Robustness Before Performance

A formally specified, governed viability kernel for decisions under disturbance

Justin D. Hart
Viridis Research
ORCID: 0009-0008-3082-2482

7 August 2026
v1 successor candidate; formal kernel independently Aristotle-reviewed and Viridis-audited

Abstract

Efficiency, specialization, and maximum output are useful objectives only inside a sufficiently trustworthy frame. Under deep disturbance, a narrowly optimized system may lose essential function because it has removed slack, hidden a common dependency, suppressed a fallback, or exported fragility. Building from Olivier Hamant's biological account of robustness, this paper defines robustness as governed viability rather than a universal score. The method first declares the system boundary, essential functions and rights, stakeholders, scales, disturbance set, time horizon, evidence state, and authority. It then expands dependency ancestry, blocks recalled evidence, classifies closed outcome intervals, analyzes sampled disturbance trajectories, rejects non-compensable breaches, compares only robustly separated profiles, and records selection separately from execution. A deterministic Python 0.6.0 reference implementation passes 35 of 35 structural tests on synthetic municipal-heat and cyber-identity cases. Its Viridis agent service passes 5 focused tests and the complete isolated fleet passes 2,136 tests across 37 agents. The same source is deployed as a digest-bound, internal-only, non-actuating service; final health, a canonical evaluation, fleet-network reachability, and an eight-second stop/start rollback drill pass. A pinned Lean 4 successor proves 22 frozen statements in nine claim families concerning dependency diversity, authority, epistemic-state separation, cascading recall, closed-interval viability, robust dominance, trajectory shortfall, aftershock bounds, and release-state separation. The independent Aristotle review preserved every frozen statement and added 19 non-degeneracy checks; the subsequent Viridis rebuild completed 930 jobs with zero proof holes, no forbidden constructs, and only the permitted foundational axioms. The runtime receipt establishes deployment separately from proof. None of these results proves empirical completeness, political legitimacy, calibrated uncertainty, adoption, revenue, or beneficial real-world outcomes.

Keywords: robustness; viability; resilience; governance; formal verification; common-mode failure; interval uncertainty; trajectory analysis

1 Problem and intellectual lineage

Performance under expected conditions and continued identity under disturbance are different properties. Highly Optimized Tolerance describes systems that are robust to designed-for uncertainty yet unusually fragile to unexpected perturbations or design errors [1, 2]. Biological robustness is likewise maintained through mechanisms that carry performance, resource, and newly exposed fragility trade-offs [5].

Hamant's contribution is especially generative for design. His plant-science argument rejects frictionless biological optimization: stochasticity, heterogeneity, delays, incoherence, and apparent inefficiency can participate in robust function [3]. In the Sismique interview that initiated this project, he frames robustness as maintaining viability amid fluctuation, preserving maneuvering room, and maintaining multiple paths instead of only a shortest path [4]. These observations motivate engineering hypotheses; they do not transfer biological facts directly into political authority, justice, building codes, or security policy.

The design problem is therefore not "maximize robustness." A harmful institution may be robust; redundant services may share one identity root; and local continuity may export risk to another population or time. The required artifact is a governed decision contract that makes identity, thresholds, stakeholders, failure domains, evidence, and authority inspectable.

2 First-principles constitution

Let a decision frame be

$$\mathcal{F} = (X, I, D, T, L, K, E, A), \tag{1}$$

where X is the system and boundary, I essential functions, rights, or identity conditions, D declared disturbances, T time horizon, L spatial and organizational scales, K affected stakeholders, E evidence and model state, and A authority and governance. A robustness statement that omits these fields is incomplete.

For candidate p, invariant i, scenario s, stakeholder or scale k, and time t, a directional normalized viability margin is

$$m_{i,k}(p, s, t) = \sigma_i \frac{o_{i,k}(p, s, t) - \theta_{i,k}}{q_{i,k}}, \qquad q_{i,k} > 0, \tag{2}$$

where $\sigma_i \in \{-1, 1\}$ orients the margin so that positive is safer, o is the declared outcome, θ the threshold, and q a diagnostic normalizer. Normalization does not make distinct rights or harms morally commensurable.

The operational definition is:

> A system is robust when specified essential functions, rights, or identity conditions remain within acceptable bounds, for specified stakeholders and scales, across a declared range of disturbances and time horizons, without exporting unacceptable fragility elsewhere.

The kernel orders decision conditions lexicographically:

$$\text{authority and legitimacy} \succ \text{rights and safety} \succ \text{mandatory viability} \succ \text{robustness profile} \succ \text{nominal performance}. \tag{3}$$

If the accountable process marks an invariant non-compensable, cost, speed, output, or average welfare may not offset its breach. Software does not decide which conditions deserve this status; it enforces the declared status without silently converting it into a weighted penalty.

3 Kernel

3.1 Dependencies and lineage

Every fallback is closed over a typed acyclic dependency graph. For expanded finite dependency sets L and R, the diagnostic independence is

$$\mathcal{I}_D(L, R) = 1 - \frac{|L \cap R|}{|L \cup R|}. \tag{4}$$

It is not a probability. Shared owners, roots of trust, control planes, locations, energy supplies, suppliers, and maintenance teams can turn nominal redundancy into a common-mode failure. Missing ancestry remains a real-world model risk.

Outcome bundles bind models and adapters through a content-addressed lineage DAG. An effective recall blocks the recalled artifact and every declared descendant before evaluation. Unrelated nodes are not blocked solely by that recall relation. Hash authenticity and recall authority remain external verification duties.

3.2 Interval viability and robust comparison

An assessment may declare a closed interval $[\ell, u]$, $\ell \le u$, rather than a single falsely precise value. For a lower-is-safer constraint such as $o \ge \theta$, the classification is

$$\begin{cases} \texttt{ROBUST\_PASS}, & \ell \ge \theta, \\ \texttt{ROBUST\_FAIL}, & u < \theta, \\ \texttt{UNCERTAIN}, & \ell < \theta \le u. \end{cases} \tag{5}$$

The four inequality operators and exact equality use their corresponding closed-boundary rules. Mandatory non-compensable uncertainty yields `HOLD`. The intervals are declared bounds, not probability distributions or calibrated confidence intervals.

For a maximize dimension, candidate a is robustly no worse than b only if $\ell_a \geq u_b$; for a minimize dimension only if $u_a \leq \ell_b$. Robust dominance requires this separation in every dimension and strict separation in at least one. Overlap remains unresolved. The relation is irreflexive and transitive under the frozen definitions.

3.3 Sampled trajectories

For ordered samples (t_j, m_j), the reference engine declares piecewise-linear interpolation between samples. Its shortfall area is

$$S = \sum_{j=0}^{n-2} \int_{t_j}^{t_{j+1}} \max(0, -\widetilde{m}_j(t))\, dt \geq 0, \tag{6}$$

where $\widetilde{m}_j$ is the linear segment. The engine separately reports possible and robust violation spans using conservative endpoint-state marking, the earliest declared guaranteed recovery after the worst sample, and a count of pass-to-nonpass adjacent transitions. A governance-declared coverage floor prevents sparse mandatory trajectories from impersonating complete temporal analysis. Samples and interpolation describe the supplied path, not the unobserved world.

3.4 Epistemic and action firewalls

The record keeps nine independent claim states: `ASSUMED`, `HYPOTHESIZED`, `NUMERICALLY_SUPPORTED`, `REPRODUCED`, `FORMALLY_PROVED`, `EMPIRICALLY_SUPPORTED`, `AUTHORIZED`, `DEPLOYED`, and `OBSERVED`. Membership in one state supplies no witness for a different state.

For selected candidate p and authority lease λ,

$$\begin{aligned} \mathrm{Exec}(p, \lambda) = \ & \mathrm{Selected}(p) \wedge \neg\lambda.\mathrm{advisoryOnly} \wedge \lambda.\mathrm{actionAllowed} \\ & \wedge \lambda.\mathrm{active} \wedge \lambda.\mathrm{scopeMatches} \wedge \lambda.\mathrm{unexpired}. \end{aligned} \tag{7}$$

The reference kernel and agent service never execute an action. They return an action boundary and the conditions that failed or passed.

4 Formal specification and proof status

The successor Lean namespace is `Viridis.RobustnessKernelV1`. Table 1 binds every theorem-shaped paper claim to a frozen claim family. The statement contract contains 22 required declarations and five additional supporting declarations.

ID	Mechanism	Machine-checked conclusion
F1	dependency diversity	shared ancestry lowers finite-set independence
F2	execution authority	execution iff selection and every lease condition pass
F3	epistemic firewall	explicit membership creates no distinct-state membership
F4	cascading recall	recalled roots and descendants block; unrelated nodes do not
F5	interval viability	pass and fail are disjoint; classification is complete and exact
F6	robust dominance	frozen separation relation is irreflexive and transitive
F7	trajectory shortfall	shortfall and marked span are nonnegative; safe traces have zero shortfall
F8	aftershock bound	transition count does not exceed adjacent sample pairs
F9	release-state separation	proof, selection, authorization, deployment, and observation have distinct witnesses

Table 1: Frozen formal claim families.

The frozen project is pinned to Lean 4.28.0 and mathlib revision `8f9d9cff6bd728b17a24e163c9402775d9e6a365`. It builds successfully in 920 jobs. A source scan reports zero proof holes or forbidden escape constructs; the axiom audit reports only `propext`, `Classical.choice`, and `Quot.sound`, and the cross-kernel witness prevents an empty

specification from passing as the result. The frozen Lean source SHA-256 is `ef373a04b19dff71bf02632d0e413487 9069c31ab716e5f669483ab33c18802f`.

Aristotle project `3e142131-f6f4-4c32-b273-873783735a69` independently reviewed the exact request whose SHA-256 begins `addd4da10c171161`. The returned kernel is byte-identical to the frozen source. Aristotle added 19 review-check theorems that demonstrate non-degenerate values and both granting and denying cases across F1–F9. Viridis then rebuilt the returned project in 930 jobs, compared all 27 theorem statements, scanned raw and comment-stripped source, and audited axioms, non-vacuity, and significance. All seven audit gates passed. The exact forge-audit SHA-256 is `fc5e6c7d21fcb3c3c2ec5ab5b06dc3590ae0aab99600fa375c9a31fca562a98e`.

These are therefore machine-checked and independently audited claims over the frozen mathematical objects. No proof result can establish that a dependency map is complete, evidence is true, intervals are calibrated, authority is legitimate, a service was deployed, or an outcome improved.

5 Reference implementation and fleet service

The deterministic Python 0.6.0 kernel validates closed JSON schemas, expands dependencies, enforces content-addressed recall, classifies intervals, analyzes trajectories, applies hard gates, builds profiles, performs robust comparison, and emits canonical decision records. On 7 August 2026, 35 of 35 reference tests passed. The municipal-heat and cyber-identity fixtures are synthetic and materially different; no domain-specific branch exists in the kernel. Their canonical record hashes are respectively `3badd072093e42958cdeb1f1df65eca12742235b235f3b4374e09f292aef7783` and `ed2e6726571d2239ed2e90839435ea06ed887f371f791c3a0cb4e43df23c1b79`.

A Viridis agent service separates the pure kernel from HTTP transport, accepts finite canonical JSON under a request-size limit, exposes description, health, and evaluation interfaces, and requires no secret, external network call, model call, or mutable database. Five focused service tests pass. The authoritative isolated fleet runner passes 2,136 tests with zero failures or errors across 37 agents.

On 7 August 2026, the exact locked AMD64 image was deployed as an internal-only service on the Viridis fleet network. Its read-only root filesystem, dropped capabilities, no-new-privileges setting, lack of published host ports, digest-bound final health, gateway reachability, and canonical cyber evaluation passed. The decision-record hash was `cb45c00a79c72c43fb86995097ae61c19be9fb08a37eae014ed80e230d4ee3cf`, with execution disabled. A stop/start drill restored the same image in eight seconds while existing fleet services remained healthy. This receipt establishes deployment, not empirical validation, adoption, authority, or beneficial outcome.

6 Cross-domain use

The kernel remains stable while accountable domain adapters supply the semantics, evidence standards, models, disturbances, and thresholds.

Primitive	Societal governance	Cybersecurity	Architecture/infrastructure	Product design
Invariants	rights, minimum services, legitimacy, ecological limits	confidentiality, integrity, availability, recovery	life safety, habitability, accessibility	user safety, core job, compliance, repairability
Disturbance	disaster, recession, epidemic, legitimacy loss	fault, compromise, ransomware, dependency loss	fire, flood, heat, utility loss, degradation	misuse, supplier loss, recall, repair scarcity
Failure domain	finance, logistics, law, information	root identity, DNS, admin plane, cloud, power	grid, water, egress, structure, contractor	supplier, firmware, platform, proprietary tool
Authority	democratic or delegated authority, due process, appeal	risk owner, security authority, incident command	occupants, regulators, engineers, community	product and safety owners, users, regulators

NIST's cyber-resiliency approach already organizes systems engineering around anticipating, withstanding, recovering from, and adapting to adverse conditions [6]. The present work does not replace domain standards. It adds a cross-domain constitutional ordering, evidence-state separation, transitive failure-domain diagnostic, non-compensation rule, and hash-bound release boundary.

7 Falsification and validation

The system should be restricted or revised if preregistered studies show that:

1. independent reviewers cannot reproduce records from frozen inputs;
2. dependency closure reveals no additional actionable common-mode exposure over competent review;
3. correctly encoded gates admit known non-compensable failures or reject safe options systematically;
4. irrelevant identifiers or input ordering change the result;
5. users confuse proof, authorization, deployment, and observation despite the explicit state model;
6. prospective use fails to improve auditability, invariant preservation, recovery, or off-boundary harm relative to competent baselines; or
7. affected stakeholders expose systematic fragility export omitted by the declared frame.

The next empirical step is a preregistered paired study: one authorized public or infrastructure decision and one authorized cyber architecture review, frozen before outcomes, each with a competent baseline, independent reviewers, withheld disturbances, stop rules, and preserved dissent. The experiment must not acquire execution authority.

8 Limitations

Biological persistence does not determine justice or purpose. Robustness can protect the wrong identity. Redundancy and slack consume energy, material, money, and attention and can add attack surface. Optimization remains valuable inside a sound admissible region. Dependency ancestry, scenario coverage, thresholds, intervals, samples, and externalities are supplied and may be wrong. Formal proof checks declared mathematics; it does not authenticate the world. Software tests check behavior on declared cases; they do not validate causal claims or cross-domain benefit. Production deployment is not empirical validation, adoption, authority, or observed benefit. Publication and canon admission require explicit rights, a real ledger row, and accountable authorization.

9 Conclusion

Hamant's provocation can be operationalized without pretending that plants write constitutions. Robustness becomes a governed relationship: what must remain viable, for whom, over which scales and disturbances, with which paths, evidence, uncertainty, and authority. The v1 kernel turns that relationship into deterministic gates, profiles, lineage controls, and action boundaries, and freezes its theorem-shaped core for independent verification. Its most important feature is restraint: performance is optimized only after viability passes, unresolved uncertainty remains visible, proof does not become authority, and a decision service does not become an actuator.

AI and formal-methods disclosure

The accountable author directed the project and remains responsible for the claims and release decisions. OpenAI Codex assisted with source synthesis, specification, code, tests, manuscript drafting, formalization, and independent audit execution. Aristotle independently reviewed the frozen Lean project and added non-degeneracy review certificates; Viridis then rebuilt and audited the return against the immutable request. AI outputs are not treated as empirical evidence, authorization, deployment, canon admission, publication, or revenue.

References

[1] J. M. Carlson and J. Doyle, "Highly optimized tolerance: robustness and design in complex systems," *Physical Review Letters*, 84, 2529–2532, 2000. doi:10.1103/PhysRevLett.84.2529.

[2] J. M. Carlson and J. Doyle, "Complexity and robustness," *Proceedings of the National Academy of Sciences*, 99(Suppl. 1), 2538–2545, 2002. doi:10.1073/pnas.012582499.

[3] O. Hamant, "Debunking the idea of biological optimisation: quantitative biology to the rescue," *Quantitative Plant Biology*, 5, e3, 2024. doi:10.1017/qpb.2024.3.

[4] J. Devaureix and O. Hamant, "Performance, turbulences et robustesse," Sismique, interview 140, recorded 30 May 2024, published 19 July 2024. transcript and audio.

[5] H. Kitano, "Biological robustness," *Nature Reviews Genetics*, 5, 826–837, 2004. doi:10.1038/nrg1471.

[6] R. Ross, V. Pillitteri, R. Graubart, D. Bodeau, and R. McQuaid, *Developing Cyber-Resilient Systems*, NIST SP 800-160 Vol. 2 Rev. 1, 2021. doi:10.6028/NIST.SP.800-160v2r1.