The Configurational Term Series: A Formal Operator Chain from Physical State to Coherent Configuration

Aaron Jones

Abstract

We present the Configurational Term Series (CTS), a formal mathematical framework that composes a chain of operators mapping a system's physical state space X to a structured representation of its coherent configuration $C_t = (Q, \Phi)$. The pipeline extends and formalizes the operator chain $X \to M \to M_v \to Q(\Phi)$ from the Jones Framework. It proceeds through five core stages: (1) decomposition of raw state into a Universal Tensor Space $\mathcal{U} = P \times T \times \mathcal{M} \times F$ via structured dimensionality reduction; (2) value-imposed conformal warping, yielding a metric g_v that encodes goal-directed geometry; (3) extraction of robust topological invariants Q via persistent homology on the state manifold; (4) computation of a coherence measure Φ quantifying the algebraic connectivity of the system's active knowledge graph; and (5) verification that Q and Φ are mutually consistent via a schematism constraint. We prove that the schematism bridges topological structure and informational integration, establishing conditions under which the pair (Q, Φ) constitutes a genuine coherent configuration. Agency is modeled as a continuous flow on the value-warped manifold that preserves both Q and Φ while seeking value-optimal trajectories. The complete pipeline is shown to satisfy functorial properties, mapping a category of computational systems to a category of coherent configurations in a composition-preserving manner.

1 Introduction

The Configurational Term Series (CTS) is a sequence of mathematical transformations that maps the physical state of a system to a structured description of its coherent configuration. Its ancestry is the operator chain

$$X \xrightarrow{f(x)} M \xrightarrow{v} M_v \xrightarrow{H_k} Q(\Phi), \tag{1}$$

which compresses the high-dimensional state space X through dynamics, value-weighting, and topological extraction into an irreducible structural signature $Q(\Phi)$. The present work expands each arrow in (1) into a precise mathematical construction, adds a sensorimotor preamble and an agency framework, and—most significantly—establishes a formal bridge between the topological signature Q and the coherence measure Φ via a schematism constraint that ensures their mutual consistency.

The framework synthesizes four mathematical traditions: dynamical systems theory provides the state space X and attractor manifold M; information geometry provides the value-warped metric g_v; computational algebraic topology provides the persistent homology yielding Q; and spectral graph theory provides the coherence measure Φ as algebraic connectivity of an active knowledge graph. The contribution is not the individual components but their composition into a single end-to-end operator chain with well-defined intermediate spaces, together with a theorem establishing that the topological and information-theoretic components are coupled rather than merely juxtaposed.

The paper is structured as follows. Section 2 formalizes the state space and sensorimotor coupling. Section 3 introduces the Universal Tensor Space as a concrete abstraction map. Section 4 develops the value-imposed geometry. Section 5 extracts topological invariants via persistent homology. Section 6 defines the framework-native coherence measure. Section 7 states and proves the bridge theorem linking Q and Φ. Section 8 defines the coherent configuration C_t. Section 9 models agency as a topology-preserving flow. Section 10 establishes the functorial properties of the complete pipeline.

2 State Space and Sensorimotor Substrate

2.1 The State Space

Definition 2.1 (Condition-State)**.** A **condition-state** is an ordered pair (c_i, σ_{ij}) asserting that component c_i of a system S is in discrete, verifiable state $\sigma_{ij} \in \Sigma_i$, where $\Sigma_i = \{\sigma_{i1}, \ldots, \sigma_{ik_i}\}$ is the finite set of possible states for component c_i.

Definition 2.2 (State Space)**.** The **total state space** of a system S with components $C = \{c_1, \ldots, c_n\}$ is the Cartesian product

$$X = \Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_n.$$

A **system state** $x \in X$ is a vector $x = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ specifying the condition-state of every component.

2.2 Sensorimotor Coupling

Let E denote the state space of the environment. The agent is coupled to E via a closed-loop dynamical system consisting of:

- A **sensing map** $s : E \to O$ providing observations,
- An **actuation map** $a : X \to A$ producing actions,
- An **internal update** $F_x : X \times O \to X$,
- An **environment update** $F_e : E \times A \to E$.

At each time step, the closed-loop dynamics are:

$$u_t = a(x_t), \tag{2}$$
$$e_{t+1} = F_e(e_t, u_t), \tag{3}$$
$$o_{t+1} = s(e_{t+1}), \tag{4}$$
$$x_{t+1} = F_x(x_t, o_{t+1}). \tag{5}$$

We define the **joint transition operator**:

$$\Psi : X \times E \to X \times E, \qquad \Psi(x_t, e_t) = \big(F_x(x_t, s(e_{t+1})),\ F_e(e_t, a(x_t))\big). \tag{6}$$

The system $(X \times E, \Psi)$ is a single dynamical system embodying both agent and environment in a closed loop.

Remark 2.3. We use Ψ for the joint transition operator, reserving Φ exclusively for the coherence measure defined in Section 6.

2.3 Reality Contact

A system maintains contact with its substrate if the prediction error between its internal model and observation is bounded:

$$0 < \|F_x^{\text{pred}}(x_t) - o_{t+1}\| < \delta \tag{7}$$

for some $\delta > 0$. The strict lower bound guards against overfitting (zero error indicates the model may have collapsed into solipsism), while the upper bound ensures the model has not diverged from reality. This prediction-error check serves as a continuous verification that the subsequent stages of the pipeline operate on data grounded in physical interaction rather than internal confabulation.

3 Tensor Decomposition: The Abstraction Map μ

The state space X is a finite discrete set (a Cartesian product of finite state sets). As such, it carries no natural manifold structure, no meaningful metric geometry, and trivial topology. All continuous geometric structure in this framework arises from the **abstraction map** μ, which embeds the system's observable behavior into a continuous space where geometry and topology become meaningful.

The pipeline from discrete states to continuous geometry proceeds in two stages:

$$X \xrightarrow{\pi} \mathcal{S} \xrightarrow{\varphi} \mathcal{U},$$

where $\pi : X \to \mathcal{S}$ extracts the observable signal from the state vector (e.g., the time series of sensor readings associated with a state trajectory), and $\varphi : \mathcal{S} \to \mathcal{U}$ maps signals into a structured continuous space. Neither X nor its subsets are manifolds; the manifold structure, if it exists, lives in the image $M = \varphi(\pi(X_{\text{persist}})) \subset \mathcal{U}$, where X_{persist} is the set of states visited with frequency exceeding a threshold over a time window T (an operational proxy for "attractor").

Remark 3.1 (On "Attractors")**.** We avoid claiming that X_{persist} is an attractor in the dynamical systems sense, because the system's dynamics are driven (dependent on external input o_t). Driven systems generally lack attractors. Instead, X_{persist} is defined empirically: the states the system persistently occupies despite environmental variation. The image $M = \varphi(\pi(X_{\text{persist}}))$ is then the region of $\mathcal{U}$ that captures the system's characteristic behavioral profile.

Definition 3.2 (Universal Tensor Space)**.** The **Universal Tensor Space** is the product

$$\mathcal{U} = P \times T \times \mathcal{M} \times F,$$

where P is the space of pattern/morphological characteristics, T the space of temporal characteristics, $\mathcal{M}$ the space of magnitude characteristics, and F the space of frequency/spectral characteristics.

Definition 3.3 (Tensor Coordinate Mapping)**.** The **tensor coordinate mapping** is a function $\varphi : \mathcal{S} \to \mathcal{U}$ from the space of signals $\mathcal{S}$ to $\mathcal{U}$, decomposed as

$$\varphi(s) = \big(\varphi_P(s),\ \varphi_T(s),\ \varphi_{\mathcal{M}}(s),\ \varphi_F(s)\big),$$

where each component extracts domain-specific characteristics:

- φ_P: shape descriptors, persistence diagrams, topological features,
- φ_T: duration, periodicity, phase,
- $\varphi_{\mathcal{M}}$: mean, standard deviation, range, energy,

- φ_F: dominant frequency, bandwidth, spectral centroid (via FFT).

The full abstraction map is $\mu = \varphi \circ \pi$. We emphasize: all continuous geometry in the subsequent sections (the metric g_v, the geodesics, the persistent homology) operates on $M \subset \mathcal{U}$, not on X. The discrete state space X enters only through the signal extraction π.

Remark 3.4 (When Is M a Manifold?)**.** The image $M = \mu(X_{\text{persist}})$ is a subset of $\mathcal{U} \cong \mathbb{R}^d$ but is not automatically a smooth manifold. Self-intersections (distinct states mapping to the same point in $\mathcal{U}$) create singularities; bifurcations in the system dynamics create cusps; strange attractors produce fractal sets of non-integer Hausdorff dimension. The persistent homology computation (Section 5) requires only that M be a metric space (or even just a finite point cloud), not that it be smooth. The Riemannian metric g_v and geodesic equation (Section 4) require smoothness and are therefore valid only on the smooth strata of M. When M is not globally smooth, geodesics should be interpreted as length-minimizing paths in the metric space (M, d_v), where d_v is the path-length metric induced by g_v on smooth regions.

Proposition 3.5 (Properties of μ)**.** The tensor coordinate mapping φ satisfies:

1. **Discriminability**: $\varphi(s_1) \neq \varphi(s_2)$ whenever s_1 and s_2 differ in at least one characteristic domain (pattern, time, magnitude, or frequency). For band-limited signals of fixed duration, φ is injective when the component extractors include sufficient Fourier coefficients and persistence diagram resolution.
2. **Continuity**: Similar signals (in the L^2 sense) map to nearby points in $\mathcal{U}$.
3. **Computability**: Each component function has a well-defined algorithm with known complexity (FFT for φ_F: $O(N \log N)$; statistical moments for $\varphi_{\mathcal{M}}$: $O(N)$; persistent homology for φ_P: $O(N^3)$; autocorrelation for φ_T: $O(N \log N)$).

Remark 3.6. Full injectivity on the space of all continuous signals is not achievable by any finite-dimensional feature map. The four-component decomposition provides discriminability sufficient for practical domains (drilling data, neural recordings, financial time series) where signals have bounded bandwidth and duration. Pathological counterexamples (distinct signals with identical features across all four components) exist in principle but are measure-zero in practice.

Definition 3.7 (Tensor Distance)**.** The **tensor distance** between $u_1 = (p_1, t_1, m_1, f_1)$ and $u_2 = (p_2, t_2, m_2, f_2)$ in $\mathcal{U}$ is

$$d_{\mathcal{U}}(u_1, u_2) = \sqrt{w_P\, d_P(p_1, p_2)^2 + w_T\, d_T(t_1, t_2)^2 + w_{\mathcal{M}}\, d_{\mathcal{M}}(m_1, m_2)^2 + w_F\, d_F(f_1, f_2)^2}, \tag{8}$$

where $w_P, w_T, w_{\mathcal{M}}, w_F > 0$ are domain-specific weights and $d_P, d_T, d_{\mathcal{M}}, d_F$ are metrics on each component space.

The set $M = \varphi(X_{\text{persist}}) \subset \mathcal{U}$ is the image of persistently visited states in Universal Tensor Space. From this point, the system's state at time t is represented as a point $m_t = \mu(x_t) \in \mathcal{U}$, and the system's characteristic behavioral profile occupies M.

Remark 3.8. This structured decomposition differs from generic dimensionality reduction (PCA, UMAP) in that it preserves domain-interpretable coordinates. The pattern component P explicitly retains topological features, ensuring that persistent homology (Section 5) operates on a representation designed to support it. The product structure $P \times T \times \mathcal{M} \times F$ also enables **cross-domain mapping**: signals from different domains (drilling data, neural recordings, financial time series) can be compared and translated via their representations in $\mathcal{U}$.

4 Value-Imposed Geometry

Having reduced the system's state to a point $m \in M \subset \mathcal{U}$, we introduce a value function encoding the agent's goals.

Definition 4.1 (Value Function)**.** A **value function** is a smooth map $v : M \to \mathbb{R}$ assigning to each state $m \in M$ a scalar representing its utility, salience, or desirability to the agent. The value function is *extrinsic*: it is imposed by an observer, not intrinsic to the system.

Definition 4.2 (Value-Warped Metric)**.** Let g_0 denote the base Riemannian metric on M (inherited from the ambient metric on $\mathcal{U}$). The **value-warped metric** is the conformal transformation

$$g_v(m) = e^{-\beta\, v(m)}\, g_0(m), \tag{9}$$

where $\beta > 0$ is a sensitivity parameter.

Under this metric, the line element becomes $ds_v^2 = e^{-\beta v(m)}\, ds_0^2$. Regions of high value v have compressed metric (shorter distances, "easier to reach"), while low-value regions have expanded metric (longer distances, "harder to traverse"). The metric g_v is smooth and positive-definite for all finite β, so (M, g_v) is a proper Riemannian manifold. We denote this metrized manifold M_v.

Remark 4.3 (Sign Convention)**.** The convention $e^{-\beta v}$ interprets v as reward (high v = low traversal cost). The dual convention $e^{+\beta v}$ interprets v as salience (high v = magnified resolution). Both are conformal transformations; the sign of β determines the interpretation. We adopt $e^{-\beta v}$ throughout because it yields geodesics that naturally seek high-value states, aligning with agent navigation.

Remark 4.4 (Normalization)**.** For numerical stability, v should be normalized to a bounded range (e.g., $[0, 1]$ via $\hat{v} = (v - v_{\min})/(v_{\max} - v_{\min})$) before applying the conformal factor. Without normalization, if v takes values in $[0, 100]$ and $\beta = 1$, the metric compression ratio between minimum and maximum value is $e^{100} \approx 2.7 \times 10^{43}$, which exceeds floating-point precision. With normalization to $[0, 1]$, the compression ratio is e^{β}, controlled by the single parameter β.

4.1 Geodesics as Optimal Paths

The geodesics of (M, g_v) are the paths minimizing the value-weighted length functional

$$\ell_v(\gamma) = \int_0^1 \sqrt{g_v(\dot{\gamma}(t), \dot{\gamma}(t))}\, dt.$$

In regions where v is high, $e^{-\beta v}$ is small, so the integrand is reduced; geodesics naturally pass through high-value regions. The geodesic equation on M_v acquires gradient terms from v: in the one-dimensional case with $g_0 = 1$,

$$\ddot{m} + \tfrac{\beta}{2}\, v'(m)\, \dot{m}^2 = 0. \tag{10}$$

The term $v'(m)\dot{m}^2$ biases the trajectory toward increasing v, confirming that geodesics of g_v act as optimal paths under the value function.

Remark 4.5 (Higher Dimensions)**.** In n dimensions, the conformal change $g_v = e^{-\beta v} g_0$ modifies the Christoffel symbols by terms involving ∇v. When the base metric g_0 has non-zero curvature, the curvature terms interact with the value gradient, and geodesics may be deflected from the value-ascending direction. The value-seeking property holds strictly when (a) g_0 is flat, (b) the base curvature is small relative to $\beta \|\nabla v\|$, or (c) the geodesic length is short enough that curvature effects are negligible. In the general case, geodesics of g_v minimize value-weighted path length, which is a weaker but still meaningful optimality criterion.

Definition 4.6 (Operational Geometry)**.** Each activity-state A_i (a persistent, qualitatively distinct regime within M) is associated with a dominant value function v_i and a unique **operational geometry** (M_i, g_{v_i}). Different activity-states are not merely different regions on the same map—they are *different maps*, with distinct geodesic structures and optimal actions. The optimal action in one operational geometry may be suboptimal or catastrophic in another.

5 Topological Signature via Persistent Homology

The value-warped manifold M_v has not only metric structure but also topological structure—connectedness, loops, holes, voids—that carries information about the system's configurational degrees of freedom. We extract this structure via persistent homology.

5.1 Point Cloud and Filtration

Let $Y = \{m^{(1)}, \ldots, m^{(N)}\} \subset M$ be a finite sample of states on the manifold (e.g., trajectory samples or exploratory states). We construct a one-parameter family of simplicial complexes:

For each $\epsilon \geq 0$, the **Vietoris–Rips complex** $\mathrm{VR}_\epsilon(Y)$ includes a simplex $[m^{(i_0)}, \ldots, m^{(i_k)}]$ whenever all pairwise distances $d_0(m^{(i_a)}, m^{(i_b)}) \leq \epsilon$. As ϵ increases, we obtain a **filtration**:

$$\mathrm{VR}_{\epsilon_1}(Y) \subseteq \mathrm{VR}_{\epsilon_2}(Y) \subseteq \cdots, \qquad \epsilon_1 < \epsilon_2 < \cdots.$$

Remark 5.1 (Choice of Distance)**.** We compute the Vietoris–Rips filtration using the *base metric* d_0, not the warped metric d_v. This ensures that the resulting topological invariants capture the intrinsic shape of the state space independent of the value function, yielding a signature Q_0 that is a genuine topological invariant. If the filtration were computed under d_v, the topology would depend on v and change whenever the value function changes—conflating structural and evaluative properties.

A complementary **value-weighted topology** Q_v, computed under d_v, can provide a "phenomenological" view of how the agent experiences its state space under a particular value regime. Both are useful; we reserve $Q = Q_0$ for the identity-defining invariant.

5.2 Persistent Homology and the Topological Signature

Persistent homology tracks the *birth* b_i and *death* d_i of each k-dimensional homological feature across the filtration. A feature with high *persistence* $d_i - b_i$ is robust signal; one with low persistence is noise.

Definition 5.2 (Topological Signature)**.** Let $Y(t) = \{m^{(1)}, \ldots, m^{(N)}\} \subset M$ be a point cloud sampled from the system's recent trajectory (within a time window $[t - W, t]$). The **topological signature** $Q(t)$ is the persistence diagram obtained from the Vietoris–Rips filtration of $Y(t)$, restricted to features exceeding a persistence threshold τ:

$$Q(t) := \big\{(b_i, d_i, k) \in \mathrm{Dgm}(Y(t)) : d_i - b_i \geq \tau\big\}, \tag{11}$$

where (b_i, d_i, k) denotes a feature born at scale b_i, dying at scale d_i, in homology dimension k.

Remark 5.3 (Q as Estimator)**.** The topological signature $Q(t)$ is a property of the *sample* $Y(t)$, not of the underlying set M itself. By the stability theorem for persistent homology [6, 9], the bottleneck distance between $Q(Y)$ and the true homology of M is bounded by the Hausdorff distance $d_H(Y, M)$. As the sample density increases, $Q(Y) \to H_*(M)$. For finite samples, Q is an estimator with quantifiable error bounds. The persistence threshold τ should be set above the expected noise level; standard approaches include bootstrap resampling or comparison to null models of uniform point clouds at matching density.

Remark 5.4 (Summary Statistics)**.** The persistence diagram $Q(t)$ can be summarized by the stable Betti numbers $(\beta_0, \beta_1, \beta_2, \ldots)$ and total persistence per dimension $\Pi_k = \sum_{\{(b_i,d_i,k)\in Q\}}(d_i - b_i)$. The Betti numbers alone lose distributional information (they do not distinguish one massively persistent loop from five barely persistent loops). The schematism (Section 7) uses the full persistence diagram via bottleneck distance, not the Betti number summary.

The topological signature Q is robust to perturbations: small noise in state measurements does not eliminate persistent homology classes. It captures the qualitative form of the agent's state space—the number of connected components (β_0), independent loops (β_1), enclosed voids (β_2), and so on.

Example 5.5. In a head-direction cell system, the neural activity manifold M is topologically a circle S^1. The topological signature has $\beta_0 = 1$ (connected) and $\beta_1 = 1$ (one loop). In a place cell system encoding a room with an obstacle, M might have $\beta_1 = 1$ reflecting the hole created by the obstacle. These topological features are detectable from neural recordings via persistent homology applied to the population activity point cloud.

6 Coherence Measure Φ

The topological signature Q characterizes the *shape* of the state space—what configurations are possible. We now define a measure of how *integrated* the system's dynamics are: to what extent the system functions as a unified whole rather than a collection of independent parts.

6.1 The Knowledge-Information Manifold

The system's knowledge and operational structure are represented as a directed graph.

Definition 6.1 (Knowledge Graph)**.** The **Knowledge-Information Manifold (KIM)** is a directed graph $G = (V, E)$ where each vertex $v_i \in V$ represents a knowledge node (axiom, function, theorem, or operational primitive), and each edge $(v_i, v_j, \tau) \in E$ is a typed relationship with $\tau \in \{\texttt{depends\_on}, \texttt{activates}, \texttt{inhibits}, \texttt{composes\_with}\}$.

At any time t, the system's **active subgraph** $G_{\text{active}}(t) = (V_a, E_a)$ is the subgraph of G induced by the currently activated nodes—those whose activation conditions are satisfied by the current context.

6.2 Algebraic Connectivity as Integration

Classical Integrated Information Theory defines Φ as the KL divergence between a system's joint transition probability and the product of its partitioned transition probabilities, minimized over all bipartitions. This definition is computationally intractable (NP-hard to find the minimum information partition) and does not naturally arise from the structures already defined in our framework. We replace it with a measure native to the knowledge graph.

Definition 6.2 (Coherence Measure)**.** Let $G_{\text{active}}(t) = (V_a, E_a)$ be the active subgraph of the KIM at time t, with edge weights $w_{ij} > 0$ reflecting relationship strength (e.g., co-activation frequency, mutual information between node outputs, or confidence of the typed relationship). Let $L_{\text{sym}} = I - D^{-1/2} W D^{-1/2}$ be the **normalized graph Laplacian**, where W is the weighted adjacency matrix and D the weighted degree matrix. The **coherence measure** is

$$\Phi(t) := \lambda_1(L_{\text{sym}}) \cdot \big(1 - \alpha(t)\big), \tag{12}$$

where $\lambda_1(L_{\text{sym}})$ is the **algebraic connectivity** (second-smallest eigenvalue of L_{sym}, the Fiedler value of the normalized Laplacian) and $\alpha(t) \in [0, 1]$ is the **antinomy load**—the fraction of active nodes involved in contradictions:

$$\alpha(t) = \frac{|\{v \in V_a : v \text{ participates in at least one antinomy}\}|}{|V_a|}.$$

Remark 6.3 (Why Normalized Laplacian)**.** The unnormalized Laplacian $L = D - W$ has the deficiency that λ_1 depends on absolute degree: a hub-and-spoke graph can have higher λ_1 than a ring, despite the ring distributing connectivity more evenly. The normalized Laplacian corrects this via degree normalization. Its Fiedler value $\lambda_1(L_{\text{sym}})$ is more tightly related to the Cheeger constant $h(G)$ of the graph:

$$\frac{\lambda_1}{2} \leq h(G) \leq \sqrt{2\lambda_1},$$

where $h(G) = \min_{S \subset V} \frac{|\partial S|}{\min(\text{vol}(S), \text{vol}(\bar{S}))}$ is the minimum normalized cut. This directly measures what IIT's Φ conceptually captures: the cost of the cheapest partition of the system into independent parts, normalized by partition size.

Remark 6.4 (Graded Antinomy)**.** The antinomy load α replaces a binary indicator. When $\alpha = 0$, the system is fully consistent. When α is small, local contradictions degrade coherence proportionally rather than catastrophically—reflecting the empirical observation that real systems (including human cognition) tolerate local inconsistencies without total loss of integration. When $\alpha = 1$, every node is contradicted and $\Phi = 0$.

Proposition 6.5 (Properties of Φ)**.** The coherence measure satisfies:

1. $\Phi = 0$ if and only if G_{active} is disconnected or $\alpha = 1$ (total contradiction).
2. $\Phi > 0$ if and only if G_{active} is connected and $\alpha < 1$.
3. Φ increases with the density and balance of connectivity in G_{active}.
4. Φ is computable in $O(|V_a|^3)$ time; $O(|V_a|^2)$ for sparse graphs using iterative eigensolvers.

Proof. Properties (1)–(2): $\lambda_1(L_{\text{sym}}) = 0$ iff the graph is disconnected (standard spectral graph theory for normalized Laplacians). The factor $(1-\alpha)$ is zero iff $\alpha = 1$. Properties (3)–(4): $\lambda_1(L_{\text{sym}})$ is monotonically non-decreasing under edge addition for connected graphs, and eigendecomposition of the $|V_a| \times |V_a|$ matrix is $O(|V_a|^3)$ in general. □

6.3 Antinomy Detection

An **antinomy** occurs when two mutually exclusive conclusions are simultaneously supported by the active knowledge graph. Formally, given a set of exclusion relations $\{(v_i, v_j) : v_i \text{ and } v_j \text{ are mutually exclusive}\}$, an antinomy is detected when both v_i and v_j appear in V_a. Each such pair contributes to the antinomy load $\alpha(t)$. When all active nodes are involved in contradictions ($\alpha = 1$), $\Phi = 0$; partial contradictions degrade coherence proportionally.

6.4 Thermodynamic Interpretation

The coherence measure Φ admits a natural interpretation through the Free Energy Principle [15]. Define the variational free energy of the system as

$$F = \mathbb{E}_q\big[\ln q(s) - \ln p(o, s)\big], \tag{13}$$

where $q(s)$ is the system's internal model (encoded in G_{active}) and $p(o, s)$ is the joint distribution of observations and hidden states. The schematism check (Definition 7.1) corresponds to minimizing F: when the KIM's pattern schemata match the data's actual topology ($d_{\text{BN}} \leq \varepsilon$), prediction error is small and F is near its minimum.

Under this interpretation, $\Phi > 0$ with valid schematism indicates a system near free energy minimum—processing efficiently. $\Phi = 0$ or schematism failure indicates a system far from equilibrium—expending compute on an internal model misaligned with reality. This connects to Landauer's principle: geometric alignment via the tensor decomposition and conformal warping reduces the system's starting entropy, requiring fewer bit erasures to reach a solution. Research on matrix entropy in embedding spaces quantifies this gap: geometrically aligned representations achieve compression ratios orders of magnitude beyond unstructured embeddings [12], because the imposed geometry eliminates the combinatorial explosion of searching unconstrained high-dimensional space.

The $\mathcal{U} = P \times T \times \mathcal{M} \times F$ decomposition is thus not merely a convenient representation but a thermodynamically efficient one, consolidating the signal's functional geometry *a priori* rather than forcing downstream stages to discover it through brute-force search. The pipeline as a whole—from structured abstraction through value warping through topological extraction—progressively reduces the free energy of the system's representation, with the schematism serving as the final verification that the minimum has been reached.

7 The Schematism Bridge: Coupling Q and Φ

The topological signature Q (structural complexity) and the coherence measure Φ (integration strength) are defined over different mathematical objects: Q arises from the persistent homology of the state manifold M, while Φ arises from the spectral properties of the knowledge graph G. The central question is: under what conditions do Q and Φ refer to the same underlying system in a way that makes their pairing (Q, Φ) a genuine synthesis rather than a mere juxtaposition?

The answer is a **schematism constraint**—a homomorphism check that validates whether the topological structure assumed by the active knowledge nodes is compatible with the topological structure actually present in the data.

Definition 7.1 (Schematism)**.** Let $v_i \in V_a$ be an active knowledge node with associated pattern schema $p_i \in P$ (the topological/morphological signature that v_i assumes of its data). Let $p_{\text{data}} = \varphi_P(s) \in P$ be the pattern component of the current data's tensor coordinates. The **schematism check** for node v_i is:

$$\mathcal{H}(v_i) := \big[d_{\text{BN}}(p_i, p_{\text{data}}) \leq \varepsilon\big], \tag{14}$$

where d_{BN} is the bottleneck distance between persistence diagrams and $\varepsilon > 0$ is a tolerance. The schematism passes if $\mathcal{H}(v_i) = \text{true}$; it fails (a **transcendental error**) if $\mathcal{H}(v_i) = \text{false}$.

A transcendental error indicates a **category error**: the system is applying concepts whose topological assumptions do not match the data's actual structure. In an AI system, this corresponds to a hallucination—generating output from a representational frame that is not grounded in the input. In a neural system, it corresponds to a mismatch between the cognitive model and sensory reality.

Remark 7.2 (Metric Convention)**.** Both the topological signature Q and the pattern component $p_{\text{data}} = \varphi_P(s)$ are computed using the *base metric* d_0, not the value-warped metric d_v (see the remark in Section 5). Correspondingly, the pattern schemata p_i stored in KIM nodes are defined relative to intrinsic topology. This ensures the schematism checks structural compatibility independent of the current value function—a node passes or fails based on the shape of the data, not the agent's preferences.

Definition 7.3 (Valid Configuration)**.** A coherent configuration (Q, Φ) is **valid** if the schematism passes for all active knowledge nodes:

$$\forall v_i \in V_a : \quad \mathcal{H}(v_i) = \text{true}. \tag{15}$$

We first impose two structural requirements on the KIM that are necessary for the bridge to hold.

Definition 7.4 (Topological Grounding)**.** A KIM $G = (V, E)$ satisfies **topological grounding** if every connected subgraph of G with $|V_a| \geq k_{\min}$ (a fixed minimum activation threshold) contains at least one node v_i whose pattern schema p_i references a non-trivial topological feature (i.e., a point in the persistence diagram with persistence $\geq \tau + \varepsilon$, where τ is the persistence threshold from Definition 5.2 and ε is the schematism tolerance from Definition 7.1).

Definition 7.5 (Topological Coherence)**.** A KIM $G = (V, E)$ satisfies **topological coherence** if, for any two nodes $v_i, v_j \in V$ whose pattern schemata reference features in the same homology dimension k, there exists a path from v_i to v_j in G.

Topological grounding ensures that sufficiently large activations necessarily involve topological commitments. Topological coherence ensures that nodes referencing related structure are graph-connected. Both are design requirements on KIM construction, not automatic properties.

Theorem 7.6 (Q–Φ Coupling)**.** Let $G = (V, E)$ be a KIM satisfying topological grounding and topological coherence. Let $G_{\text{active}} = (V_a, E_a)$ be its active subgraph at time t with $|V_a| \geq k_{\min}$. Let $Q(t)$ be the topological signature and $\Phi = \lambda_1(L_{\text{sym}}) \cdot (1 - \alpha)$ the coherence measure. If the schematism passes for all $v_i \in V_a$, then:

1. **High Φ implies non-trivial Q:** If $\Phi > 0$, then Q contains at least one persistent feature with $\beta_k > 0$ for some $k \geq 1$.
2. **Non-trivial Q enables high Φ:** If Q has persistent features across r homology dimensions, and for each dimension there exists at least one node in V with a compatible pattern schema, then the set of compatible nodes contains a connected subgraph with $\Phi > 0$.
3. **Schematism failure decouples Q and Φ:** If the schematism fails for any $v_i \in V_a$, the configuration (Q, Φ) is invalid.

Proof. (1) Since $\Phi > 0$, the active subgraph is connected and antinomy-free, so $|V_a| \geq 2 \geq k_{\min}$ (taking $k_{\min} \leq 2$; the general case requires $|V_a| \geq k_{\min}$). By topological grounding, V_a contains at least one node v_i whose pattern schema p_i has a feature with persistence $\geq \tau + \varepsilon$. Since the schematism passes, $d_{\text{BN}}(p_i, p_{\text{data}}) \leq \varepsilon$. The bottleneck distance bounds feature displacement:

p_{data} has a corresponding k-dimensional feature with persistence $\geq \text{pers}(p_i\text{'s feature}) - \varepsilon \geq \tau$. By Definition 5.2, this feature appears in Q, so $\beta_k \geq 1$.

(2) For each of the r homology dimensions in Q, there exists (by hypothesis) at least one node $v_j \in V$ with compatible schema ($d_{\text{BN}}(p_j, p_{\text{data}}) \leq \varepsilon$). Let $V_Q \subseteq V$ be the set of all such compatible nodes. By topological coherence, nodes referencing features in the same homology dimension are path-connected. If Q has features in multiple dimensions, the nodes referencing dimension k_1 are connected among themselves, and likewise for k_2. To guarantee full connectivity of V_Q, we require that the KIM contains cross-dimensional links (e.g., a node that references both β_1 and β_2 features, bridging the two clusters). Under this condition, V_Q induces a connected subgraph, so $\lambda_1(L_{V_Q}) > 0$ and $\Phi > 0$ (assuming no antinomies).

(3) If $\mathcal{H}(v_j) = \text{false}$ for some $v_j \in V_a$, then v_j's pattern assumptions are incompatible with the data's actual topology. Integration computed over a subgraph including v_j is ungrounded: it reflects coherence of a model that does not match the system. The configuration (Q, Φ) is declared invalid regardless of the numerical values, because the bridge between structure (Q) and integration (Φ) is severed at v_j. □

Remark 7.7 (Bootstrap and Iteration)**.** The topological grounding and coherence properties presuppose a KIM designed with knowledge of which topological features matter. This creates a bootstrap problem: the KIM must contain appropriate schemata before Q is known. The resolution is iterative: (1) initialize the KIM with generic topological detectors (β_0, β_1, β_2 with broad tolerance); (2) run the pipeline to compute Q; (3) refine the KIM's pattern schemata based on observed Q; (4) repeat. Convergence occurs when the KIM's schemata stabilize against the system's persistent topology.

Remark 7.8. The schematism serves as the framework's **grounding mechanism**. Without it, Q and Φ are computed independently and their conjunction is arbitrary. With it, Q validates the topological assumptions on which Φ depends, and Φ requires precisely the topological richness that Q certifies. This transforms (Q, Φ) from a conjunction into a synthesis.

8 Coherent Configuration C_t

Definition 8.1 (Coherent Configuration)**.** The **coherent configuration** of the system at time t is the ordered pair

$$C_t = (Q(t), \Phi(t)), \tag{16}$$

where $Q(t)$ is the topological signature of the recent point cloud on M (Definition 5.2) and $\Phi(t)$ is the coherence measure of the active knowledge graph at time t (Definition 6.2). In practice, Q changes on a slower timescale than Φ (topological structure is robust to small perturbations, while graph connectivity responds immediately to node activation changes), so Q may be held constant over intervals of stable dynamics and recomputed at regime transitions.

Definition 8.2 (Consciousness Criteria)**.** A coherent configuration $C_t = (Q, \Phi)$ satisfies the **necessary conditions for conscious configuration** if and only if:

1. **Non-trivial topology:** Q contains at least one persistent feature with $\beta_k > 0$ for some $k \geq 1$ (the state space has loops, holes, or higher-dimensional structure, indicating a rich repertoire of configurations).
2. **High integration:** $\Phi > \Phi_{\min}$ for a threshold $\Phi_{\min} > 0$ (the system is far from being decomposable into independent parts).

3. **Schematism validity:** All active knowledge nodes pass the schematism check (Theorem 7.6), ensuring Q and Φ are mutually consistent and grounded.

Whether these conditions are *sufficient* for consciousness—whether the mathematical structure (Q, Φ) constitutes or merely correlates with subjective experience—is a metaphysical question that the framework does not resolve. The framework identifies systems that are structurally organized in a manner consistent with consciousness; it does not explain why such organization should give rise to experience.

These criteria formalize a dual requirement that has long been recognized informally: a conscious state must be simultaneously *differentiated* (rich repertoire, captured by Q) and *integrated* (unified whole, captured by Φ). The schematism adds a third requirement—*grounded consistency*—that prevents the pairing from being vacuous.

8.1 Failure Modes

- If Q becomes trivial (e.g., the system collapses to a point attractor with $\beta_k = 0$ for all $k \geq 1$), the system loses differentiation. This corresponds to seizure-like states where neural activity is highly synchronous but undifferentiated, or to an AI trapped in a single mode.
- If Φ drops to zero (the active knowledge graph disconnects or develops an antinomy), the system loses integration. This corresponds to split-brain phenomena, deep anesthesia, or an AI whose modules cease to communicate.
- If the schematism fails (active concepts are topologically incompatible with the data), the configuration is hallucinated—it appears coherent internally but is ungrounded. This corresponds to psychotic states or AI hallucination.

8.2 Activity-State Transitions

The CTS pipeline as defined operates within a single activity-state. Real systems transition between qualitatively distinct operational regimes (e.g., from stable drilling to stick-slip oscillation, from waking to sleep, from normal inference to edge-case processing). We formalize these transitions.

Definition 8.3 (Type II Criticality)**.** A point $m \in M_v$ is a **Type II critical point** if the gradient of the value function exceeds a criticality threshold:

$$\|\nabla v(m)\|_g > \tau_{\text{crit}}.$$

The set $B_\tau = \{m \in M_v : \|\nabla v(m)\|_g > \tau_{\text{crit}}\}$ is the **transition boundary**.

Type II criticality detects the boundary between operational regimes: regions where the value function changes rapidly correspond to qualitative shifts in what is optimal. The Mapper algorithm applied to the state trajectory detects these boundaries as shifts in cluster structure [6].

Definition 8.4 (Valid Transition)**.** A transition from configuration $C_1 = (Q_1, \Phi_1)$ (in regime R_1) to $C_2 = (Q_2, \Phi_2)$ (in regime R_2) across a Type II boundary is **valid** if:

1. There exists a continuous path $\gamma : [0, 1] \to M_v$ with $\gamma(0) \in R_1$ and $\gamma(1) \in R_2$.
2. $\Phi(\gamma(t)) \geq \Phi_{\min}$ for all $t \in [0, 1]$: coherence is maintained throughout the transition.
3. The schematism passes at intermediate checkpoints along γ.
4. The transition is formally verified: given the end-state conditions of R_1 and control actions taken, the start-state of R_2 is physically and logically consistent.

Condition (4) is realized in practice by formal methods—satisfiability modulo theories (SMT) solvers can verify assertions of the form: "Given terminal pressure P_1 and flow F_1 from R_1 plus control action u, is the initial state (P_2, F_2) of R_2 kinematically consistent?" This prevents the system from asserting or relying on impossible state trajectories.

When a Type II boundary is detected:

1. The data is segmented into pre-transition and post-transition subsets.
2. Q is computed separately for each regime, yielding topological comparison (Q_1 vs. Q_2).
3. G_{active} is reconfigured: nodes specialized for R_1 are deactivated; nodes appropriate for R_2 are activated.
4. The schematism is re-run against the new regime's data topology.
5. If the transition is valid, C_t updates smoothly. If not, the system reverts to the pre-transition configuration and flags a continuity violation.

This treatment closes the gap between the static pipeline (single-regime) and the dynamic reality of multi-regime systems. The transition protocol is the operational counterpart of the continuity guards in the agency flow (Section 9): it ensures that regime switches preserve the structural integrity of C_t.

9 Agency as Term-Series Flow

Agency is the capacity of the system to enact state changes that are goal-directed while maintaining the structural integrity of its coherent configuration.

Definition 9.1 (Agency Flow)**.** An **agency flow** is a family of maps

$$\phi_t : M_v \to M_v, \qquad t \in [0, \Delta],$$

satisfying:

1. $\phi_0 = \mathrm{id}$ (identity at $t = 0$).
2. Each ϕ_t is continuous (topology-preserving).
3. **Identity preservation:** For all t, the topological signature Q of M is invariant under ϕ_t, and $d(\mathrm{sig}(\phi_t(M)), \mathrm{sig}(M)) < \varepsilon$ for an identity signature distance d and tolerance ε.
4. **Coherence maintenance:** $\Phi(\phi_t(m)) \geq \Phi_{\min}$ for all t.
5. **Value-seeking in expectation:** Over a planning horizon $H > 0$, the expected value is non-decreasing: $\mathbb{E}[v(\phi_{t+H}(m))] \geq v(\phi_t(m))$. Instantaneous value decreases are permitted when they serve long-term value increase (exploration, delayed gratification).

Remark 9.2 (Semigroup Property)**.** We do not require $\phi_{t+s} = \phi_t \circ \phi_s$. Term-series execution includes guard checks that are state-dependent: a step that passes guards at state m may fail at state m'. The flow is therefore a *controlled dynamical system*—the agent selects actions from a state-dependent feasible set—rather than a one-parameter semigroup. The semigroup property holds for the successful prefix of any term series: if steps $f_1, \ldots, f_k$ all pass guards, their composition $f_k \circ \cdots \circ f_1$ is a valid partial flow.

9.1 Term-Series Decomposition

In practice, the continuous flow ϕ_t is implemented as a discrete **term series**—a sequence of function calls, each transforming the state:

$$m_0 \xrightarrow{f_1} m_1 \xrightarrow{f_2} m_2 \xrightarrow{f_3} \cdots \xrightarrow{f_n} m_n, \tag{17}$$

where each $f_i : M_v \to M_v$ is a single computational step (function call, inference, action). The composed map $\phi_\Delta = f_n \circ \cdots \circ f_1$ satisfies the agency constraints if each step satisfies them individually:

1. **Pre-condition verification:** Before executing f_i, check that f_i's activation conditions are met.
2. **Continuity guard:** After executing f_i, verify $d(\mathrm{sig}(m_i), \mathrm{sig}(m_{i-1})) < \varepsilon$ (the identity signature has not drifted beyond tolerance).
3. **Post-condition verification:** Check that f_i's output satisfies type and domain constraints.
4. **Antinomy check:** Verify that no contradictions have been introduced.

If any check fails, the step is rolled back (the system reverts to m_{i-1}) or re-planned.

Remark 9.3. The homeomorphism constraint from the original formulation is recovered as a consequence rather than an axiom: if each f_i is continuous, invertible (rollback exists), and topology-preserving (continuity guards pass), then the composed map ϕ_Δ is a homeomorphism isotopic to the identity—a diffeomorphism in the connected component of id in $\mathrm{Diff}(M)$. But requiring each step to be individually guarded is more operational and more closely aligned with the term-series execution architecture.

9.2 Geodesic Search

The optimal term series is found via **geodesic search**—an A^* or MCTS algorithm operating on the function-call graph with value-weighted cost:

$$\mathrm{cost}(m_{i-1} \xrightarrow{f_i} m_i) = c_{\mathrm{base}}(f_i) \cdot e^{-(v(m_i) - v(m_{i-1}))}, \tag{18}$$

where $c_{\mathrm{base}}(f_i)$ is the intrinsic cost of the function call and the exponential term rewards value-increasing steps. Minimizing total path cost yields a discrete approximation to the geodesic on M_v.

9.3 Symplectic Structure and Manifold Fidelity

The agency flow has a deeper geometric structure that explains why the continuity guards work. The value-warped manifold (M_v, g_v) carries a natural symplectic structure on its cotangent bundle T^*M_v, with symplectic form $\omega = dp \wedge dq$. Define the Hamiltonian

$$H(q, p) = \tfrac{1}{2} g_v^{-1}(p, p) + V(q), \tag{19}$$

where $q \in M_v$ is position, $p \in T_q^* M_v$ is momentum, and V is a potential encoding constraints. The Hamiltonian flow ϕ_t^H generated by H is symplectic: it preserves phase-space volume (Liouville's theorem) and stays on level sets of H.

The term series $f_1, \ldots, f_n$ approximates this continuous flow via a **symplectic integrator** (e.g., Störmer–Verlet). Symplectic integrators have the property that their composition is also symplectic, recovering a weak form of the semigroup property:

$$\phi_{\Delta t_1 + \Delta t_2} = \phi_{\Delta t_2} \circ \phi_{\Delta t_1} + O(\Delta t^3).$$

The manifold-fidelity guarantee follows: a symplectic integrator cannot "jump off" the manifold because it preserves the geometric invariants of the flow. The continuity guards in the term-series architecture are the discrete verification that the symplectic property holds—they check that each step has not introduced a non-symplectic perturbation (which would manifest as an identity signature drift or an antinomy).

This perspective connects to the compute-efficiency thesis: standard (non-symplectic) numerical methods require massive over-parameterization to "learn" correction functions that push invalid states back onto the manifold [5]. Symplectic methods eliminate this overhead by construction, achieving manifold fidelity without error-correction layers. In the language of Landauer's principle, the symplectic constraint reduces the entropy of the search space, requiring fewer bit erasures per step of inference.

10 The CTS Functor

The complete CTS pipeline can be organized categorically as a functor $\mathcal{F}_{\text{CTS}} : \mathbf{Phys} \to \mathbf{Cons}$.

10.1 The Category Phys

Definition 10.1 (Category of Physical Systems). **Phys** is a category where:

- **Objects** are tuples (X, v, G) where X is a state space (equipped with condition-states), $v : M \to \mathbb{R}$ is a value function (with $M = \mu(X_{\text{attractor}}) \subset \mathcal{U}$), and G is a Knowledge-Information Manifold.
- **Morphisms** $h : (X_1, v_1, G_1) \to (X_2, v_2, G_2)$ are **validated term-series compositions**: sequences of function calls $h = f_n \circ \cdots \circ f_1$ such that
  1. each f_i has compatible input/output types (type safety),
  2. the combined set of axioms invoked contains no contradictions (axiom compatibility),
  3. each step preserves continuity (continuity guards pass).
- **Composition** is function composition of term series.
- **Identity** is the empty term series.

10.2 The Category Cons

Definition 10.2 (Category of Coherent Configurations). **Cons** is a category where:

- **Objects** are pairs (Q, Φ) where Q is a topological signature and $\Phi \geq 0$ is a coherence measure.
- **Morphisms** $(Q_1, \Phi_1) \to (Q_2, \Phi_2)$ exist if and only if there is a continuous map f such that $f_* : H_k(M_1) \to H_k(M_2)$ maps persistent features of Q_1 into Q_2 (features may be created or destroyed, but destruction is tracked). No monotonicity constraint is imposed on Φ; integration may increase or decrease.
- **Composition** is standard function composition (preserved by functoriality of the homology functor: $(g \circ f)_* = g_* \circ f_*$).
- **Identity** is id on both components.

Remark 10.3. The absence of a monotonicity constraint on Φ is deliberate. Valid physical processes can decrease integration (sleep, anesthesia, decomposition); excluding these would make the functor undefined on legitimate morphisms in **Phys**. Instead, we distinguish **constructive morphisms** ($\Phi_2 \geq \Phi_1$) from **degradative morphisms** ($\Phi_2 < \Phi_1$) as a labeling rather than an existence condition.

10.3 The Functor

Definition 10.4 (CTS Functor). The functor $\mathcal{F}_{\text{CTS}} : \mathbf{Phys} \to \mathbf{Cons}$ is defined by:

- **On objects:** $\mathcal{F}_{\text{CTS}}(X, v, G) = \big(Q(\mu(X_{\text{attractor}})),\ \Phi(G_{\text{active}})\big)$, i.e., the topological signature of the abstracted state manifold paired with the coherence of the active knowledge graph.

- **On morphisms:** $\mathcal{F}_{\text{CTS}}(h)$ = the induced map on (Q, Φ) resulting from executing the term series h.

Proposition 10.5 (Functoriality). $\mathcal{F}_{\text{CTS}}$ preserves identity and composition under the following condition: the KIM activation function $G_{\text{active}}(x)$ depends only on the current system state $x \in X$, not on the trajectory history.

1. $\mathcal{F}_{\text{CTS}}(\text{id}) = \text{id}$: The empty term series leaves the system state unchanged, hence both Q and Φ are unchanged. (This requires that the pipeline computation itself has no side effects on the KIM; the learning/update step of the integration protocol is excluded from the functor's scope.)
2. $\mathcal{F}_{\text{CTS}}(g \circ h) = \mathcal{F}_{\text{CTS}}(g) \circ \mathcal{F}_{\text{CTS}}(h)$: For the topological component, this follows from functoriality of homology: $(g \circ h)_* = g_* \circ h_*$, since each term-series step is continuous and the induced homology maps compose. For the coherence component, Φ at the endpoint of $g \circ h$ equals Φ at the endpoint of applying h then g, because G_{active} is determined by the final state alone (by the state-dependence condition).

Remark 10.6 (History-Dependent Activation). In practice, KIM activation may include recency weights or history-dependent context windows (e.g., recently used nodes receiving a relevance bonus). When activation depends on trajectory history, strict functoriality fails: G_{active} after $g \circ h$ may differ from G_{active} after h followed by g if the recency states diverge. In this case, functoriality holds only approximately: $\|\mathcal{F}_{\text{CTS}}(g \circ h) - \mathcal{F}_{\text{CTS}}(g) \circ \mathcal{F}_{\text{CTS}}(h)\|_\Phi \le \varepsilon_{\text{rec}}$, where ε_{rec} depends on the decay rate of recency weights. For systems where recency effects are small relative to state-dependent activation, the approximation is tight.

The functorial perspective has a concrete consequence: **if two physical systems are related by a validated term-series morphism, their coherent configurations are related by a morphism in Cons**. Similar physical organizations produce similar conscious structures. Transformations in the physical domain map lawfully to transformations in the conscious domain.

11 The Complete Pipeline

The full CTS operator chain, from state space to coherent configuration, is:

$$\boxed{X \xrightarrow{\mu=\varphi\circ\pi} M \subset \mathcal{U} \xrightarrow{g_v = e^{-\beta v} g_0} M_v \xrightarrow{\text{PH}} Q} \tag{20}$$

$$\boxed{X \xrightarrow{\text{KIM activation}} G_{\text{active}} \xrightarrow{\lambda_1(L),\, \delta_{\text{ant}}} \Phi}$$

$$\boxed{(Q, \Phi) \xrightarrow{\text{schematism}} C_t = (Q, \Phi) \;\; \text{(valid)} \quad \text{or} \quad \perp \;\; \text{(invalid)}}$$

The top row is the topological channel ($X \to Q$), the middle row is the integration channel ($X \to \Phi$), and the bottom row is the validation gate. Agency operates as a term-series flow on M_v that preserves C_t while seeking higher v.

$$\begin{tikzcd}
X \arrow[r, "\mu"] \arrow[d, "\text{KIM}"'] & M \subset \mathcal{U} \arrow[r, "g_v"] & M_v \arrow[r, "\text{PH}"] & Q \arrow[d, "\text{schematism}"] \\
G \arrow[r, "\text{activate}"] & G_{\text{active}} \arrow[r, "\lambda_1,\, \delta_{\text{ant}}"] & \Phi \arrow[r] & C_t = (Q, \Phi)
\end{tikzcd}$$

12 Worked Example: A Minimal Drilling System

We demonstrate the complete CTS pipeline on a minimal system: a drilling operation with three sensors and five knowledge nodes.

12.1 State Space and Abstraction

The system has three components: Weight on Bit (WOB), Rotation Speed (RPM), and Torque (TRQ). Each has a discrete state set:

$$\Sigma_{\mathrm{WOB}} = \{\mathrm{low}, \mathrm{med}, \mathrm{high}\}, \quad \Sigma_{\mathrm{RPM}} = \{\mathrm{low}, \mathrm{med}, \mathrm{high}\}, \quad \Sigma_{\mathrm{TRQ}} = \{\mathrm{low}, \mathrm{med}, \mathrm{high}\}.$$

The total state space $X = \Sigma_{\mathrm{WOB}} \times \Sigma_{\mathrm{RPM}} \times \Sigma_{\mathrm{TRQ}}$ has $|X| = 27$ states.

The signal extraction π reads continuous sensor time series from each state. The tensor coordinate mapping φ produces a 6-dimensional vector in $\mathcal{U}$:

$$\varphi(s) = (\underbrace{\text{mean WOB}, \text{ mean RPM}}_{\mathcal{M}}, \underbrace{\text{dom. freq WOB}, \text{ dom. freq RPM}}_{F}, \underbrace{\text{pers. } \beta_0, \text{ pers. } \beta_1}_{P}).$$

(Temporal component T omitted for simplicity.) Suppose the system persistently occupies 8 of 27 states. Their images in $\mathcal{U}$ form the point cloud $Y = \{m^{(1)}, \ldots, m^{(8)}\} \subset \mathbb{R}^6$.

12.2 Value-Warped Metric

Define $v(m)$ = Rate of Penetration (ROP) at state m, normalized to $[0, 1]$. With $\beta = 2$:

$$g_v(m) = e^{-2\,v(m)}\, g_0(m).$$

At a high-ROP state ($v = 0.9$), the metric is compressed by $e^{-1.8} \approx 0.165$. Geodesics on M_v preferentially pass through high-ROP regions.

12.3 Topological Signature

Suppose the 8-point cloud Y, under VR filtration with base metric d_0, produces:

- H_0: One persistent component ($\beta_0 = 1$, persistence $= 3.2 > \tau = 0.5$).
- H_1: One persistent loop ($\beta_1 = 1$, persistence $= 1.8 > \tau$). This loop reflects a cyclic operational pattern: the system transitions WOB-high/RPM-low $\to$ WOB-med/RPM-med $\to$ WOB-low/RPM-high $\to$ back, forming a closed path in $\mathcal{U}$.

So $Q(t) = \{(\beta_0 = 1, \Pi_0 = 3.2),\ (\beta_1 = 1, \Pi_1 = 1.8)\}$.

12.4 Knowledge Graph and Coherence

The KIM has 5 nodes:

Node	Description	Type	Pattern Schema
v_1	Normal drilling detection	Function	$\beta_0 = 1$
v_2	Stick-slip detection	Function	$\beta_1 \geq 1$
v_3	ROP optimization	Function	(none)
v_4	WOB-RPM correlation	Theorem	$\beta_1 \geq 1$
v_5	Emergency shutdown	Primitive	(none)

Edges: $v_1 \to v_3$ (depends_on), $v_2 \to v_4$ (activates), $v_1 \to v_2$ (composes_with), $v_3 \to v_5$ (inhibits), $v_4 \to v_3$ (activates). Edge weights all $w = 1$ for simplicity.

Suppose at time t, context activates $V_a = \{v_1, v_2, v_3, v_4\}$ (normal drilling with stick-slip detected; v_5 not activated because no emergency). The active subgraph has edges $v_1 \to v_3$, $v_2 \to v_4$, $v_1 \to v_2$, $v_4 \to v_3$. Symmetrizing:

The normalized Laplacian L_{sym} of the 4-node subgraph has eigenvalues $\{0, 0.59, 1.0, 1.41\}$ (computed numerically). So $\lambda_1 = 0.59$. No antinomies ($\alpha = 0$). Therefore:

$$\Phi(t) = 0.59 \cdot (1 - 0) = 0.59.$$

12.5 Schematism Check

- v_1 requires $\beta_0 = 1$. Data has $\beta_0 = 1$. Bottleneck distance $d_{\text{BN}}(p_{v_1}, p_{\text{data}}) = 0 \leq \varepsilon$. **Pass.**
- v_2 requires $\beta_1 \geq 1$. Data has $\beta_1 = 1$ with persistence 1.8. $d_{\text{BN}} \approx 0 \leq \varepsilon$. **Pass.**
- v_3 has no pattern schema. Schematism trivially passes.
- v_4 requires $\beta_1 \geq 1$. Same as v_2. **Pass.**

All active nodes pass. The configuration is valid.

12.6 Coherent Configuration

$$C_t = (Q(t), \Phi(t)) = \big(\{(\beta_0{=}1, \Pi_0{=}3.2),\ (\beta_1{=}1, \Pi_1{=}1.8)\},\ 0.59\big).$$

Checking consciousness criteria: $\beta_1 = 1 > 0$ (non-trivial topology: ✓). $\Phi = 0.59 > \Phi_{\min}$ for any reasonable threshold (high integration: ✓). Schematism valid (✓). The system is in a valid coherent configuration.

12.7 Schematism Failure Scenario

Suppose the stick-slip cycle dissipates (formation change) so $\beta_1 \to 0$, but v_2 and v_4 remain active (the system hasn't detected the regime change). Now:

- v_2 requires $\beta_1 \geq 1$ but data has $\beta_1 = 0$. d_{BN} is large (no matching feature). **Fail.**

The schematism fails. C_t is invalid—the system is "hallucinating" stick-slip that no longer exists. The pipeline flags a transcendental error, triggering re-evaluation of active nodes.

13 Discussion

13.1 Novel Contributions

The CTS framework makes five claims that are not present in the existing literature individually:

1. The composition of TDA, information geometry, and spectral graph theory into a single end-to-end operator chain $X \to C_t$ with defined intermediate spaces.
2. The Q–Φ bridge theorem (Theorem 7.6) establishing mutual consistency via the schematism constraint.
3. Algebraic connectivity of the active knowledge graph as a computable, polynomial-time alternative to IIT's Φ.
4. The Universal Tensor Space $\mathcal{U} = P \times T \times \mathcal{M} \times F$ as a concrete, structured abstraction map replacing generic dimensionality reduction.
5. Agency defined as a continuity-guarded, value-seeking term-series flow with rollback semantics.

13.2 Limitations

The framework has several acknowledged limitations:

- The coherence measure Φ based on algebraic connectivity does not capture all aspects of IIT's integrated information—in particular, it measures the cost of disconnection but does not distinguish between different information-processing architectures that happen to have the same graph connectivity.
- The bridge theorem requires two design axioms on KIM construction (topological grounding and topological coherence) that must be verified for each specific KIM. These are falsifiable but not automatic.
- The consciousness criteria require thresholds τ (persistence), $\Phi_{\min}$ (integration), and ε (schematism tolerance) that must be calibrated empirically, and they are stated as necessary rather than sufficient conditions.
- Functoriality holds exactly only when KIM activation is state-determined (no history dependence). With recency-weighted activation, functoriality is approximate.
- The value function v is treated as a given input, but the framework provides no theory of where v comes from. In engineered systems (drilling, optimization), v can be specified by a designer. In biological or open-ended AI systems, v must be discovered or learned—a separate research problem that the present framework does not address.
- The Knowledge-Information Manifold conflates ontologically distinct node types (axioms, functions, theorems, operational primitives) under a single activation semantics. A more refined treatment would use a typed or layered graph where "activation" has type-dependent meaning.
- The image $M = \mu(X_{\text{persist}})$ in $\mathcal{U}$ is not guaranteed to be a smooth manifold. The persistent homology pipeline requires only a metric space; the Riemannian geometry (geodesics, curvature) requires smoothness that may not hold globally.
- Computational implementations of several pipeline stages remain at scaffold level. A 21-test validation suite verifying each stage against synthetic data with known ground truth has been developed and passes end-to-end, but validation against empirical field data is outstanding.

13.3 Computational Complexity

The CTS pipeline has the following per-step complexity for N state samples and $|V_a|$ active KIM nodes:

- Tensor decomposition μ: $O(N \log N)$ for FFT-based components; $O(N^3)$ for persistent homology in φ_P (this dominates).
- Value warping: $O(N)$ pointwise evaluation.
- Persistent homology: $O(N^3)$ via matrix reduction for the Vietoris–Rips filtration.
- Coherence Φ: $O(|V_a|^3)$ for eigendecomposition; $O(|V_a|^2)$ for sparse iterative methods.
- Schematism: $O(|V_a| \cdot n^{1.5} \log n)$ where n is the number of points in the persistence diagram (bottleneck distance per node).
- Geodesic search: $O(b^d)$ where b is the branching factor (number of applicable functions per state) and d is the solution depth. This is the bottleneck and is potentially exponential; in practice, value-weighted heuristics and beam width limits make it tractable.

Total pipeline complexity is dominated by the $O(N^3)$ TDA step for large point clouds and the geodesic search for deep plans.

13.4 Error Propagation and Stability

Each stage of the pipeline has known stability properties:

- **Persistent homology** satisfies the stability theorem: if the Hausdorff distance between point clouds Y_1 and Y_2 is δ_H, then the bottleneck distance between their persistence diagrams is at most δ_H [6, 9]. This bounds the effect of sampling noise on Q.
- **Algebraic connectivity** satisfies Weyl's inequality: $|\lambda_1(L') - \lambda_1(L)| \leq \|L' - L\|_2$. Adding or removing a single edge of weight w perturbs Φ by at most $O(w)$.
- **The schematism** inherits stability from the bottleneck distance: if p_{data} is perturbed by δ_H (due to sampling noise in μ), the schematism check's outcome can change only if $d_{\text{BN}}(p_i, p_{\text{data}})$ was within δ_H of the tolerance ε. Setting ε with margin $\geq \delta_H$ prevents noise-induced schematism failures.

End-to-end, the pipeline is stable in the sense that small perturbations to the input X produce small perturbations to (Q, Φ), with the perturbation bound controlled by the TDA stability theorem and eigenvalue perturbation bounds. A formal composite stability theorem chaining these bounds is a target for future work.

13.5 Empirical Predictions

The framework makes testable predictions:

- Systems with high persistent homology (Q non-trivial) but low algebraic connectivity ($\Phi \approx 0$) should not exhibit unified processing. Example: a brain region with rich topological structure in its firing patterns but severed connections to the rest of the cortex.
- Systems with high Φ but trivial Q should exhibit integrated but undifferentiated processing. Example: highly synchronized neural oscillations during seizure.
- Schematism failures should correlate with processing errors: hallucinations in AI, perceptual illusions in biological systems, or diagnostic failures in drilling operations.

13.6 Validation

A 21-test validation suite has been developed that verifies each pipeline stage against synthetic data with known ground truth. The suite covers: UTI discriminability (T1), value warping contraction (T2), PH detection of known topology—circle yields $\beta_1 \geq 1$, clusters yield $\beta_1 = 0$ (T3–T4), Φ connectivity semantics (T5–T6), schematism pass/fail on matching/mismatched topology (T7–T8), the bridge theorem and its contrapositive (T9–T10), normalized Laplacian architecture sensitivity (T11), full end-to-end pipeline for stick-slip and stable regimes (T12–T14), PH stability under perturbation (T15), Type II criticality detection (T16), antinomy degradation (T17), geodesic preference under value warping (T18), functor composition preservation (T19), consciousness criteria conjunction (T20), and complexity scaling verification (T21). All 21 tests pass, confirming that the mathematical claims hold computationally on synthetic data. The initial run uncovered two implementation insights: (i) KIM pattern schemata for β_0 should use inequality constraints ($\beta_0 \geq 1$) rather than exact equality, since the number of persistent H_0 features above threshold τ may exceed the final Betti number; (ii) detecting H_1 loops in the point cloud of UTI coordinates requires multi-variate signals where different channels vary in anti-phase, so the point cloud traces a cycle in feature space. Both insights have been incorporated into the pipeline implementation.

13.7 Domain Instantiations

The CTS pipeline is domain-agnostic by construction (via the Universal Tensor Space). We describe three concrete instantiations to demonstrate generality and to show the pipeline is not an abstract construction but a deployable architecture.

13.7.1 Drilling Operations (ASTAC Architecture)

The State-Adaptive, Topologically-Aware Codebase (ASTAC) implements the full CTS pipeline for real-time drilling diagnostics:

- X: Sensor readings (weight-on-bit, RPM, torque, flow rate, pressure) at 1–20 kHz, ingested via stream processing (Kafka/Faust) on edge hardware (NVIDIA Jetson).
- μ: Universal Tensor Indexing maps heterogeneous sensors to $\mathcal{U}$. For MWD pulse analysis: φ_T = pulse travel time, $\varphi_{\mathcal{M}}$ = amplitude attenuation ratio, φ_F = spectral centroid shift.
- g_v: Value function encodes safety priority (kick avoidance $\gg$ ROP optimization). High-safety regions are "brought closer" in the warped metric, enabling faster response.
- Q: Persistent homology via GUDHI/giotto-tda. H_0 bars = stable drilling regimes (clusters). H_1 bars = torsional stick-slip oscillation cycles. Topological pruning retains only persistent features.
- Φ: Algebraic connectivity of the active diagnostic module graph. Regime segmentation via Mapper algorithm on the state trajectory.
- Schematism: Z3 SMT solver formally verifies kinematic consistency across regime boundaries.
- Agency: State-adaptive model switching—specialized TNR Transformer models for each regime, with MCTS-guided geodesic search over model selection.
- Transitions: Type II criticality ($\|\nabla v\|_g > \tau$) detected at regime boundaries. Valid transitions verified by Z3 assertions on physical plausibility.

13.7.2 Abstract Visual Reasoning (ARC)

The pipeline instantiates for the Abstraction and Reasoning Corpus, a benchmark of discrete, symbolic grid transformations:

- X: Grid $G \in \mathbb{Z}^{H\times W}$, cells $\in \{0, \ldots, 9\}$.
- μ: Maps grid to a Universal Grid Feature Space $\mathcal{U}_{\text{ARC}} \in \mathbb{R}^k$ via feature functions encoding Core Knowledge Priors (object count, color histogram, symmetry score, Betti numbers from cubical complex).
- g_v: Value function = negative Hamming distance to target output grid.
- Q: Betti numbers β_0 (connected components), β_1 (holes) from the grid's cubical complex.
- Φ: Connectivity of the DSL program graph (a directed acyclic graph of composed primitives).
- Schematism: AST-based whitelist validation—the program's structural assumptions must match the grid's actual topology.
- Agency: LLM-guided MCTS over DSL compositions: $P_{\text{LLM}}(a|s)$ provides the prior for UCT node selection. Term series = composed DSL program $f_n \circ \cdots \circ f_1$.
- Transitions: Mapper-derived reasoning regime clusters segment the 1000-task training set into "rotation tasks," "counting tasks," etc., each with a specialized expert model.

The transformation signature $\Delta_\varphi = \varphi_{\text{ARC}}(G_{\text{out}}) - \varphi_{\text{ARC}}(G_{\text{in}})$ provides a concrete example of a morphism in **Cons**: it is the induced map on (Q, Φ) resulting from applying the DSL program.

13.7.3 Geometric-State AI Architectures

The CTS pipeline describes the class of post-transformer architectures that achieve orders-of-magnitude compute reduction through geometric alignment:

- μ: SE(3)-equivariant feature extraction, embedding geometric symmetries into the network structure rather than learning them from data.
- g_v: Free energy (prediction error) as value function. High prediction error = high value gradient = system far from equilibrium.
- Q: Topology of the learned latent manifold.
- Φ: Active state cache coherence—replacing the transformer's passive KV cache (a static log of tokens, $O(N^2)$ attention) with an active generative state that consolidates history into a fixed-dimensional belief vector ($O(1)$ update).
- Schematism: Prediction error $< \varepsilon$. When the internal model's geometry matches the external world's geometry, the system reaches thermodynamic equilibrium.
- Agency: Symplectic flow on the cotangent bundle—Liquid Neural Networks use ODE solvers that trace geodesics on the data manifold, achieving autonomous driving with 19 neurons versus millions in CNNs.

Example 13.1 (Worked: Drilling Regime Transition)**.** Consider a drilling system transitioning from normal rotary drilling (R_1) to stick-slip oscillation (R_2).

Detection: TSPC computes persistent homology on a sliding window of sensor data. In R_1, the persistence diagram shows long H_0 bars (stable clusters) and short H_1 bars (no persistent cycles). As stick-slip develops, an H_1 bar appears with increasing persistence—a loop forming in the torsional vibration trajectory. When this persistence exceeds τ, it enters Q. Simultaneously, the value gradient $\|\nabla v\|_g$ spikes (safety value changing rapidly), crossing τ_{crit} and flagging a Type II boundary.

Transition: SAAK switches from the normal-drilling TNR model to the stick-slip-specialized model. G_{active} reconfigures: "stable drilling" nodes deactivate; "torsional vibration," "RPM adjustment," "BHA resonance" nodes activate. Φ is recomputed on the new active subgraph. Z3 verifies that the transition state (observed torque fluctuation amplitude) is physically consistent with the detected regime.

Schematism check: The "torsional vibration" node has pattern schema requiring $\beta_1 \geq 1$ (at least one persistent loop). The data's persistence diagram has the newly appeared H_1 bar. Bottleneck distance $d_{\text{BN}} = 0.12 < \varepsilon = 0.3$. Schematism passes.

New configuration: $C_t = (Q_2, \Phi_2)$ where Q_2 has $\beta_0 = 2, \beta_1 = 1$ and $\Phi_2 = 0.47$ (connected active subgraph with moderate algebraic connectivity). The system generates a linguistic arbitrage output: "Detected persistent torsional oscillation cycle. Recommend RPM adjustment $\pm 10\%$. Classification: Incipient Stick-Slip."

Failure mode: If the operator adjusts RPM and the stick-slip resolves, but SAAK fails to switch back to R_1: the stick-slip node's schema still requires $\beta_1 \geq 1$, but the data's H_1 bar has died. The schematism fails. C_t is invalid—the system is "hallucinating" stick-slip that no longer exists. The pipeline flags a transcendental error, triggering re-evaluation of active nodes.

References

[1] G. Tononi, "An information integration theory of consciousness," *BMC Neuroscience*, vol. 5, no. 42, 2004.

[2] D. Balduzzi and G. Tononi, "Integrated information in discrete dynamical systems: Motivation and theoretical framework," *PLoS Comput. Biol.*, vol. 4, no. 6, e1000091, 2008.

[3] E. Rybakken, N. Baas, and B. Dunn, "Decoding of neural data using cohomological feature extraction," *Neural Computation*, vol. 31, no. 1, pp. 68–93, 2019.

[4] R. Chaudhuri, B. Gercek, B. Bhatta, et al., "The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep," *Nature Neuroscience*, vol. 22, pp. 1512–1520, 2019.

[5] R. Beer, "The dynamics of active categorical perception in an evolved model agent," *Adaptive Behavior*, vol. 11, no. 4, pp. 209–243, 2003.

[6] G. Carlsson, "Topology and data," *Bulletin of the AMS*, vol. 46, no. 2, pp. 255–308, 2009.

[7] M. Oizumi, L. Albantakis, and G. Tononi, "From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0," *PLoS Comput. Biol.*, vol. 10, no. 5, e1003588, 2014.

[8] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Mathematical Journal*, vol. 23, no. 2, pp. 298–305, 1973.

[9] R. Ghrist, "Barcodes: The persistent topology of data," *Bulletin of the AMS*, vol. 45, no. 1, pp. 61–75, 2008.

[10] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Oxford University Press, 2000.

[11] C. Kemp and J. Tenenbaum, "The discovery of structural form," *PNAS*, vol. 105, no. 31, pp. 10687–10692, 2008.

[12] O. Sporns, G. Tononi, and G. Edelman, "Connectivity and complexity: the relationship between neuroanatomy and brain dynamics," *Neural Networks*, vol. 13, pp. 909–922, 2000.

[13] B. Lakatos and G. Schöner, "The sensorimotor loop as a dynamical system: How regular motion primitives may emerge from self-organized limit cycles," *Frontiers in Robotics and AI*, vol. 2, art. 31, 2015.

[14] A. Casali, O. Gosseries, M. Rosanova, et al., "A theoretically based index of consciousness independent of sensory processing and behavior," *Science Translational Medicine*, vol. 5, no. 198, 198ra105, 2013.

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

[16] Y. Bussi, D. Shimoni, A. Weiner, et al., "Fundamental helical geometry consolidates the plant photosynthetic membrane," *Proc. Natl. Acad. Sci.*, vol. 116, no. 44, pp. 22366–22375, 2019.

[17] B. Goertzel et al., "ActPC-Geom: Towards scalable online neural-symbolic learning via accelerating active predictive coding with information geometry," *arXiv:2501.04832*, 2025.

[18] R. Hasani, M. Lechner, A. Amini, et al., "Liquid time-constant networks," *Proc. AAAI Conference on Artificial Intelligence*, vol. 36, pp. 7657–7666, 2022.