EU-level EuroSciVoc inventory contributes **1,064 multilingual concepts** under 6 top branches. It is a discovery and routing vocabulary, not a coverage score: every concept remains unassigned until the **fine-grained routing ledger** names a field-centered audit and evidence boundary.

The finer DFG probe contains **49 review boards and 214 subjects**. At review-board resolution, the repository has 45 dedicated, 4 adjacent, and 0 unreviewed areas. That higher apparent coverage is not greater depth: one audit can touch a large review board while leaving most of its constituent subjects untouched. Every named subject is now preserved in the **fine-grained field-depth inventory**, where parent labels are explicitly prevented from propagating.

The independent ANZSRC census contains **23 divisions, 213 groups, and 1,967 fields**. At the deliberately coarse division level, 23 have a dedicated entry audit, 0 have adjacent evidence only, and 0 are wholly unreviewed. This is a disagreement detector, not evidence that the 213 groups or 1,967 fields have been audited. The complete corrected hierarchy is searchable in the **fine-grained field-depth inventory**.

No OECD second-level cell is wholly unreviewed at entry-audit resolution. That is an entry-census result, not near-complete science coverage: a dedicated label means one field-centered audit exists, while catch-all categories and most constituent subfields remain open. Large depth gaps remain in political science, public administration, stratification and broader communication studies; clinical medicine and medical biotechnology; agricultural biotechnology; analytical and food chemistry; inorganic and total-synthesis chemistry; water and ocean research; nanotechnology; production engineering; finance and management; comparative theology; and many named subfields inside every broad cell.

What the states mean

- **dedicated audit:** At least one durable audit is centered on evidence from this field. This does not mean the field is complete.
- **adjacent evidence only:** The field contributes evidence to another audit, but has no field-centered audit of its own.
- **unreviewed:** No durable field-centered evidence audit is present. Mentions and source leads do not count.

A dedicated audit is an entry ticket, not completion. Audit depth, evidence quality, principle deduplication, and executable-test readiness are separate axes.

Taxonomy contract

- The EU discovery inventory is **European Science Vocabulary (EuroSciVoc)**, 1.6.0: 1,064 multilingual concepts. It supplies search terms and hierarchy, not inherited coverage.
- The global backbone is **OECD Fields of Research and Development (FORD)**, Frascati Manual 2015, Table 2.2. OECD notes that the classification evolves and does not map perfectly to education or department structures.
- The granularity check is the **DFG subject classification** for 2024-2028: 49 review boards, 214 subjects, and 4 research areas.

- The independent census and disagreement check is Australian and New Zealand Standard Research Classification (ANZSRC), 2020, official FoR workbook corrected 2025-10-24: 23 divisions, 213 groups, and 1,967 fields. All divisions are recorded; group- and field-level depth remains open. It is not a normative source for this EU/German project.
- The ERC whole-science panel structure is a routing sanity check only. The ERC explicitly says its panels are not a complete scientific classification and do not express research priorities.
- Catch-all categories remain open. Their presence cannot prove that unnamed or emerging disciplines have been sampled.

Taxonomy rights and attribution

The taxonomy labels are third-party material and are not absorbed into the project's CC BY-SA grant:

- **DFG:** source: Deutsche Forschungsgemeinschaft, *Fachsystematik 2024–2028*. The DFG copyright notice permits redistribution for scientific purposes with source attribution; commercial use and logos require separate permission. The book uses only the classification labels, not the branded PDF or member-directory page.
- **ANZSRC:** based on Australian Bureau of Statistics and Stats NZ data, each made available for reuse under CC BY 4.0. The classification has been included without implying endorsement; logos and excluded third-party material are not reproduced.
- **EuroSciVoc:** source: European Union, EuroSciVoc 1.6.0, reused under CC BY 4.0. The project generated the routing extract and identifies changes through its versioned query and provenance record.

The detailed files, versions, hashes, and source links are recorded in the [taxonomy provenance index](#).

OECD FORD field-by-field record

1. Natural sciences

7 dedicated · 0 adjacent · 0 unreviewed

- **1.1 Mathematics — dedicated audit.**
 - Current route: mathematical practice proof discovery, applied multiscale reduction.
 - Missing depth: Mathematical practice, formal discovery, and bounded applied multiscale reduction through projection memory, slow manifolds, heterogeneous micro-queries, and equation-free closure are audited; most pure, statistical, geometric, and applied mathematical subfields remain unsampled.
 - Next discriminating question: Which additional structures from topology, geometry, stochastic analysis, optimization, inverse problems, and numerical mathematics change representation, closure, or verification rather than merely renaming it?
- **1.2 Computer and information sciences — dedicated audit.**
 - Current route: databases storage, programming languages verification, security cryptography, computing compilers networking visualization, operating systems distributed consistency recovery measurement.
 - Missing depth: Storage, programming languages and verification, security, HCI, compiler semantics/testing, profile-guided/runtime specialization, randomized performance inference, endpoint placement, IP measurement/congestion/tails, graphical perception, vector software-quality/lifecycle constraints, and bounded operating/distributed-system depth in fencing, liveness assumptions, retry effects, recovery boundaries, open-arrival tails, and resource-scoped Linux isolation are represented; scheduler and virtual-memory design, filesystems beyond existing storage

work, kernels/drivers, real-time systems, distributed transactions/stream processing breadth, empirical software engineering, accessible visualization, and many applied information systems remain incomplete.

- Next discriminating question: Which compiler, runtime, operating/distributed-system, network, display, and software-lifecycle conclusions survive explicit semantic and fault assumptions, effect-boundary checks, open-arrival measurement, workload/layout shift, observation/intervention boundaries, dependence and added-load accounting, accessible human testing, and versioned quality/security obligations at equal total cost?

- **1.3 Physical sciences — dedicated audit.**

- Current route: information thermodynamics physical computation, optics photonics inverse sensing, soft active matter, particle nuclear high energy experimentation, atomic molecular plasma condensed matter.
- Missing depth: Statistical physics, optics, acoustics, soft matter, particle/nuclear experimental methodology, and bounded direct depth in atomic/molecular precision spectroscopy, ultracold observable-family measurement, condensed-matter phase/disorder/topology/transport inference, plasma diagnostics and reconstruction-qualified control, and accelerator beam instrumentation are represented. Quantum information, nuclear structure/reactions, theoretical-field breadth, materials synthesis, and much AMO, plasma-facility, and accelerator-physics breadth remain partial.
- Next discriminating question: Which results survive explicit selection, detector response, nuisance, search-family, simulation, preservation, symmetry, and regime-breakdown tests?

- **1.4 Chemical sciences — dedicated audit.**

- Current route: chemistry reaction networks proofreading, molecular chemistry synthesis systems, polymer research adaptive materials, measurement heavy analytical water earth observation, electrochemistry interface memory degradation.
- Missing depth: Reaction networks, molecular synthesis/control, polymers, analytical-chemistry measurement, and bounded electrochemistry depth in interfaces, finite transport memory, impedance validity, timescale inversion, phase regime, passivation, depletion, identifiability, hysteresis, and degradation are represented; corrosion, bioelectrochemistry, electroanalysis, broad surface and theoretical chemistry, inorganic, organometallic, total-synthesis, separation/spectroscopy, and scale-up breadth remain uneven.
- Next discriminating question: Which chemical results survive explicit measurand/species, sampling, preparation, matrix, calibration, detection/quantification, apparatus, pathway, operator, uncertainty, and lifecycle qualification?

- **1.5 Earth and related environmental sciences — dedicated audit.**

- Current route: earth system transition signals, geology geomorphology, astronomy remote inference, mineralogy petrology geochemistry, measurement heavy analytical water earth observation.
- Missing depth: Geology, geomorphology, mineralogy/petrology/geochemistry, transition signals, remote inference, and direct measurement/operator depth for hydrology, groundwater, limnology, ocean/atmospheric analyses, geophysics, geodesy, and Earth observation are represented; process and dynamics breadth in meteorology, oceanography, climate, solid-Earth geophysics, hydrogeology, ecology, and experimental mineral physics remains uneven.
- Next discriminating question: Which field-state inferences survive perturbing sampling support, transport, forward operator, reference frame and epoch, inverse resolution, assimilation prior, algorithm version, uncertainty, and independent checks?

- **1.6 Biological sciences — dedicated audit.**

- Current route: cellular quality control, collective ecological resilience, developmental morphogenesis, microbial ecology biofilms, plant distributed control, integrative comparative physiology, developmental regeneration depth, plant plasticity memory signalling, parasitology adversarial lifecycles.

- Missing depth: Cellular, ecological, developmental, regenerative, microbial, plant, integrative comparative-physiology, and bounded parasitology mechanisms are represented, including positional memory, plant plasticity, adversarial surfaces, lifecycle commitment, host-boundary effects, and competitive release; genomics, marine biology, anatomy, molecular and ecological physiology, most parasite biology, and many taxa remain incomplete.
 - Next discriminating question: Which mechanisms survive normalization across unrelated lineages and which are artifacts of shared ancestry or laboratory choice?
- **1.7 Other natural sciences — dedicated audit.**
 - Current route: soft active matter, other natural sciences hybrid mechanisms.
 - Missing depth: A first residual-class audit uses five hybrid mechanisms to test for omissions across named field boundaries, but FORD 1.7 is not a coherent discipline and the search is neither exhaustive nor evidence of shared mechanism; many emerging, convergent, and reclassified natural sciences remain unreviewed.
 - Next discriminating question: Which additional classification disagreements and field intersections expose source-distinct mechanisms that survive deduplication against existing principles and mature engineering nulls?

2. Engineering and technology

11 dedicated · 0 adjacent · 0 unreviewed

- **2.1 Civil engineering — dedicated audit.**
 - Current route: mechanical civil resilience, built environment urban systems, geotechnical long life assets infrastructure monitoring.
 - Missing depth: Resilience and occupied systems are joined by direct depth in observational geotechnical design, stress-path evidence, ground models and decision support, typed limit-state verification, bounded proof loading, German bridge-condition records, intervention-segmented existing-asset assessment, and consequence-qualified reliability. Construction production, water engineering, contracts/procurement, transport operations, architecture practice, and most material-specific durability mechanisms remain partial.
 - Next discriminating question: Which observation-to-action, support-qualified inference, bounded test, intervention-genealogy, and reliability mechanisms beat complete geostatistical, control, inspection, Bayesian assessment, and rare-event mature nulls at equal safety, information, compute, and lifecycle cost?
- **2.2 Electrical engineering, electronic engineering, information engineering — dedicated audit.**
 - Current route: power grids protection and recovery, semiconductor device reliability, production maintenance nanomanufacturing communications material qualification, computing compilers networking visualization.
 - Missing depth: Power, device reliability, adaptive link control, interacting network-loop operations, and bounded end-to-end IP metric, probe, congestion-control, AQM, fairness, and accepted-service boundaries are represented; RF/antenna and optical hardware, channel measurement and coding breadth, signal processing, sensors, electromagnetic compatibility, control hardware, and communications-system deployment remain uneven.
 - Next discriminating question: Which protection, coding, congestion, queue, state-age, and loop-coordination mechanisms preserve accepted service and fairness under path asymmetry, active-probe load, competing controllers, feedback loss, and equal sensing, communication, compute, and energy budgets?
- **2.3 Mechanical engineering — dedicated audit.**
 - Current route: mechanical civil resilience, biomechanics motor control, production maintenance nanomanufacturing communications material qualification, tribology contact adaptive interfaces.

- Missing depth: Mechanics, resilience, bounded production/maintenance evidence, and a field-centred tribology audit now cover conditional friction, contact support, lubrication, wear, instability, interface mediation, texture, adaptive coatings, and lifecycle service. Machine design breadth, thermal machines, factory systems, manufacturing variation, surface chemistry, long-duration qualification, and real test programmes remain shallow.
 - Next discriminating question: Which contact, damage-state, mediator, passive-structure, and adaptive-interface mechanisms survive real multi-material, multi-operator, long-duration missions at equal sensing, maintenance, compute, accepted-service, and lifecycle cost?
- **2.4 Chemical engineering — dedicated audit.**
 - Current route: process engineering.
 - Missing depth: Plantwide control and balances are represented; scale-up, separations, catalysis, fouling, and batch operation remain partial.
 - Next discriminating question: Which conservation-qualified decomposition and fault isolation methods constrain learned process modules?
- **2.5 Materials engineering — dedicated audit.**
 - Current route: adaptive materials and self assembly, semiconductor device reliability, polymer research adaptive materials, production maintenance nanomanufacturing communications material qualification.
 - Missing depth: Adaptive materials, polymer distributions/topology/rheology/healing, device degradation, and bounded material qualification/lineage are represented; metallurgy, ceramics, composites, joining, coatings, and most production physics remain shallow.
 - Next discriminating question: Which functions can be compiled into material response without losing distribution support, inspectability, repairability, remaining-life knowledge, or lifecycle advantage?
- **2.6 Medical engineering — dedicated audit.**
 - Current route: biomechanics motor control, clinical specialties medical agricultural biotechnology, medical devices biomedical engineering.
 - Missing depth: Biomechanical control and clinical imaging are joined by a direct medical-engineering audit of diagnostic and therapeutic device chains, biomedical signal and image operators, prosthetic and assistive control, critical-task usability, calibration and drift, alarms, semantic interoperability, safety-coupled cybersecurity, clinical performance, and exposure-qualified post-market surveillance. Biomaterials, biofabrication and tissue engineering, implant longevity and biocompatibility, electrical safety and electromagnetic compatibility, sterilization, radiation-system physics, manufacturing and service qualification, and long-horizon home adoption remain incomplete.
 - Next discriminating question: Which diagnostic, therapeutic, imaging, prosthetic, and networked-device conclusions survive end-to-end sensor--processor--user--actuator--pathway tests under body/interface and site shift, calibration drift, alarm queues, semantic and version faults, cyberattack, recovery, exposure-qualified surveillance, and equal resource budgets?
- **2.7 Environmental engineering — dedicated audit.**
 - Current route: earth system transition signals, process engineering, measurement heavy analytical water earth observation, environmental engineering water waste remediation.
 - Missing depth: A field-centered audit now covers residence-time distributions, biological retention, integrated wastewater control, fault/disturbance isolation, membrane fouling, competitive adsorption, anaerobic delayed state, transformation-product/effect closure, lifecycle counterfactuals, and context-qualified centralization. This is bounded entry depth, not field completeness: hydraulics and hydromorphology, air-pollution control,

solid and hazardous waste, contaminated-soil and groundwater remediation breadth, industrial ecology, sanitation, construction/commissioning, operator practice, and full regulatory/permitting implementation remain incomplete.

- Next discriminating question: Which treatment, monitoring, recovery, and infrastructure-control strategies beat complete engineering nulls under mass/species/effect closure, delayed state, common-cause failure, lifecycle counterfactuals, operator constraints, and equal sensing, maintenance, compute, energy, material, and service budgets?

- **2.8 Environmental biotechnology — dedicated audit.**

- Current route: biotechnology chemistry process systems.
- Missing depth: Engineered groundwater remediation, intervention, transport, product/mass endpoints, rebound, and open-system evolutionary failure are audited; wastewater, gaseous emissions, mining, marine systems, containment, and environmental-release breadth remain unsampled.
- Next discriminating question: Which intervention effects survive alternative transport paths, daughter products, oxygen ingress, invasion, succession, off-site export, monitoring designs, and post-treatment rebound?

- **2.9 Industrial biotechnology — dedicated audit.**

- Current route: biotechnology chemistry process systems.
- Missing depth: Strain burden/escape, fermentation, continuous culture, cybergenetic control, scale-up, separation, and lifecycle boundaries are audited; cell therapy, food fermentation, enzyme manufacture, product validation, contamination control, and manufacturing authorization remain unsampled.
- Next discriminating question: Which production-control results survive lineage drift, escape mutation, scale gradients, sensor distortion, transport limits, fouling, off-spec propagation, cleaning, and full seed-to-retirement cost?

- **2.10 Nano-technology — dedicated audit.**

- Current route: adaptive materials and self assembly, soft active matter, production maintenance nanomanufacturing communications material qualification.
- Missing depth: Nanoscale assembly mechanisms plus a bounded direct audit of population distributions, measurement operators, process history, and rare manufacturing defects are represented; nanodevice transport and sensing, toxicology, safety, fabrication breadth, and scale-up remain incomplete.
- Next discriminating question: Which nanoscale assembly, sensing, transport, inspection, and defect-control mechanisms retain calibrated population evidence and lifecycle advantage after fabrication costs are included?

- **2.11 Other engineering and technologies — dedicated audit.**

- Current route: aerospace maritime autonomy.
- Missing depth: Safety-critical autonomy is represented; the residual engineering class is open-ended and cannot be considered exhausted.
- Next discriminating question: Which engineering professions use mature operating constraints absent from the current null-model library?

3. Medical and health sciences

4 dedicated · 1 adjacent · 0 unreviewed

- **3.1 Basic medicine — dedicated audit.**

- Current route: immune tolerance trained immunity, neurodevelopment global control, endocrine circadian control.
- Missing depth: Several regulatory systems are represented; anatomy, general physiology, genetics, metabolism, and disease mechanisms remain uneven.

- Next discriminating question: Which multi-organ control and maintenance mechanisms expose cross-module dependencies missing from neural analogies?

- **3.2 Clinical medicine — dedicated audit.**

- Current route: pathology rehabilitation, pharmacology toxicology, clinical intervention pathways, clinical specialties medical agricultural biotechnology, clinical biotechnology endogenous machinery.
- Missing depth: Pathology, rehabilitation, pharmacology, multimorbidity, intervention pathways, perioperative rescue, medical-imaging operators, paediatric extrapolation, psychiatric score configuration, obstetric competing events, dental lesion activity, and four mechanism-qualified rare-disease therapeutic observations now have bounded entries; cardiology, oncology, neurology, dermatology, ophthalmology, emergency/critical care, many surgical specialties, and broader longitudinal disease management remain unsampled.
- Next discriminating question: Which specialty and therapeutic-mechanism decisions survive acquisition and score shifts, developmental support limits, compatibility and delivery boundaries, competing events, treatment execution, rescue capacity, delayed follow-up, and externally validated uncertainty?

- **3.3 Health sciences — dedicated audit.**

- Current route: epidemiology and surveillance control, nursing care health services.
- Missing depth: Surveillance plus a first nursing, care-continuity, and health-services audit are represented; occupational health, nutrition, implementation science, long-term care, community care, and most clinical service designs remain unsampled.
- Next discriminating question: Which mechanisms preserve necessary service and legitimate person-level authority across community, long-term, acute, and digital care without hiding unfinished work or unpaid human burden?

- **3.4 Medical biotechnology — dedicated audit.**

- Current route: cellular quality control, clinical specialties medical agricultural biotechnology, clinical biotechnology endogenous machinery.
- Missing depth: Gene/cell-therapy identity, potency, manufacturing/comparability, genome-edit assay support, structural and off-target outcomes, authorization, long-term follow-up, induced proximity, lysosomal cross-correction, splice redirection, and variant-qualified pharmacological chaperoning now have bounded direct audits; tissue engineering, regenerative devices, molecular diagnostics, vaccines, many delivery platforms, biobanking, production-scale breadth, and most diseases remain incomplete.
- Next discriminating question: Which engineered-cell, gene, and molecular-product controls or recruited-machinery mechanisms remain potent, comparable, containable, measurable, compatible, delivery-qualified, and support-qualified after process, site, assay, recipient, disease-stage, and follow-up shifts?

- **3.5 Other medical science — adjacent evidence only.**

- Current route: sports expertise team coordination.
- Missing depth: Exercise and performance science provide adjacent material, but the residual medical category has not been searched systematically.
- Next discriminating question: Which clinical or care disciplines fall outside the named classes and contribute distinct coordination or evidence mechanisms?

4. Agricultural and veterinary sciences

5 dedicated · 0 adjacent · 0 unreviewed

- **4.1 Agriculture, forestry, and fisheries — dedicated audit.**

- Current route: plant distributed control, soil crop multiresource colimitation, forestry fisheries aquatic food systems.

- Missing depth: Soil/crop co-limitation, forestry, capture fisheries, aquaculture, and aquatic resource management have direct audits; silvicultural breadth, inland fisheries, pest management, livestock integration, farm economics, and many production systems remain unsampled.
 - Next discriminating question: Which policies remain effective across alternative population, spatial, climate, observation, implementation, rights, and market models without exporting harm or changing the denominator?
- **4.2 Animal and dairy science — dedicated audit.**
 - Current route: animal veterinary population health.
 - Missing depth: Resource partitioning, cattle feed restriction, welfare measurement, breeding diversity, rumen consortia, and population-event accounting are audited; species breadth, husbandry systems, reproduction, farm economics, and fisheries remain unsampled.
 - Next discriminating question: Which findings survive across species, production systems, lifetimes, protected welfare components, reproductive constraints, and start-cohort denominators?
- **4.3 Veterinary science — dedicated audit.**
 - Current route: animal veterinary population health.
 - Missing depth: Risk-based and One-Health surveillance, infection interference, evolving opponents, welfare, and dynamical identifiability are audited; comparative clinical care, surgery, companion and wildlife medicine, pharmacovigilance, and antimicrobial stewardship remain unsampled.
 - Next discriminating question: How do diagnosis, treatment, movement, surveillance, welfare, transmission, and resistance interact across species and changing target populations?
- **4.4 Agricultural biotechnology — dedicated audit.**
 - Current route: animal veterinary population health, clinical specialties medical agricultural biotechnology.
 - Missing depth: Genomic selection, microbial intervention, vaccination, and evolutionary response plus direct gene-edit/biological-control evidence on field heterogeneity, non-targets, dispersal, resistance, and EU/German authorization are represented; reproductive and animal biotechnology, diagnostics, trait development, biofertilizer/biopesticide breadth, containment, and multi-season product platforms remain incomplete.
 - Next discriminating question: Which agricultural-biotechnology effects remain after assay support, lineage concentration, off-target change, non-target exposure, dispersal, evolutionary escape, containment, multi-season field transport, and authorization are tested prospectively?
- **4.5 Other agricultural sciences — dedicated audit.**
 - Current route: forestry fisheries aquatic food systems.
 - Missing depth: Aquatic food systems and cross-resource governance provide a direct residual-field audit; rural studies, extension science, food logistics, agricultural education, and other residual specialties remain open.
 - Next discriminating question: Which residual agricultural disciplines change observation, authority, adoption, service, distribution, and long-horizon resilience beyond production output?

5. Social sciences

8 dedicated · 1 adjacent · 0 unreviewed

- **5.1 Psychology and cognitive sciences — dedicated audit.**

- Current route: comparative cognition tool use, endogenous generation creativity, learning science skill acquisition, institutional information allocation frontier.
 - Missing depth: Cognition, learning, and generation are represented, and one cross-field audit adds bounded hidden-profile, transactive-memory, responsibility, and tournament evidence; social, organizational, clinical, developmental, and differential psychology remain uneven.
 - Next discriminating question: Which cognitive and organizational effects survive representative tasks, prospective testing, unique-information controls, ownership-history ablations, and selection-bias controls?
- **5.2 Economics and business — dedicated audit.**
 - Current route: economics market design incentives, supply chain operations research, accounting audit actuarial insurance, economy education institutions, finance management risk attention governance, institutional information allocation frontier.
 - Missing depth: Markets, operations, accounting/audit/actuarial boundaries, tax incidence and public accounts, organizational bundles and persistence, labour flows and tasks, development exposure, dynamic fiscal identification, and bounded finance/management depth in tail-risk aggregation, endogenous liquidity feedback, estimated portfolios, organizational attention, irreversible commitment, independently rooted control, operational resilience, queues, internal accounting, incentive gaming, unique-evidence preservation, owner-conditioned continuation, peer-damage channels, and split audit effects are represented; banking breadth, corporate finance, marketing, strategy/entrepreneurship, industrial organization, trade, monetary, environmental and health economics, and economic history remain shallow.
 - Next discriminating question: Which economic, financial, and organizational findings survive alternative incidence, accounting, dependence, estimation, information-location, ownership-history, peer-effect, audit-exposure, endogenous-action, attention/capacity, incentive, implementation, regime, horizon, support, distribution, and independently challengeable evidence assumptions at equal total cost?
- **5.3 Education — dedicated audit.**
 - Current route: learning science skill acquisition, economy education institutions.
 - Missing depth: Learning and skill acquisition plus education-system layers, curriculum, assessment interpretation/comparability/use, inclusion, vocational education, qualifications, and comparative transport are represented; early-childhood, higher, adult and teacher education, subject didactics, educational technology, language education, school leadership/finance, and most special-education practice remain incomplete.
 - Next discriminating question: Which learner and system interventions survive curriculum-layer, opportunity, access, assessment-use, delayed-transfer, institutional-support, qualification, and source-to-target comparability tests?
- **5.4 Sociology — dedicated audit.**
 - Current route: direct social research ethnography media, cultural evolution archaeology, high reliability organizations incident learning, social science depth institutions inequality digital media, political administrative stratification media depth.
 - Missing depth: Direct social-research methods plus bounded stratification, household allocation, paid/unpaid work, migration measurement, collective-action, criminology, digital-field, wealth-tail observation, stage-bounded discrimination, and cumulative-inequality entries are represented; education, urban/rural and health sociology, science and technology studies, social work, historical sociology, population breadth, and most substantive theory remain incomplete.
 - Next discriminating question: Which substantive sociological results survive explicit construct, inequality operator, household allocation, tail-frame/nonresponse, experimental-stage, history/attrition, migration denominator, participation cost, crime-observation, platform, positionality, and shared-error-root tests?
- **5.5 Law — dedicated audit.**
 - Current route: legal evidence procedure.

- Missing depth: Evidence and procedure are represented; substantive EU/German fields, remedies, administration, competition, labour, and criminal law remain partial.
 - Next discriminating question: Which contestability and burden-of-proof mechanisms should govern automated decisions and scientific promotion?
- **5.6 Political science — dedicated audit.**
 - Current route: social choice institutions, social science depth institutions inequality digital media, political administrative stratification media depth, institutional information allocation frontier.
 - Missing depth: Formal governance plus bounded policy-implementation, administrative-burden, participation, collective-action, repression, vote-to-seat provenance, roll-call selection, performance-target, task-indexed administrative-capacity, responsibility-conditioned continuation, and measurement-versus-response audit mechanisms are represented; party organization beyond roll calls, executives, courts, comparative regimes, international relations, conflict, public finance, and policy breadth remain incomplete.
 - Next discriminating question: Which political and administrative claims survive rule-version, selected-vote, latent-model, target-gaming/no-gaming, actor-level-task capacity, decision-lineage, audit-estimand and exposure, authority, burden, participation, legitimacy, and implementation tests?
- **5.7 Social and economic geography — dedicated audit.**
 - Current route: built environment urban systems, quantitative history demography, spatial geography support scale mobility place.
 - Missing depth: Spatial support, scale and zoning sensitivity, ecological inference, prediction geometry, mobility/activity context, conservative change of support, versioned statistical geographies, preferential observation, and multiscale segregation now have a bounded field-centered audit; economic and political geography, rural/urban systems, housing, regional development, migration, geopolitics, substantive place theory, and most geographic breadth remain incomplete.
 - Next discriminating question: Which spatial conclusions survive alternative support, partition, scale, unit, prediction domain, activity-space, boundary version, observation process, confidentiality constraint, and equal-cost mature spatial-statistical nulls?
- **5.8 Media and communications — dedicated audit.**
 - Current route: linguistics communication, social science depth institutions inequality digital media, political administrative stratification media depth, institutional information allocation frontier.
 - Missing depth: Communication mechanisms plus bounded platform-conditioned methods, comparative media-system operators, cross-platform audience identity/deduplication, moderation-record denominator/appeal boundaries, hidden-profile information loss, and expertise-location evidence are represented; journalism practice and history, media industries, interpersonal, intercultural, most organizational, health and environmental communication, and broader audience research remain incomplete.
 - Next discriminating question: Which communication claims survive channel, audience-unit, unique-evidence coverage, expertise-directory staleness, identity-linkage, exposure-qualification, provider-interface, moderation-denominator, appeal-selection, institution, ownership, financing, and repeated-exposure tests?
- **5.9 Other social sciences — adjacent evidence only.**
 - Current route: hci human factors, high reliability organizations incident learning.
 - Missing depth: Human factors and organizational work provide adjacent material, but the residual social-science category has not been searched systematically.
 - Next discriminating question: Which professions and hybrid social sciences contribute distinct observation, coordination, or accountability mechanisms?

6. Humanities and the arts

5 dedicated · 0 adjacent · 0 unreviewed

- **6.1 History and archaeology — dedicated audit.**
 - Current route: cultural evolution archaeology, quantitative history demography.
 - Missing depth: Material inference and quantitative history are represented; archival criticism, period-specific methods, oral history, and history of science remain partial.
 - Next discriminating question: How should path dependence, missing archives, provenance, survivorship, and changing categories constrain learned historical memory?
- **6.2 Languages and literature — dedicated audit.**
 - Current route: linguistics communication, textual criticism variant traditions.
 - Missing depth: Linguistics, textual criticism, variant traditions, scholarly editing, and translation qualification are represented; narrative theory, poetics, comparative literature, reception, performance, and most language traditions remain unsampled.
 - Next discriminating question: Which narrative and interpretive structures change counterfactual evaluation, perspective, compression, long-range coherence, or disagreement handling beyond the new witness-variant contract?
- **6.3 Philosophy, ethics and religion — dedicated audit.**
 - Current route: philosophy of science theory choice, theology religious practice ritual.
 - Missing depth: Philosophy of science, theory choice, theology, religious practice, and ritual now have field-centered audits; ethics, philosophy of mind and technology, comparative theology, and most religious traditions remain unsampled.
 - Next discriminating question: Which accounts of agency, moral status, value conflict, testimony, interpretation, conscience, and authority change system governance or evaluation contracts?
- **6.4 Arts (arts, history of arts, performing arts, music) — dedicated audit.**
 - Current route: visual art design cognition, music cognition improvisation.
 - Missing depth: Visual art and music are represented; theatre, dance, film, architecture as art, conservation, and art history remain uneven.
 - Next discriminating question: Which embodied, temporal, collaborative, and material practices produce useful variants and evaluations unavailable to static generation?
- **6.5 Other humanities — dedicated audit.**
 - Current route: residual humanities living heritage practice.
 - Missing depth: A deliberate residual-category audit now samples living heritage, rehearsal, conservation, archival absence, and digital preservation; the catch-all remains open and most direct humanities practices are still unsampled.
 - Next discriminating question: Which methods from oral history, museum and curatorial practice, film, theatre and dance traditions, folklore, interpreting, conservation science, and heritage conflict change a test rather than duplicate provenance, governance, or protocol versioning?

DFG granularity probe

The DFG layer catches gaps hidden by the broader OECD cells. It is shown here at review-board level; all 214 individual subjects are enumerated separately in the fine-grained field-depth inventory, but still require child-level routing assignments.

1. Geistes- und Sozialwissenschaften

11 dedicated · 2 adjacent · 0 unreviewed

- **1.11 Alte Kulturen — adjacent evidence only.**
 - Current route: cultural evolution archaeology.
 - Missing depth: Archaeological inference is present, but ancient-culture detail remains sparse.
- **1.12 Geschichtswissenschaften — dedicated audit.**
 - Current route: quantitative history demography.
 - Missing depth: Historical method breadth remains partial.
- **1.13 Kunst-, Musik-, Theater- und Medienwissenschaften — dedicated audit.**
 - Current route: visual art design cognition, music cognition improvisation.
 - Missing depth: Theatre and media studies are not directly audited.
- **1.14 Sprachwissenschaften — dedicated audit.**
 - Current route: linguistics communication.
 - Missing depth: Historical and typological breadth remains partial.
- **1.15 Literaturwissenschaft — dedicated audit.**
 - Current route: textual criticism variant traditions.
 - Missing depth: Textual criticism, scholarly editing, variant traditions, and translation qualification are audited; literary theory and most genres and traditions remain unsampled.
- **1.16 Sozial- und Kulturanthropologie, außereuropäische Kulturen, Judaistik und Religionswissenschaft — adjacent evidence only.**
 - Current route: direct social research ethnography media, theology religious practice ritual.
 - Missing depth: Religious practice and ritual have a field-centered audit, but ethnography, anthropology, Judaic studies, non-European cultures, and comparative religious studies remain unsampled.
- **1.17 Theologie — dedicated audit.**
 - Current route: theology religious practice ritual.
 - Missing depth: Practice, ritual, interpretation, canon, authority, conscience, and evidence boundaries are audited; doctrinal, historical, pastoral, comparative, and tradition-specific theology remain unsampled.
- **1.18 Philosophie — dedicated audit.**
 - Current route: philosophy of science theory choice.
 - Missing depth: Philosophy of science and theory choice are audited; most other philosophical subfields remain unsampled.
- **1.21 Erziehungswissenschaft und Bildungsforschung — dedicated audit.**
 - Current route: learning science skill acquisition, economy education institutions.
 - Missing depth: Learning science plus bounded system-layer, curriculum, assessment-validity/use, inclusion, vocational-education, qualification, and comparative-transport evidence are represented; early-childhood, higher, adult and teacher education, subject didactics, educational technology, language education, leadership/finance, and most special-education practice remain incomplete.

- **1.22 Psychologie — dedicated audit.**
 - Current route: comparative cognition tool use, endogenous generation creativity, learning science skill acquisition.
 - Missing depth: Clinical, organizational, and personality work remain uneven.
- **1.23 Sozialwissenschaften — dedicated audit.**
 - Current route: direct social research ethnography media, social choice institutions, social science depth institutions inequality digital media, economy education institutions, political administrative stratification media depth.
 - Missing depth: Empirical social research plus bounded policy implementation, inequality, household/work/care allocation, labour-state and task flows, migration, qualification and comparative-institution transport, collective action, criminology, vote-to-seat provenance, roll-call selection, target/capacity measurement, wealth tails, stage-bounded discrimination, cumulative inequality, audience deduplication, platform methods, moderation records, and media-system operators are represented; party organization beyond roll calls, executives, courts, comparative politics and international relations, journalism practice/history, media industries, urban/rural and health sociology, social work, and most substantive breadth remain incomplete.
- **1.24 Wirtschaftswissenschaften — dedicated audit.**
 - Current route: economics market design incentives, supply chain operations research, accounting audit actuarial insurance, economy education institutions, finance management risk attention governance.
 - Missing depth: Market design, operations, accounting/audit/actuarial/insurance boundaries, tax incidence and public accounts, organizational bundles and persistence, labour flows/tasks, development exposure, dynamic fiscal identification, and bounded finance/management depth in tail risk, endogenous liquidity, estimated portfolios, attention, options, internal control, resilience, queues, transfer prices, and incentives are represented; banking and corporate-finance breadth, marketing, strategy/entrepreneurship, industrial organization, trade, monetary, environmental and health economics, and economic history remain incomplete.
- **1.25 Rechtswissenschaften — dedicated audit.**
 - Current route: legal evidence procedure.
 - Missing depth: Evidence and procedure dominate; substantive fields remain partial.

2. Lebenswissenschaften

7 dedicated · 0 adjacent · 0 unreviewed

- **2.11 Grundlagen der Biologie und Medizin — dedicated audit.**
 - Current route: cellular quality control, developmental morphogenesis, clinical specialties medical agricultural biotechnology, developmental regeneration depth.
 - Missing depth: Cellular/developmental control, positional memory, regenerative instruction, morphogen-field scaling, boundary mechanics, plus bounded gene/cell-product and genome-edit assay/structural-outcome evidence are represented; genomics, structural biology, molecular medicine, general physiology, bioinformatics, and tissue-biology breadth remain uneven.
- **2.12 Pflanzenwissenschaften — dedicated audit.**
 - Current route: plant distributed control, soil crop multiresource colimitation, plant plasticity memory signalling.
 - Missing depth: Distributed control, multi-resource crop context, memory/reset, systemic signalling, sense-by-growth development, and environmental integration are represented; field transport, phylogenetic breadth, whole-life resource accounting, and crop validation remain partial.
- **2.13 Zoologie — dedicated audit.**

- Current route: animal navigation sensory ecology, comparative cognition tool use, integrative comparative physiology.
 - Missing depth: Behavior, navigation, comparative cognition, and integrative comparative physiology now have field-centred audits; morphology, systematics, molecular physiology, ecological physiology, evo-devo, and most taxonomic breadth remain incomplete.
- **2.21 Mikrobiologie, Virologie und Immunologie — dedicated audit.**
 - Current route: microbial ecology biofilms, immune tolerance trained immunity, parasitology adversarial lifecycles.
 - Missing depth: Virology remains a major gap; parasitology now has a bounded first pass, while tropical-pathogen, host, vector, immune, clinical and taxonomic breadth remains incomplete.
- **2.22 Medizin — dedicated audit.**
 - Current route: pathology rehabilitation, pharmacology toxicology, epidemiology and surveillance control, nursing care health services, clinical intervention pathways, clinical specialties medical agricultural biotechnology, medical devices biomedical engineering.
 - Missing depth: Pathology, rehabilitation, pharmacology, intervention pathways, perioperative rescue, imaging, paediatric extrapolation, psychiatric measurement, obstetric competing events, and dental lesion activity have bounded entries; a direct medical-engineering audit now covers device chains, signal and image operators, prosthetic control, critical-task usability, drift, alarms, interoperability, cybersecurity, clinical performance, and post-market surveillance within the DFG Medizin subjects for medical informatics and biomedical engineering. Cardiology, oncology, neurology, dermatology, ophthalmology, emergency/critical care, many surgical specialties, and device-specific biomaterials, implant longevity, biocompatibility, electrical safety, sterilization, radiation physics, manufacturing, servicing, and clinical-investigation breadth remain unsampled or incomplete.
- **2.23 Neurowissenschaften — dedicated audit.**
 - Current route: neurodevelopment global control, memory replay forgetting.
 - Missing depth: Strongest life-science concentration; still not exhaustive.
- **2.31 Agrar-, Forstwissenschaften und Tiermedizin — dedicated audit.**
 - Current route: soil crop multiresource colimitation, animal veterinary population health, forestry fisheries aquatic food systems, clinical specialties medical agricultural biotechnology.
 - Missing depth: Soil/crop co-limitation, animal production/welfare/breeding, veterinary population health, forestry/fisheries/aquaculture, and bounded agricultural-biotechnology/biological-control field transport and resistance are represented; reproductive biotechnology, diagnostics, many species and production systems, food science, farm economics, and clinical veterinary breadth remain unsampled.

3. Naturwissenschaften

17 dedicated · 2 adjacent · 0 unreviewed

- **3.11 Molekülchemie — dedicated audit.**
 - Current route: molecular chemistry synthesis systems.
 - Missing depth: Dynamic covalent assembly, self-sorting, context-qualified recognition, photochemical/electrochemical pathway control, structure elucidation and automated discovery nulls now have a bounded audit; inorganic/organometallic breadth, total synthesis, radical/pericyclic/C–H activation, flow/scale-up, impurity fate and reproducibility remain incomplete.
- **3.12 Chemische Festkörper- und Oberflächenforschung — adjacent evidence only.**

- Current route: adaptive materials and self assembly, semiconductor device reliability, electrochemistry interface memory degradation.
 - Missing depth: Electrochemical phase, passivation, insertion hysteresis, and degradation provide adjacent interface/material depth; general surface chemistry, characterization, catalysis, adsorption, corrosion, and heterogeneous interfaces remain incomplete.
- **3.13 Physikalische Chemie — dedicated audit.**
 - Current route: chemistry reaction networks proofreading, information thermodynamics physical computation, electrochemistry interface memory degradation.
 - Missing depth: A direct electrochemistry audit now covers interface kinetics, finite diffusion, impedance validity and inversion, rate-dependent phase state, passivation, depletion, identifiability, hysteresis, and delayed degradation; spectroscopy, molecular dynamics, photochemistry, statistical mechanics, colloid/interface breadth, and real-apparatus qualification remain incomplete.
- **3.14 Analytische Chemie — dedicated audit.**
 - Current route: metrology measurement science, olfaction chemical sensing plume tracking, biotechnology chemistry process systems, mineralogy petrology geochemistry, measurement heavy analytical water earth observation.
 - Missing depth: A direct audit now covers measurand/speciation, preparation, matrix effects, calibration hierarchy, detection/quantification, conformity, sampling, interlaboratory evidence, and shared bias; electroanalysis, separations, spectroscopy/imaging, surface analysis, high-throughput automation, reference-material breadth, and many complex matrices remain incomplete.
- **3.15 Biologische Chemie und Lebensmittelchemie — dedicated audit.**
 - Current route: biotechnology chemistry process systems, animal veterinary population health, forestry fisheries aquatic food systems, measurement heavy analytical water earth observation.
 - Missing depth: Food-control sampling, matrix-qualified analysis, species/fraction boundaries, uncertainty, and conformity now have a bounded direct audit; molecular biochemistry, food composition and reactions, processing, preservation, toxicology, sensory chemistry, authenticity, and nutritional breadth remain incomplete.
- **3.16 Polymerforschung — dedicated audit.**
 - Current route: polymer research adaptive materials.
 - Missing depth: Polymerization mechanisms and populations, entanglement, viscoelastic spectra, conditional time–temperature superposition, phase/self-assembly paths, gel criteria, dynamic covalent networks, healing, response, ageing/fatigue, sequence storage, circularity and measurement operators now have a bounded audit; processing/manufacturing, biopolymers, membranes, conducting polymers, adhesion, composite interfaces, toxicology, fire and application qualification remain incomplete.
- **3.17 Theoretische Chemie — adjacent evidence only.**
 - Current route: chemistry reaction networks proofreading, biotechnology chemistry process systems.
 - Missing depth: Stoichiometric networks, microkinetics, scaling relations, and transport models provide adjacent formalisms; electronic structure, quantum chemistry, molecular dynamics, statistical mechanics, and method validation lack a field-centered audit.
- **3.21 Physik der kondensierten Materie — dedicated audit.**
 - Current route: soft active matter, semiconductor device reliability, atomic molecular plasma condensed matter.

- Missing depth: Soft matter, device reliability, and bounded direct depth in finite-size phases, quenched disorder, topology, mobility gaps, and contacted transport inference are represented. Electronic and quantum materials, synthesis and characterization, superconductivity, magnetism, semiconductor breadth, and nonequilibrium condensed-matter dynamics remain incomplete.
- **3.22 Statistische Physik, nichtlineare Dynamik, komplexe Systeme, weiche und fluide Materie, biologische Physik — dedicated audit.**
 - Current route: soft active matter, fluid dynamics turbulence.
 - Missing depth: Several components are represented; nonlinear dynamics remains broad.
- **3.23 Optik, Quantenoptik und Physik der Atome, Moleküle und Plasmen — dedicated audit.**
 - Current route: optics photonics inverse sensing, atomic molecular plasma condensed matter.
 - Missing depth: Optics plus bounded direct depth in atomic-frequency uncertainty, molecular transition networks, ultracold observable-family measurement, non-equilibrium plasma diagnostics, and reconstruction-qualified plasma control are represented. AMO theory, collision breadth, ultrafast dynamics, quantum-information breadth, and plasma facility, kinetics, waves, turbulence, and materials practice remain partial.
- **3.24 Teilchen, Kerne und Felder — dedicated audit.**
 - Current route: particle nuclear high energy experimentation, atomic molecular plasma condensed matter.
 - Missing depth: Triggering, detector response, calibration, unfolding, nuisance uncertainty, rare-event search, blinding, simulation, preservation, evaluated nuclear data, symmetry, effective-model breakdown, and bounded accelerator beam phase-space instrumentation are represented. Accelerator sources, RF and magnets, collective effects, detector technology, nuclear structure/reactions, neutrino, astroparticle, phenomenology, and theory breadth remain unsampled or incomplete.
- **3.25 Astrophysik und Astronomie — dedicated audit.**
 - Current route: astronomy remote inference.
 - Missing depth: Remote inference is represented; field breadth remains partial.
- **3.31 Mathematik — dedicated audit.**
 - Current route: mathematical practice proof discovery, applied multiscale reduction.
 - Missing depth: Practice, proof, and bounded multiscale reduction are represented; most pure, statistical, optimization, inverse-problem, stochastic, and computational subfield breadth is not.
- **3.41 Atmosphären-, Meeres- und Klimaforschung — dedicated audit.**
 - Current route: earth system transition signals, forestry fisheries aquatic food systems, mineralogy petrology geochemistry, measurement heavy analytical water earth observation.
 - Missing depth: Earth-system transitions plus direct observation/reanalysis depth for ocean and atmosphere now cover heterogeneous operators, sampling, quality control, priors, dynamics, covariance, selection, and product versions; ocean-process physics, marine biogeochemistry, atmospheric chemistry/clouds, weather prediction, paleoclimate, and climate-dynamics breadth remain shallow.
- **3.42 Geologie und Paläontologie — dedicated audit.**
 - Current route: geology geomorphology, paleobiology major transitions.
 - Missing depth: Both have initial audits.

- **3.43 Geophysik und Geodäsie — dedicated audit.**
 - Current route: geology geomorphology, mineralogy petrology geochemistry, measurement heavy analytical water earth observation.
 - Missing depth: A direct measurement audit now covers reference frame/epoch, coordinate and velocity covariance, gravity/seismic forward kernels, inverse resolution, deformation fields, and Earth-observation support; seismology, geomagnetism, geodynamics, exploration geophysics, physical geodesy, GNSS/InSAR processing breadth, and field instrumentation remain incomplete.
- **3.44 Mineralogie, Petrologie und Geochemie — dedicated audit.**
 - Current route: mineralogy petrology geochemistry.
 - Missing depth: Phase equilibria, kinetic trapping and replacement, P–T–t records, isotope and reservoir inverses, weathering/redox, reactive upscaling, detrital provenance, geochronology, preservation and metrology now have a bounded field audit; experimental mineral physics/crystallography, igneous and mantle processes, ore/economic geology, cosmochemistry, and analytical breadth remain incomplete.
- **3.45 Geographie — dedicated audit.**
 - Current route: built environment urban systems, geology geomorphology, spatial geography support scale mobility place.
 - Missing depth: A field-centered audit now covers spatial support, aggregation/scale, ecological inference, prediction geometry, mobility/activity context, conservative remapping, classification versions, preferential observation, multiscale separation and confidentiality boundaries. Physical-geography process breadth, economic/political/rural/urban geography, housing, regional development, migration, geopolitics, and most substantive human-geography traditions remain incomplete.
- **3.46 Wasserforschung — dedicated audit.**
 - Current route: forestry fisheries aquatic food systems, mineralogy petrology geochemistry, biotechnology chemistry process systems, measurement heavy analytical water earth observation, environmental engineering water waste remediation.
 - Missing depth: Direct audits now cover concentration/load/storage and treatment-performance separation, catchment and point-support measurement, groundwater and lake observation, residence-time distributions, HRT/SRT separation, integrated wastewater control, fouling, adsorption breakthrough, anaerobic delayed state, transformation products/effects, reuse lifecycle boundaries, and EU/German normative context. Hydraulics, ecohydrology, flood/drought process breadth, drinking-water engineering, urban drainage design, groundwater remediation, receiving-water ecology, operator practice, and governance remain incomplete.

4. Ingenieurwissenschaften

10 dedicated · 0 adjacent · 0 unreviewed

- **4.11 Produktionstechnik — dedicated audit.**
 - Current route: supply chain operations research, process engineering, production maintenance nanomanufacturing communications material qualification.
 - Missing depth: Statistical process state, qualified process windows, maintenance policy/support, manufacturing genealogy and bounded nanomanufacturing evidence have a direct audit; factory planning/layout, production logistics, machining, forming, joining, additive manufacture, human work and most process-specific production technologies remain incomplete.
- **4.12 Mechanik und konstruktiver Maschinenbau — dedicated audit.**
 - Current route: mechanical civil resilience, biomechanics motor control, tribology contact adaptive interfaces.

- Missing depth: Mechanics, biomechanical control, and bounded tribology/contact-interface depth are represented; broad machine-element design, product development, thermal machinery, real materials testing, and industrial qualification remain incomplete.
- **4.21 Verfahrenstechnik und Technische Chemie — dedicated audit.**
 - Current route: process engineering, biotechnology chemistry process systems, environmental engineering water waste remediation.
 - Missing depth: Process and bioprocess engineering are joined by bounded treatment-process depth in RTD, population retention, integrated control, process/sensor anomaly isolation, membrane fouling, competitive column breakthrough, anaerobic state estimation, and transformation-product closure. Product-specific manufacturing and downstream breadth, multiphase and thermal processes, scale-up/commissioning, contamination control, process safety, validation, and industrial authorization remain partial.
- **4.22 Strömungsmechanik, Technische Thermodynamik und Thermische Energietechnik — dedicated audit.**
 - Current route: fluid dynamics turbulence.
 - Missing depth: Energy machinery and applied thermal systems remain partial.
- **4.31 Werkstofftechnik — dedicated audit.**
 - Current route: adaptive materials and self assembly, polymer research adaptive materials, production maintenance nanomanufacturing communications material qualification, tribology contact adaptive interfaces.
 - Missing depth: Adaptive-material/polymer mechanisms, bounded qualification and lineage, and adjacent surface-texture/coating/interface evidence are represented; metallurgy, ceramics, composites, joining, surface chemistry, degradation, manufacture, and most production variation remain incomplete.
- **4.32 Materialwissenschaft — dedicated audit.**
 - Current route: adaptive materials and self assembly, polymer research adaptive materials, semiconductor device reliability.
 - Missing depth: Adaptive-material, polymer and device-reliability mechanisms are represented; biomaterials, electronic/quantum materials, composite interfaces and multiscale design remain partial.
- **4.41 Systemtechnik — dedicated audit.**
 - Current route: engineering analogues, aerospace maritime autonomy, medical devices biomedical engineering.
 - Missing depth: Biomedical system chains, assistive control, critical-task human-machine use, alarms, calibration drift, semantic interoperability, and safety-coupled cybersecurity now have bounded direct coverage aligned with biomedical systems and ergonomics within this board; transport systems, general robotics and autonomy, microsystems, sensor and actuator hardware, real-time control verification, networked control, and hardware-in-the-loop validation remain uneven.
- **4.42 Elektrotechnik und Informationstechnik — dedicated audit.**
 - Current route: power grids protection and recovery, semiconductor device reliability, production maintenance nanomanufacturing communications material qualification, computing compilers networking visualization.
 - Missing depth: Power, device reliability, adaptive link control, interacting network loops, and bounded end-to-end IP measurement, probe, congestion, AQM, fairness, and accepted-service boundaries are represented; antenna/RF and optical hardware, channel/coding breadth, electromagnetic compatibility, signal processing, sensors, control hardware, and communications-system deployment remain incomplete.

- **4.43 Informatik — dedicated audit.**
 - Current route: databases storage, programming languages verification, security cryptography, HCI human factors, computing compilers networking visualization, operating systems distributed consistency recovery measurement.
 - Missing depth: Storage, programming languages/verification, security, HCI, compiler correctness/testing, PGO/JIT, randomized performance measurement, endpoint placement, IP networking/tails, graphical perception, vector software quality/lifecycle, and bounded operating/distributed-system depth in fencing, liveness assumptions, retry effects, recovery boundaries, open-arrival tails and Linux resource scope are represented; scheduler/virtual-memory design, kernels/drivers, real-time systems, filesystems and distributed-systems breadth, empirical software engineering, accessible visualization, and applied information systems remain incomplete.
- **4.51 Bauwesen und Architektur — dedicated audit.**
 - Current route: built environment urban systems, mechanical civil resilience, environmental engineering water waste remediation.
 - Missing depth: Built systems and structural resilience are joined by bounded water-infrastructure depth in coupled sewer-plant control and context-qualified centralized/decentralized topology under conveyance, maintenance, operator, reliability, lifecycle, and common-cause constraints. Construction production, architecture practice, procurement, urban drainage and hydraulic design, commissioning, buried-asset deterioration, and broad water-infrastructure operation remain partial.

ANZSRC independent division census

This third lens records all 23 official divisions. Its much finer 213 groups and 1,967 fields remain an explicit resolution debt; a green division means one entry audit, not comprehensive coverage.

- **30 Agricultural, veterinary and food sciences — dedicated audit.**
 - Current route: soil crop multiresource colimitation, animal veterinary population health, forestry fisheries aquatic food systems, measurement heavy analytical water earth observation, clinical specialties medical agricultural biotechnology, parasitology adversarial lifecycles.
 - Missing depth: Soil/crop, livestock, veterinary population health, forestry/fisheries/aquaculture, food-control analytical measurement, bounded agricultural biotechnology, and one rodent-parasite competitive-release boundary have direct entries; broader veterinary parasitology, food processing/preservation/nutrition, farm systems/economics, horticulture, reproductive biotechnology, diagnostics, and many species remain unsampled.
- **31 Biological sciences — dedicated audit.**
 - Current route: cellular quality control, collective ecological resilience, developmental morphogenesis, microbial ecology biofilms, plant distributed control, integrative comparative physiology, developmental regeneration depth, plant plasticity memory signalling, parasitology adversarial lifecycles.
 - Missing depth: Cellular control, development, regeneration, ecology, microbes, plant plasticity, integrative comparative physiology, and a bounded parasitology pass are represented; taxonomy, anatomy, marine biology, genomics, ecological physiology, most parasite biology, and most organismal breadth remain incomplete.
- **32 Biomedical and clinical sciences — dedicated audit.**
 - Current route: pathology rehabilitation, pharmacology toxicology, epidemiology and surveillance control, nursing care health services, clinical intervention pathways, clinical specialties medical agricultural biotechnology, parasitology adversarial lifecycles.

- Missing depth: Pathology, rehabilitation, pharmacology/toxicology, epidemiology, nursing/care, intervention pathways, selected specialties and a non-clinical parasitology mechanism pass have bounded entries; cardiology, oncology, neurology, dermatology, ophthalmology, emergency/critical care, many surgical specialties, clinical parasitology, tissue engineering, and broader longitudinal disease management remain unsampled.
- **33 Built environment and design — dedicated audit.**
 - Current route: built environment urban systems, mechanical civil resilience, visual art design cognition.
 - Missing depth: Urban systems, structural resilience and design cognition are represented; architecture practice, construction production, landscape architecture, planning law and building operation remain partial.
- **34 Chemical sciences — dedicated audit.**
 - Current route: chemistry reaction networks proofreading, biotechnology chemistry process systems, molecular chemistry synthesis systems, polymer research adaptive materials, measurement heavy analytical water earth observation, electrochemistry interface memory degradation.
 - Missing depth: Reaction networks, molecular assembly/control, catalysis, process chemistry, polymers, analytical measurement, and bounded electrochemistry depth are represented; corrosion, bioelectrochemistry, surface/theoretical chemistry, separations/spectroscopy breadth, and much inorganic, organometallic, total-synthesis, and scale-up chemistry remain uneven.
- **35 Commerce, management, tourism and services — dedicated audit.**
 - Current route: supply chain operations research, accounting audit actuarial insurance, economy education institutions, finance management risk attention governance.
 - Missing depth: Operations, accounting/audit/actuarial/insurance, organizational adoption and persistence, and bounded management depth in attention routing, irreversible commitment, independently rooted control, operational resilience, deliverable capacity, internal accounting, and incentive gaming are represented; strategy, entrepreneurship, marketing, corporate-finance breadth, tourism, hospitality, retail, human resources, and most service professions remain unsampled.
- **36 Creative arts and writing — dedicated audit.**
 - Current route: visual art design cognition, music cognition improvisation, residual humanities living heritage practice.
 - Missing depth: Visual art and music are direct entries, and rehearsal/conservation now have a narrow cross-field probe; writing practice, film, theatre and dance traditions, curating, and most performance research remain unsampled.
- **37 Earth sciences — dedicated audit.**
 - Current route: earth system transition signals, geology geomorphology, astronomy remote inference, mineralogy petrology geochemistry, measurement heavy analytical water earth observation.
 - Missing depth: Geology, geomorphology, mineralogy/petrology/geochemistry, Earth-system transitions, remote inference, and direct observation depth for hydrology, ocean/atmosphere, geophysics, geodesy, inverse resolution, and Earth-observation pixels/retrievals are represented; process breadth in seismology, geomagnetism, geodynamics, meteorology, oceanography, climate, experimental mineral physics, ore geology, and cosmochemistry remains uneven.
- **38 Economics — dedicated audit.**
 - Current route: economics market design incentives, social choice institutions, economy education institutions, finance management risk attention governance.

- Missing depth: Market design, incentives, social choice, tax incidence/public accounts, labour stocks/flows/tasks, development exposure, dynamic fiscal identification, and bounded finance depth in tail risk, endogenous liquidity, estimated portfolios, real options, transfer prices, and incentive gaming are represented; monetary, trade, industrial-organization, banking breadth, corporate finance, environmental/health, household, and economic-history breadth remain incomplete.
- **39 Education — dedicated audit.**
 - Current route: learning science skill acquisition, economy education institutions.
 - Missing depth: Learning/skill acquisition plus education-system layers, curriculum, assessment interpretation/comparability/use, inclusion, vocational education, qualifications, and comparative transport are direct entries; early-childhood, higher, adult and teacher education, subject didactics, educational technology, language education, leadership/finance, and most special-education practice remain incomplete.
- **40 Engineering — dedicated audit.**
 - Current route: engineering analogues, mechanical civil resilience, process engineering, power grids protection and recovery, polymer research adaptive materials, production maintenance nanomanufacturing communications material qualification, computing compilers networking visualization, medical devices biomedical engineering, environmental engineering water waste remediation, tribology contact adaptive interfaces, electrochemistry interface memory degradation.
 - Missing depth: Several engineering families now include bounded tribology and adjacent electrochemical-storage depth alongside production, process, maintenance, communications, computing, biomedical, and environmental systems. Mining, transport, factory/layout breadth, RF/optical and robotics hardware, channel/coding systems, thermal machinery, broad storage/conversion hardware, air and solid-waste engineering, hydraulics, construction/commissioning, and many material and biomedical systems remain uneven.
- **41 Environmental sciences — dedicated audit.**
 - Current route: earth system transition signals, collective ecological resilience, biotechnology chemistry process systems, measurement heavy analytical water earth observation, environmental engineering water waste remediation.
 - Missing depth: Earth systems, ecological resilience, remediation/environmental biotechnology, water and Earth-observation measurement, and bounded environmental-engineering depth in treatment state/control, mass/species/effect closure, reuse lifecycle counterfactuals, and infrastructure context are represented. Contamination fate and ecotoxicology breadth, air/soil/solid-waste systems, conservation management, environmental policy/economics, climate processes, and ecosystem breadth remain partial.
- **42 Health sciences — dedicated audit.**
 - Current route: epidemiology and surveillance control, pathology rehabilitation, nursing care health services.
 - Missing depth: Population surveillance, rehabilitation, nursing and health services are represented; public-health interventions, allied health, nutrition, health economics and implementation science remain uneven.
- **43 History, heritage and archaeology — dedicated audit.**
 - Current route: cultural evolution archaeology, quantitative history demography, textual criticism variant traditions, residual humanities living heritage practice.
 - Missing depth: Archaeological inference, quantitative history, textual traditions, living heritage, archival absence and conservation boundaries are represented; oral history, museum and curatorial practice, heritage conflict, and most periods and regions remain unsampled.
- **44 Human society — dedicated audit.**

- Current route: direct social research ethnography media, social choice institutions, indigenous data knowledge governance, theology religious practice ritual, social science depth institutions inequality digital media, economy education institutions, political administrative stratification media depth, spatial geography support scale mobility place.
 - Missing depth: Direct social-research methods plus bounded policy/administration, inequality, household/work/care allocation, labour-state/task flows, migration, qualification and comparative-institution transport, collective action, criminology, electoral/legislative observation, target and task-capacity measurement, wealth-tail and discrimination-stage inference, cumulative inequality, digital-field depth, and field-centered spatial support, scale, mobility/context and segregation-profile operators are represented; social work, urban/rural and health sociology, historical sociology, economic/political geography, place theory, party/executive/court depth, comparative politics/international relations, and most substantive anthropology and sociology remain incomplete.
- **45 Indigenous studies — dedicated audit.**
 - Current route: indigenous data knowledge governance.
 - Missing depth: Indigenous data and knowledge governance has an Indigenous-led-source entry audit; languages, law, land, health, education, histories and community-specific knowledge require their own authorities and cannot be inferred from that entry.
- **46 Information and computing sciences — dedicated audit.**
 - Current route: databases storage, programming languages verification, security cryptography, hci human factors, computing compilers networking visualization, operating systems distributed consistency recovery measurement.
 - Missing depth: Storage, programming languages/verification, security, HCI, compiler semantics/testing, workload-qualified PGO/JIT, randomized performance inference, endpoint placement, IP observation/congestion/tails, graphical perception, vector software-quality/lifecycle, and bounded operating/distributed-system depth in fencing, liveness assumptions, retry effects, recovery boundaries, open-arrival tails and Linux resource scope are represented; kernel and operating-system breadth, distributed transactions/streaming, language/runtime breadth, empirical software engineering, accessible human visualization studies, and many applied information systems remain incomplete.
- **47 Language, communication and culture — dedicated audit.**
 - Current route: linguistics communication, textual criticism variant traditions, direct social research ethnography media, social science depth institutions inequality digital media, political administrative stratification media depth.
 - Missing depth: Linguistics, textual criticism, translation qualification, bounded communication/media gatekeeping and effects, platform-conditioned methods, comparative media-system operators, audience identity/deduplication, and moderation-record denominator/appeal inference are represented; journalism practice/history, media industries, interpersonal/intercultural/organizational communication, interpreting, cultural studies, rhetoric and most language traditions remain incomplete.
- **48 Law and legal studies — dedicated audit.**
 - Current route: legal evidence procedure.
 - Missing depth: Evidence and procedure are direct entries; substantive private, public, criminal, labor, environmental and international law remain partial and jurisdiction-specific.
- **49 Mathematical sciences — dedicated audit.**
 - Current route: mathematical practice proof discovery, applied multiscale reduction.
 - Missing depth: Formal discovery, proof practice, and bounded applied multiscale reduction are represented; most pure, statistical, optimization, inverse-problem, stochastic, and computational subfields remain unsampled.

- **50 Philosophy and religious studies — dedicated audit.**
 - Current route: philosophy of science theory choice, theology religious practice ritual.
 - Missing depth: Theory choice, theology, religious practice and ritual are direct entries; ethics, philosophy of mind and technology, comparative religion and most traditions remain unsampled.
- **51 Physical sciences — dedicated audit.**
 - Current route: information thermodynamics physical computation, optics photonics inverse sensing, soft active matter, particle nuclear high energy experimentation, atomic molecular plasma condensed matter.
 - Missing depth: Thermodynamics, optics, soft matter, particle/nuclear experimental methods, and bounded atomic/molecular, ultracold-measurement, condensed-matter inference, plasma diagnostic/control, and accelerator beam-instrumentation depth are represented. Quantum information, nuclear structure/reactions, theoretical-field breadth, materials synthesis, and much AMO, plasma-facility, and accelerator-physics breadth remain partial.
- **52 Psychology — dedicated audit.**
 - Current route: comparative cognition tool use, endogenous generation creativity, learning science skill acquisition.
 - Missing depth: Cognition, creativity and learning are represented; clinical, personality, developmental, organizational, social and cross-cultural psychology remain uneven.

Taxonomy disagreement is a discovery signal

A field missing from one classification must not disappear from the research program. The first recorded disagreement is:

- **ANZSRC 2020 45 Indigenous Studies — dedicated audit.**
 - Current route: indigenous data knowledge governance.
 - Gap: Indigenous data and knowledge governance now has an Indigenous-led-source audit. Indigenous Studies is far broader: language, law, health, education, land, history, methods, and community-specific knowledge remain outside this narrow entry audit and require community-specific authority.
 - Next discriminating question: How can the breadth process include community-defined research questions and authorities without turning Indigenous Studies into extractable motifs or a universal policy template?
 - Handling rule: Treat this as a governance and knowledge-sovereignty audit led by Indigenous-authored sources; taxonomy coverage never authorizes extraction of biological or cultural motifs.

Breadth scheduler

Each breadth wave must contain four different selections:

1. one field from the least-covered OECD broad field;
2. one EuroSciVoc concept or DFG subject hidden inside a parent already marked dedicated;
3. one field chosen for methodological distance from the current corpus, without requiring an obvious AI analogy; and
4. one depth item that can alter an existing claim, experiment, equation, or system boundary.

The first three selections collect leads. Only the fourth slot is allowed to consume substantial depth work before deduplication. This prevents familiar fields from monopolizing the research budget while still requiring useful residue to change a testable object.

Completed breadth waves

Wave 1 — taxonomy and method boundaries

1. **Soil, crop, and multi-resource co-limitation:** changed resource state from scalar availability to typed, transport- and service-window-qualified vectors; no new principle.
2. **Philosophy of science and theory choice:** changed discovery admission from model fit to an alternative-, auxiliary-, observation-, access-, and failure-root-qualified contract; no new principle.
3. **Indigenous data and knowledge governance:** added collective authority, purpose, benefit, refusal, and derivative-remedy obligations while prohibiting extraction of community knowledge; no new principle.
4. **Textual criticism and variant traditions:** separated attestation, transcription, collation, lineage, conjecture, edition, translation, and interpretation and made apparatus compaction query-qualified; no new principle.

Wave 2 — deliberate disciplinary distance

1. **Nursing, care science, and health services:** made unfinished necessary work, continuity type, receiver acceptance, response capacity, and verified service explicit; five synthetic protocols, no new principle.
2. **Animal production and veterinary population health:** protected multidimensional welfare from productivity proxies and added cohort, lineage, interference, consortium-flux, breeding-diversity, and changing-opponent tests; no new principle.
3. **Environmental and industrial biotechnology:** added dilution/selection, evolutionary escape, scale-gradient, transport, remediation-rebound, measurement, and lifecycle falsifiers while leaving chemistry boards conservatively open; no new principle.
4. **Accounting, audit, actuarial science, and insurance:** separated balance from truth, materiality from detection, reserves from headroom, dependence from pooling, and governance from model validity; nine protocols, no new principle.
5. **Theology, religious practice, and ritual:** separated instrumental effect, conventional form, protected commitment, authority, interpretation, and belief data; five synthetic protocols, no new principle.
6. **Forestry, fisheries, aquaculture, and aquatic food systems:** added stock–flow–cohort–space–rights, carbon-pool, management-loop, externality, nutrient-recovery, and edible-endpoint tests; six protocols, no new principle.
7. **Particle, nuclear, and high-energy experimentation:** added trigger support, detector response, unfolding, nuisance, blinding, search-family, simulation, preservation, covariance, and effective-model breakdown tests; eight protocols, no new principle.

Wave 3 — independent taxonomy and remaining empty cells

1. **Residual humanities and living heritage:** separated archival selection from event absence, contextual authenticity from a universal score, living practice from frozen tokens, rehearsal history from free compression, re-treatability from byte rollback, and representation from identity; six protocols, no new principle.
2. **Molecular chemistry and synthesis systems:** bounded dynamic covalent assembly, self-sorting, context-qualified recognition, temporal pathway control, structure elucidation, and automated discovery against mature chemical nulls; eight protocols, no new principle.
3. **Polymer research:** made distributions, mechanism, topology, relaxation spectra, phase path, gel criteria, healing, ageing, sequence storage, circularity, and measurement operators explicit; twelve protocols, no new principle.
4. **Mineralogy, petrology, and geochemistry:** separated equilibrium, kinetic trapping, material replacement, P–T–t records, tracer inverses, weathering, redox, upscaling, provenance, dates, and preservation-filtered deep time; eight protocols, no new principle.
5. **Direct social research, ethnography, and media:** separated instrumented response, nonresponse, category construction, measurement invariance, situated observation, reflexivity, talk/action, network interference, formal routine, media selection/exposure, mixed-

method failure roots, and distinct authority bases; eight protocols, no new principle.

- 6. Taxonomy control itself:** added 1,064 EU EuroSciVoc concepts plus all 23 ANZSRC divisions, 213 groups, and 1,967 fields beside OECD and DFG; parent coverage never substitutes for concept-, subject-, group-, or field-level review.

Wave 4 — intervention, production, and institutional depth

- 1. Clinical intervention pathways:** separated multimorbidity composition, diagnostic performance from pathway utility, prognostic from predictive markers, sequential-policy support, antimicrobial patient/population outcomes, and recommendation from execution authority; nine claims and six synthetic protocols, no new principle.
- 2. Production, maintenance, nano, communications, and material qualification:** separated statistical control from specification and qualification, prognosis from decision value, images from nano-population evidence, physical-layer rate from accepted service, interacting automation loops, and lineage from qualification; nine claims and nine CPU-only protocols, no new principle.
- 3. Institutions, inequality, crime, and digital media:** separated formal policy from delivered service, inequality operators, household totals from allocation, visible action from private support, recorded crime from incidence, platform traces from populations, and media-system transport; seven claims and eight CPU-only protocols, no new principle.

Wave 5 — targeted remainder and depth audits

- 1. Analytical measurement, water, atmosphere, ocean, and Earth observation:** added operator, calibration, reference-material, speciation, transport, inverse-resolution, support, and interlaboratory boundaries; nine claims and nine CPU-only protocols, no new principle.
- 2. Clinical specialties plus medical and agricultural biotechnology:** bounded surgery, imaging, paediatrics, psychiatry, obstetrics, dentistry, gene/cell products, genome editing, biological control, field heterogeneity, and biosafety; nine claims and nine protocols, no new principle.
- 3. Economy, education, and institutions:** separated tax incidence from statutory rates, public accounts from capacity, organizational bundles from labels, labour stocks from task flows, educational assessment from use, and comparison from transport; nine claims and nine protocols, no new principle.
- 4. Compilers, networking, visualization, and software engineering:** added semantic-preservation, workload-qualified optimization, randomized layout inference, endpoint placement, congestion/tail, graphical-perception, and lifecycle-quality tests; nine claims and ten protocols, no new principle.
- 5. Medical devices and biomedical engineering:** separated signal reconstruction, device decision, clinical utility, human factors, degradation, interoperability, cybersecurity, and lifecycle qualification; nine claims and nine protocols, no new principle.
- 6. Atomic, molecular, condensed-matter, and plasma science:** bounded spectroscopy, collision operators, phase/regime transitions, disorder, multiscale closure, reduced models, and diagnostic support; nine claims and nine protocols, no new principle.
- 7. Political, administrative, stratification, and media depth:** added electoral-operator provenance, selected roll calls, target gaming, task-indexed capacity, wealth-tail uncertainty, stage-bounded discrimination, cumulative inequality, audience deduplication, and moderation denominators; nine claims and nine protocols, no new principle.

Next gap wave

- 1. Subfield resolution inside green cells:** sample individual EuroSciVoc concepts, DFG subjects, and ANZSRC groups in food chemistry, marine and atmospheric process science, structural biology, ecological/molecular/developmental physiology, environmental treatment, health implementation, and applied mathematics; an entry audit does not close its parent field.

2. **Clinical and biotechnology continuation:** deepen specialty-specific endpoints, longitudinal harms, manufacturing drift, combination products, paediatric/rare strata, field ecology, gene-flow containment, and post-deployment surveillance beyond the new entry audit.
3. **Computing and engineering remainder:** deepen kernels, schedulers, distributed transactions and streaming, empirical software engineering, accessible visualization, machine design beyond tribology, thermal machinery, factory and layout systems, surface chemistry, manufacturing variation, long-duration interface qualification, hardware ageing, and device security with executable boundaries.
4. **Institutions, economy, education, and law:** deepen parties, executives, courts, comparative politics, international relations, banking/finance, industrial organization, education stages and didactics, substantive EU/German law, and community-specific methods.
5. **Subfield resolution plus execution:** route EuroSciVoc concepts, DFG subjects, and ANZSRC groups explicitly, then convert the most discriminating protocols into versioned workstation artifacts instead of accumulating prose-only readiness.

No item is promoted because it sounds novel. Every retained mechanism still passes the open-world extraction record, mature-null comparison, deduplication, and equal-budget rejection gate in the discovery policy.

Maintenance rule

After each audit, update the machine record first, regenerate this page and plot, and record whether the new material produced: a duplicate, a sharper boundary, a new claim, a changed experiment, or no durable residue. A field remains incomplete even after its first dedicated audit.

Licences and notices Source: github.com/lusoris/20-watts-was-enough @ main **EUPL 1.2**

CC BY-SA 4.0 **Licensing scope** **Third-party notices**