Error as Obstruction

Flyxion

Independent Researcher

April 21, 2026

Abstract

We develop a unified geometric account of error, learning, and event-sourced computation in which failure of coherence is identified not with deviation from a ground-truth value but with a topological impossibility of extension. The central claim is that an event log, a training dataset, or any preserved history generates a constraint sheaf over a domain; admissible states are precisely global sections of that sheaf; and what is ordinarily called an error is a cohomological obstruction to the existence of such a section. We formalise this claim through the *Error as Obstruction* theorem, which equates the convergence of stochastic projected relaxation dynamics with the triviality of a derived cohomology class. The framework synthesises the RSVP field-space construction, the Spherepop event-calculus, and the TARTAN sheaf-consistency architecture into a single programme: intelligence is the organisation of an overcomplete constraint manifold under boundary preservation, and error is the failure of that manifold to admit a globally coherent section.

Contents

1. Introduction

1.1. The Bottleneck Assumption and Its Discontents

For most of its history, the theory of learning has been organised around scarcity. The core assumption is that the capacity to represent is limited, that experience exceeds what can be retained, and that generalisation is therefore the art of productive forgetting—of compressing many instances into a compact rule. Under this view, memory and generalisation stand in fundamental tension: a system that preserves too much detail overfits, mistaking the texture of particular examples for the structure of the underlying process. The standard prescription follows directly: reduce, simplify, discard.

This assumption is so deeply embedded that it shapes not only formal models but practical pedagogy. A struggling student is told to take fewer courses. A craftsperson is advised to specialise. A programmer is counselled to keep the codebase small. The recommendation in every case is subtraction: reduce the inputs until the system stabilises.

The present work begins from a different direction. There is a class of learning systems—and a class of learners—for which this prescription is not merely unhelpful but actively wrong. For these systems, adding more structured experience does not increase the burden; it increases the coherence. Consider a builder who knows electrical, plumbing, construction, finishing, and installation. When tasked with sorting materials, organising offcuts, or evaluating whether a scrap fitting will be useful, that person is not performing five separate assessments. Every fragment is already indexed by function, hazard, sequence, and reuse potential. The narrow specialist sees disconnected objects; the cross-trained worker sees a structured landscape. A narrow assistant either asks constantly or proceeds blindly—not because they lack facts in isolation, but because they lack the surrounding geometry that would tell them what matters, what can wait, what is hazardous, and what they do not yet know they do not know.

The same pattern appears at different scales. Switching between subjects when studying is not avoidance; it is coordinate-frame rotation, a way of keeping the system active while preventing any single track from becoming locally stale. In a namespace already crowded with thousands of projects, one does not need to remember every prior name; enough exposure to the distribution internalises it, so that a new choice can be evaluated as a position in a landscape whose rough topology is already known. The choice is no longer arbitrary but navigational.

These examples share a common structure. In each case, the system is operating above a threshold beyond which additional structured input reshapes the space of possible organisation rather than merely occupying it. Memorisation and generalisation are not in tension because the system is not a fixed container filling up; it is a geometry being refined. Capacity here is not the size of a buffer but the dimensionality of an internal coordinate system, and that dimensionality can *increase* with exposure.

1.2. The Formal Problem

The standard picture of learning identifies error with a scalar distance between prediction and target. This scalar is differentiable, and gradient descent drives it toward zero. The picture is clean, but it conceals the structural question that motivates the present work: *why does perfect interpolation not destroy generalisation?* In the excess-capacity regime—neural networks with far more parameters than training examples, event stores with richer constraint vocabularies than the log they hold—the system memorises everything and yet continues to extrapolate sensibly [3, 2, 1]. Compression cannot explain this; the system has not compressed.

The answer we develop here is geometric rather than statistical. Once the system reaches exact fit, learning does not stop; it continues as motion *within* the admissible set, driven by stochastic relaxation dynamics that seek regions of low curvature. In this regime, error is not a gradient to be cancelled but a structural impossibility: the log or dataset has imposed locally consistent constraints that cannot be glued into any globally coherent configuration. That impossibility is an obstruction in the sheaf-theoretic sense, and it is the central object of the paper.

The novelty of the claim is not that memorisation and generalisation can coexist. Any practitioner who has worked in a richly cross-constrained domain already knows this from the inside. The novelty is that their coexistence is not accidental. It is a *predictable consequence* of operating in a regime where the space of possible representations is larger than the data that constrains it—a regime with a precise geometric characterisation, a well-defined dynamical theory, and a natural notion of failure.

That notion of failure is the obstruction of the title. When a history cannot be made globally coherent—when locally consistent constraints cannot be glued into any smooth global configuration—the system does not merely give a wrong answer. It inhabits a space in which no coherent answer exists. Error, on this account, is not a mismeasurement. It is a topological fact about the structure of the record.

The exposition proceeds as follows. Section 2 establishes the mathematical setting: the RSVP field space, the admissible fit set, and the three geometric regimes of learning. Section 3 introduces the soft action functional and derives projected gradient flow on regular strata. Section 4 extends the construction to stochastic dynamics and defines the stationary density of wisdom. Section 5 develops the role of tangential entropy and shows that selection in the excess regime is a volumetric phenomenon. Section 6 articulates the principle of Log Sovereignty. Section 7 formalises the log as a boundary operator and proves historical closure. Section 8 identifies the Spherepop event-calculus as a discrete realisation of these principles. Section 9 analyses singular strata and degenerate geometry as the precursor to full obstruction. Section 10 introduces the constraint sheaf and reformulates admissibility as global section existence. Section 11 states and proves the main theorem together with the Stability–Obstruction Dichotomy. Section 12 isolates the triple diagnostic signature. Section 13 classifies obstruction types and develops a hierarchy of repair operations. Section 14 makes the viscoelastic picture exact. Section 15 connects the framework to contemporary machine learning. Section 16 states the central thesis in its strongest form. An appendix recounts the phenomenological motivation in more concrete terms for readers whose primary interest is practical rather than formal.

2. Mathematical Setting

2.1. The RSVP Field Space

Let $\Omega \subset \mathbb{R}^n$ be a domain. A *field configuration* is a triple $\Phi = (\phi, \mathbf{v}, S)$ where $\phi : \Omega \to \mathbb{R}$ is a scalar density field, $\mathbf{v} : \Omega \to \mathbb{R}^n$ is a vector flow field, and $S : \Omega \to \mathbb{R}_{\geq 0}$ is an entropy density. The configuration space $\mathcal{X}$ is a Fréchet manifold whose local charts are determined by Sobolev norms on each component.

A readout operator $R : \mathcal{X} \to \mathbb{R}^m$ extracts observable predictions from a configuration. The experience set $D = \{(x_i, y_i)\}_{i=1}^N$ specifies N input-output pairs. We write $\hat{y}_i = R(\Phi)(x_i)$ for the prediction at x_i.

Definition 2.1 (Admissible Fit Set)**.** The *admissible fit set* for D is

$$\mathcal{M}(D) = \left\{ \Phi \in \mathcal{X} \;\middle|\; R(\Phi)(x_i) = y_i \text{ for all } i = 1, \dots, N \right\}.$$

$\mathcal{M}(D)$ is the zero-level set of the constraint map $\Phi \mapsto (R(\Phi)(x_i) - y_i)_{i=1}^N$ and, for smooth R, is generically a submanifold of $\mathcal{X}$ whose codimension equals the number of active constraints.

2.2. The Three Geometric Regimes

Constrained regime. The constraints over-determine $\mathcal{X}$: $\mathcal{M}(D) = \emptyset$, and the system can only minimise a soft penalty.

Sufficient regime. The number of constraints matches the effective degrees of freedom: $\mathcal{M}(D)$ is generically a discrete set of isolated points.

Excess regime. The system has far more degrees of freedom than constraints: $\mathcal{M}(D)$ is a high-dimensional manifold, and learning continues as navigation *within* $\mathcal{M}(D)$. This is not a pathology but a phase transition in the geometry of the problem [2, 1].

3. Projected Gradient Flow

3.1. The Soft Action Functional

Let $\mathcal{A}_{\text{soft}} : \mathcal{X} \to \mathbb{R}_{\geq 0}$ be a smooth, coercive *soft action functional* encoding inductive preferences beyond exact fit.

Definition 3.1 (Restricted Hessian)**.** For $\Phi \in \mathcal{M}(D)$, let $T_\Phi \mathcal{M}(D)$ denote the tangent space and let Π_Φ denote orthogonal projection onto it. The *restricted Hessian* is

$$\text{Hess}_\Sigma \mathcal{A}_{\text{soft}}(\Phi) = \Pi_\Phi \, D^2 \mathcal{A}_{\text{soft}}(\Phi) \, \Pi_\Phi.$$

Definition 3.2 (Flat Region)**.** A stratum $U \subset \mathcal{M}(D)$ is *flat* if $\text{Hess}_\Sigma \mathcal{A}_{\text{soft}}(\Phi) = 0$ for all $\Phi \in U$, and *low-curvature* if its operator norm is bounded by $\varepsilon \ll 1$ uniformly on U.

3.2. Projection-Relaxation Dynamics

The intramanifold dynamics are governed by the *projected gradient flow*:

$$\dot{\Phi}(t) = -\,\Pi_{\Phi(t)}\, \nabla \mathcal{A}_{\text{soft}}(\Phi(t)).$$

This suppresses motion normal to $\mathcal{M}(D)$—directions that would violate log constraints—while allowing free motion in tangential directions. The normal component of the gradient is the *restriction bias*; the tangential component is the *soft inductive bias.*

Proposition 3.3 (Monotonicity)**.** *Along any trajectory of the projected gradient flow,* $\frac{d}{dt}\mathcal{A}_{\text{soft}}(\Phi(t)) \leq 0$*, with equality if and only if* $\Phi(t)$ *is a critical point of* $\mathcal{A}_{\text{soft}}|_{\mathcal{M}(D)}$.

Proof. By the chain rule, $\frac{d}{dt}\mathcal{A}_{\text{soft}}(\Phi) = \langle \nabla \mathcal{A}_{\text{soft}}(\Phi), \dot{\Phi}\rangle = -\|\Pi_\Phi \nabla \mathcal{A}_{\text{soft}}(\Phi)\|^2 \leq 0.$ □

4. Stochastic Relaxation and the Stationary Density of Wisdom

4.1. The Stochastic Projected Relaxation Equation

We consider the *stochastic projected relaxation equation*

$$d\Phi_t = -\Pi_{\Phi_t}\, \nabla \mathcal{A}_{\text{soft}}(\Phi_t)\, dt + \sqrt{2\tau}\, \Pi_{\Phi_t}\, dW_t,$$

where W_t is a cylindrical Wiener process on $\mathcal{X}$ and $\tau > 0$ is a temperature parameter.

4.2. The Fokker–Planck Equation on a Smooth Stratum

For a compact smooth stratum $\Sigma \subset \mathcal{M}(D)$, the density satisfies

$$\partial_t \rho_\tau = \operatorname{div}_\Sigma(\rho_\tau\, \nabla_\Sigma \mathcal{A}_{\text{soft}}) + \tau\, \Delta_\Sigma \rho_\tau,$$

where ∇_Σ and Δ_Σ are the intrinsic gradient and Laplace–Beltrami operator on Σ. As $\tau \to 0$, the stationary measure concentrates on the sub-level sets of $\mathcal{A}_{\text{soft}}|_\Sigma$ [20, 17].

Definition 4.1 (Wisdom). A configuration $\Phi \in \mathcal{M}(D)$ is *wise* if it lies in a low-curvature region $U \subset \mathcal{M}(D)$ on which ρ_τ concentrates uniformly for all sufficiently small $\tau > 0$. A system is wise if its long-run state distribution is concentrated in flat strata.

The geometric principle is: *a stable world-state is not one that is simply true, but one that is truest across the widest range of possible tangential interpretations.*

5. Entropy, Measure, and Selection

5.1. Entropy as Tangential Volume

The stochastic projected relaxation dynamics induce a probability measure ρ_τ supported on $\mathcal{M}(D)$. While $\mathcal{A}_{\text{soft}}$ governs energy, the geometry of $\mathcal{M}(D)$ itself determines entropy. In the excess regime the admissible set carries a nontrivial volume form, and the system's long-run behaviour reflects a balance between energetic preference and geometric multiplicity.

Definition 5.1 (Tangential Entropy). Let $\Sigma \subset \mathcal{M}(D)$ be a smooth stratum. The *tangential entropy* at $\Phi \in \Sigma$ is

$$\mathcal{S}_{\tan}(\Phi) = -\int_{U_\Phi} \rho_\tau(\Psi)\, \log \rho_\tau(\Psi)\, d\Psi,$$

where U_Φ is a neighbourhood in Σ and $d\Psi$ is the Riemannian volume element induced by the ambient metric on $\mathcal{X}$.

Flat regions maximise $\mathcal{S}_{\tan}$ by supporting large volumes of near-equivalent configurations. Stochastic dynamics therefore do not merely descend energy gradients; they select regions of maximal admissible volume.

5.2. Selection as Measure Concentration

The stationary density

$$\rho_\tau(\Phi) \;\propto\; \exp\!\Big(-\frac{\mathcal{A}_{\text{soft}}(\Phi)}{\tau}\Big)$$

assigns weight not only by energy but by the measure of neighbourhoods. Regions of low curvature dominate because they occupy disproportionately large volume in Σ. This establishes the following principle.

Proposition 5.2 (Volumetric Selection)**.** *In the excess regime, the stationary measure ρ_τ concentrates on low-curvature strata of $\mathcal{M}(D)$ as $\tau \to 0$, independently of whether those strata contain the global minimum of $\mathcal{A}_{\text{soft}}$.*

Proof. By Laplace's method on the manifold Σ, the mass assigned to a region $U \subset \Sigma$ under ρ_τ scales as $\exp(-\mathcal{A}_{\text{soft}}(\Phi_0)/\tau) \cdot \text{vol}(U) \cdot (1 + O(\tau))$, where Φ_0 is the energy minimiser in U and $\text{vol}(U)$ is the Riemannian volume. A flat region with large volume can dominate a sharp minimum with small volume whenever the volume ratio exceeds $\exp((\mathcal{A}_{\text{soft}}(\Phi_{\text{sharp}}) - \mathcal{A}_{\text{soft}}(\Phi_{\text{flat}}))/\tau)$. □

Generalisation is therefore a volumetric phenomenon. A configuration is stable not because it minimises $\mathcal{A}_{\text{soft}}$ absolutely, but because it lies in a region that remains invariant under perturbation [4, 5]. The excess-capacity regime is exactly the regime in which such regions exist and are accessible; the constrained regime collapses them to isolated points.

6. Log Sovereignty

Definition 6.1 (Log Sovereignty)**.** A log D exercises *sovereignty* over a system if D is treated as a boundary condition rather than a compressible object: $\mathcal{M}(D)$ is defined by exact equality constraints derived from D, and the system's dynamics remain within $\mathcal{M}(D)$ at all times.

Under Log Sovereignty, forgetting is not a mechanism of generalisation. The log is preserved entirely—a choice with roots in the thermodynamic principle that erasure of information is the only irreversible operation [14, 15]. What generalises is the system's capacity to maintain coherence across the high-dimensional admissible structure: to find and remain in regions where $\mathcal{A}_{\text{soft}}$ is low. This dissolves the standard tension between memorisation and generalisation—in the excess regime, $\mathcal{M}(D)$ is large enough to contain both, and the stochastic relaxation selects the stable configurations by navigating within the constraint manifold toward its flattest regions.

7. Boundary Conditions and Historical Closure

7.1. The Log as a Boundary Operator

The event log D is not merely a set of constraints but a *boundary operator* acting on the field space $\mathcal{X}$. It defines the admissible set as a boundary-constrained subspace:

$$\mathcal{M}(D) = \ker \partial_D,$$

where $\partial_D : \mathcal{X} \to \mathbb{R}^N$ maps configurations to constraint violations, $\partial_D(\Phi) = (R(\Phi)(x_i) - y_i)_{i=1}^N$.

This perspective makes explicit that D does not describe the state; it *restricts* the space of possible states. The state is never the log; the log is the boundary within which the state lives.

7.2. Historical Closure

Definition 7.1 (Historical Closure)**.** A system exhibits *historical closure* if all admissible configurations are determined by D as a boundary condition, and all dynamics preserve membership in $\mathcal{M}(D)$ for all time.

Under historical closure, the past is not stored as data to be retrieved but enforced as geometry. The admissible manifold is the closure of history under the dynamics.

Proposition 7.2 (Invariance under Projected Relaxation)**.** *Historical closure is equivalent to the invariance of $\mathcal{M}(D)$ under the projected relaxation dynamics.*

Proof. The projected gradient flow $\dot{\Phi} = -\Pi_\Phi \nabla \mathcal{A}_{\text{soft}}(\Phi)$ removes all components of the gradient normal to $\mathcal{M}(D)$ by construction. Therefore, if $\Phi(0) \in \mathcal{M}(D)$, then $\Phi(t) \in \mathcal{M}(D)$ for all $t \geq 0$. Conversely, any dynamics that preserve $\mathcal{M}(D)$ must remove normal components, which is exactly the projection condition. □

This formalises Log Sovereignty at the level of dynamics rather than intent. The log is not a policy that the system chooses to respect; it is a geometric constraint that the dynamics cannot exit. Sovereignty is structural, not volitional.

7.3. Remark on Discrete Logs

In the continuous field-theoretic setting, ∂_D is a smooth operator and $\mathcal{M}(D)$ is a smooth submanifold at regular values. In the discrete event-sourced setting of Section 8, the same structure holds locally: each event e_i contributes one constraint, and the admissible set is the intersection of the corresponding zero-level sets. The boundary operator formalism is therefore not an artefact of the continuous setting but the common structural core of both.

8. Spherepop as Discrete Projected Relaxation

8.1. The Event Log as Boundary Condition

In the Spherepop calculus, a world-state is defined by its *identity as event history*. Let $D = (e_1, \ldots, e_N)$ be an ordered event sequence. Each e_i imposes the constraint that Φ is consistent with e_i. The admissible world-state set

$$\mathcal{M}(D) = \{\, \Phi \in \mathcal{X} \mid \Phi \text{ is consistent with } e_i \text{ for all } i \,\}$$

is the construction of Section 2. Reconstruction of state from the log is not deterministic replay but a search for an admissible configuration.

8.2. The Dual-Rate Projection-Relaxation Cycle

When a new event e_{N+1} arrives, the admissible set transitions from $\mathcal{M}(D)$ to $\mathcal{M}(D \cup \{e_{N+1}\})$. This update has two distinct phases.

The Normal Snap (fast). The system projects its current configuration onto the updated admissible set:

$$\Phi' = \operatorname*{argmin}_{\Psi \in \mathcal{M}(D \cup \{e_{N+1}\})} \|\Phi - \Psi\|_{\mathcal{X}}.$$

This enforces logical consistency with the new event immediately and makes the system *truthful* with respect to the log.

The Tangential Drift (slow). Starting from Φ', the system runs stochastic projected relaxation on $\mathcal{M}(D \cup \{e_{N+1}\})$, seeking low-curvature regions. This makes the system *stable* under perturbation.

The Normal Snap alone does not determine behaviour. A configuration can be admissible while residing in a high-curvature, brittle region of $\mathcal{M}(D)$. The Tangential Drift converts correctness into wisdom. Systems that treat event replay as deterministic and terminal—rather than as initialisation for ongoing relaxation—conflate the two phases and thereby sacrifice stability.

9. Singular Strata and Degenerate Geometry

9.1. Stratification of the Admissible Set

In general, $\mathcal{M}(D)$ is not a smooth manifold but a *stratified space*:

$$\mathcal{M}(D) = \bigsqcup_k \Sigma_k,$$

where each Σ_k is a smooth manifold of dimension k. Singularities arise at configurations where the rank of the constraint map ∂_D drops below its generic value: these are the points at which multiple constraint surfaces are tangent, or at which the Jacobian of ∂_D loses rank. The regular stratum is the open dense subset on which the rank is maximal; all other strata are lower-dimensional and singular.

9.2. Breakdown of Tangential Structure

At a singular point $\Phi \in \mathcal{M}(D)$, the tangent space $T_\Phi\mathcal{M}(D)$ is not well-defined as a linear subspace. Instead one has a *tangent cone*: a closed set of possible limiting tangent directions that need not span a subspace. The orthogonal projection Π_Φ becomes multivalued or discontinuous, and the projected gradient flow fails to be well-posed.

Definition 9.1 (Geometric Degeneracy)**.** A configuration $\Phi \in \mathcal{M}(D)$ is *degenerate* if the projection operator Π_Φ fails to define a unique, continuous map from $T_\Phi\mathcal{X}$ into any subspace.

Degeneracy is the geometric precursor to obstruction. It is the point at which the system can no longer distinguish admissible tangential directions from inadmissible normal ones; the two have become mixed.

9.3. Dynamical Consequences at Singular Strata

Near singular strata, stochastic projected relaxation exhibits characteristic pathological behaviour:

- rapid fluctuations between competing tangent directions as the cone has multiple extreme rays;
- amplification of noise from the loss of curvature control, since the restricted Hessian $\mathrm{Hess}_\Sigma \mathcal{A}_{\mathrm{soft}}(\Phi)$ is ill-defined and cannot provide a restoring force;
- failure of local linearisation, so that first-order approximations to the dynamics are qualitatively incorrect.

These effects constitute an early warning signal: singular strata are the regions in which the dynamical consequences of potential obstruction first manifest, even before the full cohomological obstruction described in Section 11 is reached.

Remark 9.2. Regular strata support smooth relaxation; singular strata disrupt it. The transition between them is the geometric onset of error. A system whose trajectories repeatedly enter singular strata without settling is exhibiting the precursor signature of an obstructed log.

9.4. Relation to Obstruction

The singular strata of $\mathcal{M}(D)$ are precisely the configurations at which the constraint sheaf (Section 10) is most fragile. Near a singularity, the local sections of $\mathcal{F}_D$ over nearby patches become difficult to glue, because the local geometry no longer provides a consistent normal direction to guide the gluing. Full obstruction—the non-triviality of the Čech class—is the global version of the local degeneracy diagnosed here.

10. Constraint Sheaves and Global Sections

10.1. The Constraint Sheaf

Let $\mathcal{U} = \{U_\alpha\}$ be a cover of the index set $\{1, \dots, N\}$ by overlapping patches, where each U_α represents a local context in which the events $\{e_i : i \in U_\alpha\}$ are simultaneously relevant.

Definition 10.1 (Constraint Sheaf)**.** The *constraint sheaf* $\mathcal{F}_D$ assigns to each patch

U_α the locally admissible configurations:

$$\mathcal{F}_D(U_\alpha) = \left\{ \Phi|_{U_\alpha} \in \mathcal{X}|_{U_\alpha} \;\middle|\; \Phi|_{U_\alpha} \text{ is consistent with } e_i \text{ for all } i \in U_\alpha \right\},$$

together with the natural restriction maps on overlaps.

Proposition 10.2. *$\mathcal{M}(D)$ is nonempty if and only if $\mathcal{F}_D$ admits a global section.*

10.2. The Čech Cohomology of Incoherence

The obstruction to extending locally consistent sections to a global section is measured by $\check{H}^1(\mathcal{U};\mathcal{F}_D)$. A class $[\sigma] \in \check{H}^1(\mathcal{U};\mathcal{F}_D)$ is represented by a 1-cocycle $\sigma = (\sigma_{\alpha\beta})$ satisfying the cocycle condition on triple overlaps but not cobounding—not writable as $\sigma_{\alpha\beta} = \tau_\alpha - \tau_\beta$ for any assignment of local sections τ_α.

When $[\sigma] \neq 0$, local sections cannot be glued into a global one. A logical error, a hallucination, or a coherence failure is, under this analysis, a non-trivial Čech class: a mismatch on overlaps that cannot be absorbed by any choice of global configuration [9, 10, 22, 23].

11. The Error as Obstruction Theorem

Theorem 11.1 (Error as Obstruction)**.** *Let D be an event log inducing a constraint sheaf $\mathcal{F}_D$ over a cover $\mathcal{U}$, and let $\mathcal{M}(D)$ be the admissible fit set of global sections of $\mathcal{F}_D$. Let $\mathcal{A}_{\text{soft}}$ be a smooth, coercive soft action functional on $\mathcal{X}$, and consider the stochastic projected relaxation dynamics on regular strata of $\mathcal{M}(D)$.*

Then the following are equivalent:

(i) There exists a nonempty regular stratum $\Sigma \subset \mathcal{M}(D)$ supporting stable stochastic projected relaxation with a well-defined stationary measure.

(ii) The constraint sheaf $\mathcal{F}_D$ admits a global section, equivalently $[\sigma] = 0$ in $\check{H}^1(\mathcal{U};\mathcal{F}_D)$.

Conversely, if $[\sigma] \neq 0$ in $\check{H}^1(\mathcal{U};\mathcal{F}_D)$, then:

(a) $\mathcal{M}(D)$ is either empty or consists entirely of singular configurations at which $\Pi_{T_\Phi \mathcal{M}(D)}$ is not well-defined.

(b) The stochastic projected relaxation fails to converge to any stable distribution;

trajectories exhibit persistent high-action behaviour.

(c) *The action functional satisfies* $\inf_{\Phi \in \mathcal{X}} \mathcal{A}_{\mathrm{soft}}(\Phi) > 0$.

Proof. (i) ⇒ (ii). A regular stratum is a smooth submanifold; its existence implies $\mathcal{M}(D) \neq \emptyset$. By Proposition 10.2, this implies $[\sigma] = 0$.

(ii) ⇒ (i). Suppose $[\sigma] = 0$, so a global section $\Phi^* \in \mathcal{M}(D)$ exists. The implicit function theorem, applied to the constraint map $\Phi \mapsto (R(\Phi)(x_i) - y_i)$, shows that near any regular value the admissible set is a smooth submanifold. This submanifold is nonempty and compact in appropriate Sobolev topologies when $\mathcal{A}_{\mathrm{soft}}$ is coercive. Standard results on stochastic differential equations on compact manifolds guarantee the existence of a unique stationary measure for the stochastic projected relaxation.

Converse. Suppose $[\sigma] \neq 0$. Then $\mathcal{M}(D)$ contains no smooth point: any candidate global section fails the compatibility condition on at least one overlap, introducing a geometric defect. At such a defect, $T_\Phi\mathcal{M}(D)$ does not exist and $\Pi_{T_\Phi\mathcal{M}(D)}$ is undefined. The stochastic trajectories, unable to project onto a consistent tangent, accumulate action rather than dissipating it—the persistent high-action behaviour of (b). The lower bound in (c) follows from the coercivity of $\mathcal{A}_{\mathrm{soft}}$ and the absence of admissible configurations at which $\mathcal{A}_{\mathrm{soft}}$ can reach zero. □

Remark 11.2. Theorem 11.1 identifies error with the nonexistence of a smooth global section compatible with the log. This is not approximation error, prediction error, or deviation from a target value; it is structural impossibility. The system does not fail to find the right answer; it fails to inhabit a space in which any answer is coherent.

Theorem 11.3 (Stability–Obstruction Dichotomy)**.** *Let D be an event log inducing a constraint sheaf $\mathcal{F}_D$, and let $\mathcal{A}_{\mathrm{soft}}$ be a smooth, coercive soft action functional on $\mathcal{X}$. Consider the stochastic projected relaxation dynamics restricted to $\mathcal{M}(D)$.*

Then exactly one of the following holds:

(i) ***Stability.*** *The cohomology class vanishes, $[\sigma] = 0$ in $\check{H}^1(\mathcal{U}; \mathcal{F}_D)$. There exists a nonempty regular stratum $\Sigma \subset \mathcal{M}(D)$ on which the dynamics admit a unique stationary measure ρ_τ, and trajectories converge in distribution toward low-curvature regions of Σ.*

(ii) ***Obstruction.*** *The cohomology class is non-trivial,* $[\sigma] \neq 0$*. The admissible set* $\mathcal{M}(D)$ *contains no smooth stratum supporting stable dynamics, and stochastic projected relaxation fails to converge, exhibiting persistent action floor, tangent instability, and non-localising trajectories.*

These cases are mutually exclusive and exhaustive.

Proof. Mutual exclusivity. By Theorem 11.1, existence of a regular stratum with stable dynamics is equivalent to $[\sigma] = 0$, while non-triviality of the cohomology class implies the absence of such strata. The two conditions therefore cannot hold simultaneously.

Exhaustivity. For any D, either $\mathcal{F}_D$ admits a global section or it does not. In the first case $\mathcal{M}(D)$ contains smooth points and supports projected relaxation with a stationary measure by standard results on stochastic flows on compact manifolds [21]. In the second case $\mathcal{M}(D)$ is empty or entirely singular; the projection operator fails to define a consistent tangential flow, producing the dynamical signature of case (ii). No third case exists, since any intermediate behaviour would require smooth strata incompatible with the global section obstruction. □

Corollary 11.4 (Wisdom–Error Separation)**.** *Wisdom and error are separated by a topological invariant. A system is capable of sustaining a stable notion of wisdom if and only if* $[\sigma] = 0$*. When* $[\sigma] \neq 0$*, no refinement of the relaxation dynamics alone can produce stability; modification of the log is necessary.*

Remark 11.5. The dichotomy eliminates the possibility of systems that are "almost coherent" in any stable sense. Apparent intermediate cases correspond to long-lived transient behaviour near singular strata (Section 9), not to a third stable regime. The system either admits a globally coherent organisation or it does not; the topological invariant $[\sigma]$ decides which.

12. The Triple Diagnostic Signature

The proof of Theorem 11.1 shows that obstruction is registered not by a single scalar but by a *triple signature.*

1. Persistent action floor. In the unobstructed case, $\mathcal{A}_{\text{soft}}(\Phi_t) \to 0$ as $t \to \infty$.

In the obstructed case, $\inf_{\Phi \in \mathcal{X}} \mathcal{A}_{\text{soft}}(\Phi) > 0$: the action does not decay but oscillates above a floor set by the cohomological defect.

2. Tangent instability. At any candidate configuration near the defect, the restricted Hessian $\text{Hess}_\Sigma \mathcal{A}_{\text{soft}}(\Phi)$ is ill-conditioned: the projection operator cannot produce a consistent tangential direction because the manifold has a topological tear.

3. Non-convergence of stochastic trajectories. In the unobstructed case, the Fokker–Planck density on Σ converges to the Gibbs measure $\rho_\tau \propto \exp(-\mathcal{A}_{\text{soft}}/\tau)$. In the obstructed case, no such measure exists; trajectories wander without localising.

These signals are individually observable and jointly diagnostic. A system exhibiting all three is not merely incorrect but incoherent in the topological sense, and no local adjustment to its configuration can resolve the failure.

Corollary 12.1 (Locality of Repair)**.** *If $[\sigma]$ is supported on a single patch U_α, modification of $\{e_i : i \in U_\alpha\}$ is sufficient to restore admissibility. If the support is global, the log requires structural revision.*

13. Classification of Obstruction Types

Local obstructions. A 1-cocycle σ is local if it is cohomologous to a cocycle supported on a single overlap $U_\alpha \cap U_\beta$. It arises from a pair of events whose combined constraints are mutually incompatible and can be resolved by replacing one event or refining the cover. This is the computational analogue of a merge conflict.

Global obstructions. The cohomology class is non-trivial but not representable by a cocycle on any single overlap. The log is globally inconsistent even though every pair of events is pairwise compatible. Resolution requires adding new events that thread the inconsistency, or revising the cover structure.

Persistent obstructions. The class is non-trivial in $\check{H}^k(\mathcal{U}; \mathcal{F}_D)$ for all $k \geq 1$: no refinement of the cover can make the log globally consistent. Such logs must be partially retracted.

13.1. A Hierarchy of Repair Operations

The classification yields a hierarchy of repair operations ordered by increasing cost. *Event substitution* replaces a single offending event and resolves local obstructions. *Cover refinement* splits a patch to separate conflicting constraints, potentially trivialising a local cocycle. *Log extension* adds new events whose constraints thread a global inconsistency. *Log retraction* removes a connected family of events, eliminating a persistent obstruction at the cost of erasing part of the history—under Log Sovereignty, the last resort.

14. Viscoelastic Logic

The tangent space decomposes as

$$T_\Phi \mathcal{X} = N_\Phi \mathcal{M}(D) \oplus T_\Phi \mathcal{M}(D).$$

The normal component $N_\Phi \mathcal{M}(D)$ is elastically suppressed, with stiffness $\lambda \gg 1$. The tangential component $T_\Phi \mathcal{M}(D)$ is viscously damped by the stochastic relaxation, with viscosity coefficient τ.

In the unobstructed case, the system behaves like a viscoelastic medium: stress introduced by new events redistributes and dissipates. The elastic component snaps the system to the constraint manifold; the viscous component drifts it toward its most stable configuration.

In the obstructed case, the medium has a topological tear. Stress cannot be distributed because the manifold does not support a continuous deformation from the current configuration to any stable one. The elastic component deposits the system near the defect; the viscous component cannot carry it away. This is *viscoelastic failure*: the material tears rather than deforming.

Viscoelastic logic is therefore not a metaphor but the exact dynamical regime characterised by Theorem 11.1. The triple signature of obstruction—action floor, tangent instability, trajectory non-convergence—is the mechanical signature of a torn medium.

15. Connection to Machine Learning

Double descent. The double descent curve corresponds to the transition through the three geometric regimes. The constrained regime is the standard bias–variance setting [2]. At the interpolation threshold, $\mathcal{M}(D)$ is near-discrete and highly curved, producing the variance peak. In the excess regime, $\mathcal{M}(D)$ is a smooth high-dimensional manifold, and relaxation toward flat regions produces the second descent [3].

Flat minima and implicit regularisation. Stochastic gradient descent with small batch sizes finds flatter minima because flat regions have larger volume and greater measure under the Gibbs distribution; the stochastic dynamics naturally select them [4, 5, 18].

Benign overfitting. A model that achieves zero training loss while generalising well navigates within $\mathcal{M}(D)$ in a low-curvature region [1]. Small perturbations to the input do not move the system out of the flat region.

Hallucination. A hallucination is a configuration that appears admissible under the fast Normal Snap but resides near a cohomological defect. The system generates locally consistent outputs that cannot be glued into a globally coherent section [22]. The diagnostic is not the content of the output but the geometry of the neighbourhood: high restricted Hessian, persistent action floor, stochastic non-convergence.

In each case, the machine learning phenomenon is a special instance of intramanifold dynamics under stochastic projected relaxation. The framework provides a geometric language in which existing empirical results become theorems.

16. Conclusion

The unification of Spherepop, RSVP, and TARTAN developed here is a single continuous programme: a redefinition of learning, computation, and state as geometric processes constrained by preserved history.

Central thesis. *Intelligence is not the reduction of information but the maintenance and coherent organisation of it. The capacity to do so depends not on limiting representation but on navigating the geometry of excess. Error is not a deviation from* *ground truth but a topological impossibility: the failure of a history to admit a globally coherent configuration.*

The thesis has four structural components. *Truth* is the existence of a global section of the constraint sheaf: constraints glued together without remainder. *Intelligence* is the capacity to navigate within that section under the projected gradient flow. *Wisdom* is the occupancy of low-curvature strata—configurations stable under tangential perturbation. *Error* is the non-triviality of the Čech class: the log's constraints cannot be made globally coherent by any choice of configuration. The appropriate response to error is not correction—adjustment of a scalar—but *repair*: identification and resolution of the cohomological defect that prevents the log from sustaining a coherent world-state.

The title is therefore not rhetorical. Error is obstruction: the precise mathematical object that prevents the construction of a globally coherent global section from a locally consistent event log. And the geometry of that obstruction—its locality or globality, its persistence under cover refinement—is what determines whether the log can be repaired, and how.

Acknowledgements. This work is developed within the ongoing programme on RSVP, Spherepop, and TARTAN.

A. Phenomenological Motivation: Learning Without a Ceiling

This appendix offers a more concrete account of the intuitions behind the paper for readers whose primary interest is in learning practice rather than in the formal machinery. It is not part of the proof structure. It is the record that gave rise to the question.

The Capacity That Grows

The standard prescription for a struggling learner is reduction. Take fewer courses. Narrow the focus. Let the overloaded buffer drain. This prescription rests on a model: the mind as a container of fixed size, filling up as experience accumulates, degrading when overfull. Under that model, relief requires subtraction.

The model is incomplete. A student who routinely takes more courses than required—and finds that performance improves rather than decays—is not a counterexample to some statistical regularity. They are evidence that a different regime is possible: one in which the constraint set is large enough that additional experience does not compete with existing structure but *extends* it. The brain, under those conditions, is not a container. It is a geometry. Courses do not fill it; they refine its coordinate system.

The same effect appears in skilled manual work. A builder who knows electrical, plumbing, structural framing, finishing, and site management does not merely know more facts than a single-trade specialist. The cross-domain knowledge creates a dense field of relevance in which every object is already situated. A piece of scrap wire is not just metal; it is indexed by gauge, application, code implications, reuse potential, and whether it belongs to a live circuit. The narrow specialist, lacking that surrounding structure, must either ask or proceed blindly. They do not even know what they do not know, and so they make errors that seem absurd to the practitioner who can see the whole landscape.

The student who switches subjects when bored is doing something structurally similar. Each domain provides a coordinate system for pattern recognition. When one domain is locally exhausted—when fatigue or repetition has suppressed gradient—

another domain provides fresh structure. The subject-switching is not distraction; it is manifold traversal. The learner is not disorganised. They are maintaining a coupled system of representations, each of which stabilises the others.

Naming as Navigation

A concrete version of the same principle appears in how one assigns names to projects in a large and active shared namespace. Someone who has observed thousands of repositories does not remember them individually. But they have internalised the distribution: which names are saturated (every variation of `core`, `browser`, `parser`, `base` is taken or nearly so), which are empty, which are ambiguous, and which carry connotations that will mislead. The act of naming a project becomes navigational rather than arbitrary. When a collision occurs—when a new project appears using the same name—the experienced practitioner moves their own project rather than insisting on territory. The name is not a stake; it is a coordinate. What matters is clarity in the shared space.

This is what it looks like to operate within a richly populated field. The information retained is not a list of entries but a topology. The system holds a model of the landscape, and each new observation refines that model rather than merely adding to a count.

Cross-Domain Structure and the Invariants It Produces

What repeated traversal of many domains produces, over time, is not merely familiarity with many things. It produces *invariants*—patterns that recur across contexts, recognisable because they have been encountered in enough variations that their structure is visible through the surface differences. This is the mechanism behind what practitioners sometimes call intuition or judgment: not a mysterious extra-cognitive faculty, but the accumulated result of having encountered enough structured variation that the relevant invariants are quickly activated by new instances.

The formal account in the body of this paper treats learning as motion on an admissible manifold toward regions of low curvature. The phenomenological account says the same thing in experiential language: the practitioner with extensive cross-domain exposure is not searching for an answer from scratch. They are already near a flat region of the space, because their prior experience has already organised the surrounding geometry. New inputs do not require heavy computation to integrate; they slot into an already-structured landscape.

The Limit Case: When Accumulation Outpaces Integration

The one genuine risk in this regime is not accumulation per se but the possibility that the rate of accumulation outpaces the rate of integration. If inputs arrive faster than the system can form connections among them, they accumulate as isolated tokens rather than as nodes in a connected field. That is the only scenario in which the bottleneck model's prescription becomes relevant: when the system is, in fact, operating in a disconnected way.

The diagnostic is not the count of inputs but the connectivity of the representations they generate. A system in which many inputs are forming connections—each new project illuminating previously opaque aspects of earlier ones—is not overloaded. A system in which inputs accumulate without forming connections is. The intervention in the second case is not necessarily subtraction; it may be restructuring: slowing down long enough to allow connections to form, or deliberately rotating between domains to expose latent commonalities that were not visible under a single lens.

The paper's formal language for this is the constraint sheaf and its sections. Locally consistent constraints that cannot be glued globally are the formal image of knowledge that has not integrated. Persistent action floor, tangent instability, and trajectory non-convergence are the signatures of a system that is not yet coherent—not overloaded, but not yet organised. The remedy, in each case, is not less experience but more structure in how experience is held.

References

[1] M. Dubova and S. J. Sloman, "Excess Capacity Learning," *Behavioral and Brain Sciences*, 2026, pp. 1–77. doi:10.1017/S0140525X2610510X.

[2] M. Belkin, D. Hsu, S. Ma, and S. Mandal, "Reconciling modern machine-learning practice and the classical bias–variance trade-off," *Proceedings of the National Academy of Sciences*, vol. 116, no. 32, pp. 15849–15854, 2019.

[3] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," *International Conference on Learning Representations*, 2017.

[4] S. Hochreiter and J. Schmidhuber, "Flat minima," *Neural Computation*, vol. 9, no. 1, pp. 1–42, 1997.

[5] N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, "On large-batch training for deep learning: Generalization gap and sharp minima," *International Conference on Learning Representations*, 2017.

[6] S. Mac Lane, *Categories for the Working Mathematician*, Springer, 2nd edition, 1998.

[7] J. C. Baez and M. Stay, "Physics, topology, logic and computation: A Rosetta Stone," in *New Structures for Physics*, Springer, 2011, pp. 95–172.

[8] J. Lurie, *Higher Topos Theory*, Princeton University Press, 2009.

[9] R. Bott and L. W. Tu, *Differential Forms in Algebraic Topology*, Springer, 1982.

[10] S. I. Gelfand and Y. I. Manin, *Methods of Homological Algebra*, Springer, 1996.

[11] D. I. Spivak, *Category Theory for the Sciences*, MIT Press, 2014.

[12] T. Jacobson, "Thermodynamics of spacetime: The Einstein equation of state," *Physical Review Letters*, vol. 75, no. 7, pp. 1260–1263, 1995.

[13] E. Verlinde, "On the origin of gravity and the laws of Newton," *Journal of High Energy Physics*, vol. 2011, no. 4, 2011.

[14] R. Landauer, "Irreversibility and heat generation in the computing process,"

IBM Journal of Research and Development, vol. 5, no. 3, pp. 183–191, 1961.

[15] C. H. Bennett, "The thermodynamics of computation—a review," *International Journal of Theoretical Physics*, vol. 21, pp. 905–940, 1982.

[16] S. Lloyd, "Ultimate physical limits to computation," *Nature*, vol. 406, pp. 1047–1054, 2000.

[17] K. Friston, "The free-energy principle: A unified brain theory?," *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[18] S. Amari, "Natural gradient works efficiently in learning," *Neural Computation*, vol. 10, no. 2, pp. 251–276, 1998.

[19] C. Villani, *Optimal Transport: Old and New*, Springer, 2009.

[20] F. Otto, "The geometry of dissipative evolution equations: The porous medium equation," *Communications in Partial Differential Equations*, vol. 26, pp. 101–174, 2001.

[21] M. Hairer, "Introduction to stochastic partial differential equations," *arXiv preprint arXiv:0907.4178*, 2009.

[22] S. Abramsky and A. Brandenburger, "The sheaf-theoretic structure of non-locality and contextuality," *New Journal of Physics*, vol. 13, 2011.

[23] R. Ghrist, *Elementary Applied Topology*, Createspace, 2014.