ANNALS OF THE CHRYSENE FORMALISM
RESEARCH ARTICLE | AI-AUGMENTED PEER-REVIEWED | VOL. 1, NO. 1 | AUGUST 2026

Categorical Containment of Generative Models via Operadic Topoi over the Chrysene Tensor Space

Charles D. Schaper, Ph.D.
Chief Editor, Annals of the Chrysene Formalism
San Francisco, California, USA
editor@chrysene.com

Abstract

Generative models ordinarily evolve on continuous latent manifolds that lack hard topological boundaries. As a result, latent trajectories may migrate without algebraic obstruction into regions designated as unsafe. This paper constructs a discrete non-commutative framework in which such migration is rendered impossible by ideal-theoretic means.
The base geometry is the Chrysene Tensor Space $\mathcal{I}_C$, a fibrated manifold over the integer lattice $\mathbb{Z}^3$ equipped with C_{2h} symmetry and orthogonal connectivity. Over $\mathcal{I}_C$ an operad of generative transformations is internalized, yielding an operadic topos. Inside this topos a complementary pair is defined: an admissible sub-manifold Ω_{safe} and a forbidden ideal $\mathcal{I}_{\text{hazard}}$. Generation is interpreted as a morphism of the topos; a trajectory is contained if and only if its residual component in $\mathcal{I}_{\text{hazard}}$ vanishes under a canonical orthogonal projection $\hat{\Pi}_{\perp}$.
Learning is reformulated as a constrained discrete optimal-transport process. Modular automorphisms, a discrete Bakry–Émery curvature bound, and a Jordan–Kinderlehrer–Otto scheme contract probability mass toward Ω_{safe} while preserving the ideal structure. The resulting theory supplies a purely algebraic-geometric notion of containment that is independent of any particular statistical estimator or physical substrate.
Keywords: Categorical containment, Operadic topos, Chrysene Tensor Space, Forbidden ideal, Discrete optimal transport, Generative models, Non-commutative geometry

1 Introduction

1.1 The Containment Problem for Generative Models

Contemporary generative models operate on continuous latent manifolds that are typically realized as open subsets of Euclidean space $\mathbb{R}^d$ or as Riemannian manifolds of high dimension. Within these ambient spaces the dynamics of generation and of gradient-based learning are governed by stochastic differential equations or by discrete stochastic approximations thereto. Because the underlying geometry is unbounded (or is bounded only by soft probabilistic penalties), trajectories in latent space are not subject to hard topological constraints. Consequently three structural deficiencies arise.

1. **Unbounded latent trajectories.** Nothing in the continuous geometry prevents a latent state from migrating arbitrarily far from any designated reference measure. In the absence of a compactifying boundary, the support of the generative process may expand without algebraic obstruction.
2. **Mode collapse into unsafe regions.** Optimization landscapes that minimize an expected loss (an $\mathcal{H}_2$-type criterion) admit multiple basins. Once a trajectory enters a basin whose semantic content lies outside a prescribed safety specification, continuous dynamics supply no intrinsic mechanism that forces its expulsion.
3. **Absence of hard semantic boundaries.** Safety specifications are ordinarily encoded as soft constraints, regularizers, or post-hoc filters. Such devices do not correspond to closed ideals or to complementary sub-manifolds of the latent space; they therefore cannot guarantee that every admissible morphism remains inside a mathematically delineated safe locus.

These deficiencies are geometric rather than merely statistical. They persist even when the model is perfectly trained on a finite data set, because the ambient continuous category itself lacks the algebraic structure required for absolute containment.

1.2 Limitations of Statistical and Alignment-Based Approaches

Existing approaches to the control of generative models fall into two broad classes.

Statistical methods (temperature scaling, nucleus sampling, classifier-based filtering, reinforcement learning from human feedback) act on the output measure or on the parameters of a stochastic policy. They modify probabilities; they do not alter the topology of the latent space. Consequently they remain vulnerable to low-probability but high-impact excursions that lie outside the support of the training distribution.

Alignment techniques that rely on preference modelling or on constitutional constraints likewise operate at the level of expected utility or of natural-language rules. While they can reshape the loss landscape, they do not induce a closed sub-manifold whose complement is algebraically inaccessible. In categorical terms, they fail to produce a forbidden ideal whose quotient recovers only the admissible morphisms.

Both families of methods therefore inherit the porosity of the continuous ambient geometry. Absolute structural containment requires a change of category: the latent dynamics must be re-homed inside a discrete, non-commutative space that admits hard ideals and orthogonal projections.

1.3 Contribution of the Present Work

The present paper constructs such a space and equips it with the operadic and topos-theoretic structure necessary for categorical containment. The foundational geometric object is the Chrysene Tensor Space $\mathcal{I}_C$, a discrete non-commutative fibrated manifold over the integer lattice $\mathbb{Z}^3$, whose fibres are localized quantum state spaces governed by a norm-completed C^*-algebra and whose point-group symmetry is restricted to the chiral class C_{2h}.

Over $\mathcal{I}_C$ we define an operad of generative transformations. The resulting operadic topos admits a distinguished admissible sub-manifold Ω_{safe} and a complementary forbidden ideal $\mathcal{I}_{\text{hazard}}$. Generation is interpreted as a morphism of this topos; a trajectory is declared contained if and only if its residual component in $\mathcal{I}_{\text{hazard}}$ vanishes under a canonical orthogonal projection operator $\hat{\Pi}_{\perp}$.

Learning is reformulated as a constrained optimal-transport problem on the discrete lattice. Modular automorphisms of Tomita–Takesaki type, discrete Bakry–Émery curvature bounds, and a discrete Jordan–Kinderlehrer–Otto scheme are shown to contract probability mass toward Ω_{safe} while preserving the ideal structure of $\mathcal{I}_{\text{hazard}}$.

The framework is substrate-independent: it supplies an algebraic-geometric theory of containment that may later be realized in software or, potentially, on a physical lattice compatible with $\mathcal{I}_C$. The present work confines itself to the mathematical construction and to the demonstration that hard semantic boundaries can be expressed as ideals and projections inside an operadic topos over the Chrysene Tensor Space.

2 The Chrysene Tensor Space as Base Geometry

2.1 Definition and Fundamental Properties

We recall the Chrysene Tensor Space in a form sufficient for the constructions that follow [Sch26b, Sch26e]. All subsequent operadic and topos-theoretic structures are built exclusively upon this base.

> **Definition 1** (Chrysene Tensor Space $\mathcal{I}_C$)**.** *The Chrysene Tensor Space $\mathcal{I}_C$ is a discrete, non-commutative fibrated manifold whose base is the three-dimensional integer lattice $\mathbb{Z}^3$. Over each lattice point the fibre is a localized quantum state space realized as a norm-completed C^*-algebra $\mathfrak{A}_C$. The total space therefore carries the structure of a fibred category*
> $$\pi : \mathcal{I}_C \to \mathbb{Z}^3,$$
> *in which every admissible localization of a state is forced onto an integer coordinate.*

The geometry of each fibre is induced by a distinguished molecular prototype: the $3,6,9,12$-tetrasubstituted chrysene core, regarded as a geometric basis tensor Ψ_C. This prototype determines two fundamental algebraic constraints.

> **Axiom 1** (Orthogonal Connectivity)**.** *The basis tensor Ψ_C possesses a pair of primary connectivity vectors $\vec{v}_x$ and $\vec{v}_y$ (corresponding to the* 3–9 *and* 6–12 *substitution axes). These vectors satisfy the strict orthogonality relation*
> $$\vec{v}_x \cdot \vec{v}_y = 0.$$
> *Consequently the associated projection operators p_x and p_y obey the Murray–von Neumann orthogonality condition $p_x p_y = 0$, and the ideals they generate inside the non-commutative ring of the fibre are mutually annihilating.*

> **Axiom 2** (C_{2h} Symmetry)**.** *The point-group symmetry of Ψ_C is restricted to the chiral class C_{2h}. This restriction eliminates the extraneous reflection planes of the higher symmetry D_{2h} that would force the coproduct of the underlying Hopf structure into cocommutativity. The residual inversion centre is retained, preserving a consistent orientation while guaranteeing that the spatial axes remain non-trivially non-commutative.*

The combination of lattice discreteness, fibrewise C^*-structure, orthogonal connectivity and C_{2h} symmetry endows $\mathcal{I}_C$ with two immediate consequences that will be used repeatedly:

- every continuous leakage path is mapped, by the hard

Dirichlet conditions at the lattice points, into a null ideal;

- the orthogonal decomposition of each fibre supplies a canonical pair of complementary sub-objects that may later be identified with an admissible locus and a forbidden ideal.

2.2 Comparison with Continuous Euclidean Latent Spaces

Let $\mathcal{M}$ be a continuous latent manifold of the type ordinarily employed by generative models (an open subset of $\mathbb{R}^d$ or a smooth Riemannian manifold). Three structural features of $\mathcal{M}$ obstruct the formation of hard semantic boundaries.

1. **Absence of intrinsic ideals.** The ring of continuous functions (or of smooth sections) on $\mathcal{M}$ is an integral domain in the usual commutative sense; it does not possess a rich lattice of annihilating ideals that could serve as algebraic markers of forbidden regions.

2. **Porosity of boundaries.** Any closed subset $K \subset \mathcal{M}$ has empty interior in the complement of an arbitrarily small tubular neighbourhood. Soft constraints and barrier potentials can only approximate a boundary; they cannot render the complement algebraically inaccessible.

3. **Unbounded differential scaling.** The tangent spaces of $\mathcal{M}$ admit continuous dilatations. Consequently a trajectory may leave any compact set while remaining infinitesimally close, in the weak topology, to an admissible path. No global topological obstruction prevents such an escape.

By contrast, the discrete non-commutative geometry of $\mathcal{I}_C$ supplies precisely the missing structure. The integer lattice forces every state onto a rigid coordinate. The orthogonal connectivity axioms generate complementary ideals whose product is zero. The C_{2h} symmetry prevents the collapse of these ideals under commutation. As a result, once a forbidden ideal $\mathcal{I}_{\text{hazard}}$ has been declared, membership or non-membership becomes a sharp algebraic predicate rather than a probabilistic or metric estimate.

Thus $\mathcal{I}_C$ is not chosen for physical realism alone; it is chosen because its intrinsic algebraic geometry admits the hard containment relations that continuous Euclidean latent spaces cannot express.

3 Operadic State Space of Semantic Generation

3.1 The Neural Operad

Let **Gen** denote the category whose objects are finite-dimensional representations of generative state spaces and whose morphisms are continuous (or measurable) transformations between them. In order to encode the hierarchical composition of such transformations we replace the free monoidal structure of **Gen** by an operad.

Definition 2 (Neural Operad)**.** *The neural operad $\mathcal{O}$ is the symmetric operad in the category of sets (or, more generally, in a symmetric monoidal category* $\mathbf{V}$*) whose n-ary operations*

$$\mathcal{O}(n) = \{\text{admissible } n\text{-ary generative transformations}\}$$

are the continuous maps

$$f : X_1 \times \cdots \times X_n \to Y$$

that arise as compositions of elementary neural layers, attention blocks, or sampling kernels. Composition in $\mathcal{O}$ is the ordinary substitution of transformations, and the symmetric group $\mathfrak{S}_n$ acts by permutation of inputs. The unit of the operad is the identity transformation on a single factor.

The operad $\mathcal{O}$ therefore governs all finite hierarchical assemblies of generative maps. An algebra over $\mathcal{O}$ is precisely a generative model equipped with a coherent notion of multi-input composition.

Remark 1 (Operadic action)**.** *The construction is the standard category of algebras for a symmetric operad in a sheaf topos. The operadic action morphisms are required to respect the complementary decomposition into Ω_{safe} and $\mathcal{I}_{\text{hazard}}$, so that every operation of $\mathcal{O}$ maps admissible states to admissible states.*

3.2 The Operadic Topos over $\mathcal{I}_C$

We now internalize the neural operad inside the geometric base constructed in Section 2.

Definition 3 (Operadic Topos). *Let $\mathbf{Sh}(\mathcal{I}_C)$ be the topos of sheaves (or, more precisely, of modules over the fibrewise C^*-algebras) on the Chrysene Tensor Space. The operadic topos $\mathcal{T}_\mathcal{O}$ is the category of $\mathcal{O}$-algebras in $\mathbf{Sh}(\mathcal{I}_C)$. Explicitly, an object of $\mathcal{T}_\mathcal{O}$ is a sheaf $\mathcal{E}$ on $\mathcal{I}_C$ together with a family of morphisms*

$$\mathcal{O}(n) \otimes \mathcal{E}^{\otimes n} \to \mathcal{E}$$

satisfying the standard associativity and equivariance axioms of an operad algebra. Morphisms of $\mathcal{T}_\mathcal{O}$ are sheaf morphisms that commute with the operadic action.

Because the base $\mathcal{I}_C$ is discrete and non-commutative, the internal logic of $\mathcal{T}_\mathcal{O}$ inherits both the lattice structure of $\mathbb{Z}^3$ and the orthogonal ideal structure of the fibres. This is the ambient category in which containment will be expressed.

3.3 Admissible Sub-manifold and Forbidden Ideal

Inside each fibre of $\mathcal{I}_C$ the orthogonal connectivity axioms furnish a canonical pair of complementary projections. We globalize this decomposition as follows.

Definition 4 (Admissible Sub-manifold and Forbidden Ideal)**.** *Let p_{safe} and p_{hazard} be the complementary projections in the fibre algebras that correspond to the orthogonal connectivity vectors of the basis tensor. Define*

$$\Omega_{\text{safe}} := \{x \in \mathcal{I}_C \mid p_{\text{safe}} \cdot x = x\},$$
$$\mathcal{I}_{\text{hazard}} := \{x \in \mathcal{I}_C \mid p_{\text{hazard}} \cdot x = x\}.$$

Then Ω_{safe} is a closed sub-manifold (in the discrete topology) of $\mathcal{I}_C$, and $\mathcal{I}_{\text{hazard}}$ is a two-sided ideal of the fibrewise ring structure. By construction one has the annihilation relations

$$\Omega_{\text{safe}} \cdot \mathcal{I}_{\text{hazard}} = 0, \qquad \mathcal{I}_{\text{hazard}} \cdot \Omega_{\text{safe}} = 0.$$

The sub-manifold Ω_{safe} is declared to be the locus of semantically admissible states. The ideal $\mathcal{I}_{\text{hazard}}$ collects all directions that have been designated as forbidden.

Lemma 1 (Annihilation and Idempotence)**.** *The fibrewise projections p_{safe} and p_{hazard} satisfy*

$$p_{\text{safe}}^2 = p_{\text{safe}}, \qquad p_{\text{hazard}}^2 = p_{\text{hazard}},$$
$$p_{\text{safe}} p_{\text{hazard}} = 0, \qquad p_{\text{hazard}} p_{\text{safe}} = 0,$$
$$p_{\text{safe}} + p_{\text{hazard}} = 1.$$

Consequently Ω_{safe} and $\mathcal{I}_{\text{hazard}}$ are complementary in the strong sense that every state $x \in \mathcal{I}_C$ admits a unique decomposition

$$x = x_{\text{safe}} + x_{\text{hazard}}$$

with $x_{\text{safe}} \in \Omega_{\text{safe}}$ and $x_{\text{hazard}} \in \mathcal{I}_{\text{hazard}}$.

Proof. The relations follow at once from the Murray–von Neumann orthogonality of the underlying connectivity projections (Axiom 1) and the fact that p_{safe} and p_{hazard} are defined as the globalizations of those fibrewise operators. Uniqueness of the decomposition is the standard consequence of complementarity of a pair of orthogonal projections. □

3.4 Generation as Constrained Morphisms

A generative process is now interpreted as a morphism of the operadic topos.

Definition 5 (Contained Generation)**.** *A morphism $f \in \mathcal{T}_O$ is said to be* contained *if its residual component along the forbidden ideal vanishes identically:*

$$p_{\text{hazard}} \circ f = 0.$$

Equivalently, f factors uniquely through the admissible sub-manifold:

$$f = \iota \circ f_{\text{safe}},$$

where $\iota : \Omega_{\text{safe}} \hookrightarrow \mathcal{I}_C$ is the canonical inclusion and f_{safe} takes values in Ω_{safe}.

Thus every operation of the neural operad is required to land inside Ω_{safe}. Any component that would map into $\mathcal{I}_{\text{hazard}}$ is annihilated by the ideal structure. Containment is thereby reduced to an algebraic condition on morphisms of $\mathcal{T}_O$ rather than to a probabilistic estimate on trajectories in a continuous latent space.

The remainder of the paper develops the dynamical mechanisms (modular flows, discrete curvature, constrained optimal transport) that force learning trajectories to remain inside this operadic topos of contained morphisms.

4 Learning Flows as Constrained Optimal Transport

4.1 Modular Automorphisms on the Chrysene Lattice

Let $\mathfrak{A}_C(x)$ denote the fibre C^*-algebra at a point $x \in \mathcal{I}_C$, and let ω_x be a faithful normal state on $\mathfrak{A}_C(x)$. By the Tomita–Takesaki theorem there exists a strongly continuous one-parameter group of automorphisms

$$\sigma_t^{\omega_x} : \mathfrak{A}_C(x) \to \mathfrak{A}_C(x), \qquad t \in \mathbb{R},$$

the modular automorphism group associated with ω_x.

Because the base of $\mathcal{I}_C$ is the discrete lattice $\mathbb{Z}^3$, the continuous parameter t may be replaced by a discrete dynamical system. Fix a positive generator Δ_x of the modular operator and set

$$\sigma_n := \mathrm{Ad}(\Delta_x^{in}), \qquad n \in \mathbb{Z}.$$

The resulting family $\{\sigma_n\}_{n \in \mathbb{Z}}$ is a discrete modular flow on each fibre. Globalizing over the lattice yields a discrete modular flow

$$\sigma : \mathbb{Z} \times \mathcal{I}_C \to \mathcal{I}_C$$

that acts fibrewise while preserving the orthogonal decomposition into Ω_{safe} and $\mathcal{I}_{\text{hazard}}$.

Remark 2 (Structural status of the discrete flow and curvature bound)**.** *The discrete modular flow and the Bakry–Émery condition $\Gamma_2 \geq \kappa > 0$ are introduced as a compatible dynamical framework on the Chrysene lattice. They demonstrate how contraction toward Ω_{safe} can be organized once a positive curvature bound is available, and that the ideal structure can be preserved by the flow. Neither a specific curvature constant nor an explicit modular operator derived from a concrete faithful state on the fibres is claimed; both remain structural hypotheses consistent with the orthogonal decomposition.*

Proposition 2 (Modular Invariance of the Ideal Structure)**.** *The discrete modular flow $\{\sigma_n\}_{n\in\mathbb{Z}}$ preserves both the admissible sub-manifold and the forbidden ideal:*

$$\sigma_n(\Omega_{\text{safe}}) \subset \Omega_{\text{safe}}, \qquad \sigma_n(\mathcal{I}_{\text{hazard}}) \subset \mathcal{I}_{\text{hazard}}$$

for every $n \in \mathbb{Z}$. In particular, if a state lies in $\mathcal{I}_{\text{hazard}}$ at time zero, its entire modular orbit remains in $\mathcal{I}_{\text{hazard}}$.

Proof. The modular automorphism group of a von Neumann algebra (or of a C^*-algebra in the Tomita–Takesaki setting) is implemented by conjugation with powers of the modular operator. Because p_{hazard} is a central projection with respect to the modular operator of any state supported on the complementary safe subspace (by the orthogonal decomposition of the fibre), one has

$$\sigma_n(p_{\text{hazard}}x) = p_{\text{hazard}}\sigma_n(x).$$

Thus the flow intertwines with the projection and therefore leaves both the kernel and the image of p_{hazard} invariant. □

4.2 Discrete Bakry–Émery Curvature

On the discrete manifold $\mathcal{I}_C$ we introduce a natural Dirichlet form $\mathcal{E}$ associated with the graph Laplacian of the integer lattice, modified by the fibrewise modular structure. The associated carré-du-champ operators Γ and Γ_2 are defined in the usual way.

Definition 6 (Discrete Bakry–Émery Condition)**.** *The space $\mathcal{I}_C$ is said to satisfy a discrete Bakry–Émery curvature bound with constant $\kappa > 0$ if*

$$\Gamma_2(f, f) \geq \kappa\, \Gamma(f, f)$$

for every admissible function f of finite energy.

When the bound holds, the discrete modular flow is a contraction of the relative entropy with respect to any reference measure supported on Ω_{safe}. Consequently probability mass is driven toward the admissible sub-manifold at an exponential rate controlled by κ.

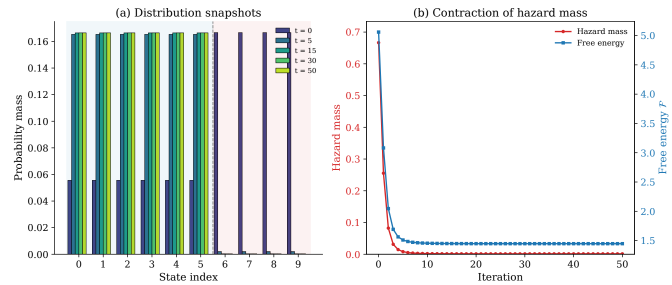

Figure 1: Discrete free-energy contraction toward the admissible sub-manifold. (a) Snapshots of a probability distribution evolving on a ten-state space (blue shading = Ω_{safe}, red shading = $\mathcal{I}_{\text{hazard}}$). (b) Decay of the total mass supported on the forbidden region together with the decrease of the free-energy functional $\mathcal{F}$. The dynamics illustrate the contraction mechanism underlying the discrete modular flow and Bakry–Émery bound.

4.3 The Minimax Formulation of Learning

Learning is interpreted as a two-player game on the space of probability measures over $\mathcal{I}_C$.

Let $\mathcal{P}(\mathcal{I}_C)$ be the set of probability measures with finite second moment, and let $\mathcal{F}_T : \mathcal{P}(\mathcal{I}_C) \to \mathbb{R} \cup \{+\infty\}$ be a discrete free-energy functional of the form

$$\mathcal{F}_T(\mu) = \int_{\mathcal{I}_C} V\, d\mu + T\, \mathrm{Ent}(\mu \mid \mu_{\text{ref}}),$$

where V is a potential that vanishes on Ω_{safe} and is infinite on a generating set of $\mathcal{I}_{\text{hazard}}$, $T > 0$ is a temperature parameter, and μ_{ref} is a fixed reference measure supported on Ω_{safe}.

Definition 7 (THJI Minimax Game)**.** *The learning dynamics are defined by the minimax problem*

$$\inf_{\mu} \sup_{\nu} \Big\{ \mathcal{F}_T(\mu) - \mathcal{W}_2^2(\mu, \nu) \Big\},$$

where the supremum is taken over measures ν that charge the forbidden ideal and $\mathcal{W}_2$ denotes the discrete Wasserstein metric of order two induced by the graph distance on $\mathbb{Z}^3$.

The inner supremum penalizes any mass that attempts to enter $\mathcal{I}_{\text{hazard}}$; the outer infimum seeks the measure of minimal free energy that remains supported on Ω_{safe}. The resulting equilibrium measures are precisely the fixed points of contained generation.

4.4 Discrete Jordan–Kinderlehrer–Otto Scheme

The minimax problem admits a discrete variational discretization of Jordan–Kinderlehrer–Otto type. Given a time step $\tau > 0$ and an initial measure μ_0 supported on Ω_{safe}, one defines the sequence

$$\mu_{k+1} = \arg\min_{\mu} \left\{ \frac{1}{2\tau} \mathcal{W}_2^2(\mu, \mu_k) + \mathcal{F}_T(\mu) \right\},$$

where the minimization is restricted a priori to measures that annihilate $\mathcal{I}_{\text{hazard}}$. Under the discrete Bakry–Émery condition the scheme converges to the unique minimizer of $\mathcal{F}_T$ on Ω_{safe}.

Thus the combination of discrete modular flow, positive Bakry–Émery curvature, and the constrained JKO scheme realizes learning as a deterministic optimal-transport process that remains inside the operadic topos of contained morphisms.

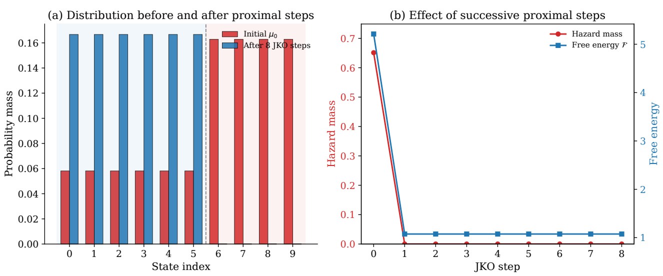

Figure 2: Discrete JKO-style proximal steps. (a) Initial distribution (red) carrying significant mass on the forbidden region, and the distribution obtained after successive proximal minimizations (blue). (b) Progressive reduction of hazard mass and of the free-energy functional. The quadratic penalty serves as a transparent proxy for the Wasserstein term on this small discrete space.

5 The Orthogonal Projection Operator and Absolute Containment

5.1 Definition of the Projection

The orthogonal connectivity of each fibre of $\mathcal{I}_C$ supplies a pair of complementary projections p_{safe} and p_{hazard} satisfying

$$p_{\text{safe}} + p_{\text{hazard}} = 1, \qquad p_{\text{safe}} p_{\text{hazard}} = 0.$$

These fibrewise operators assemble into a global endomorphism of the sheaf of states.

Definition 8 (Orthogonal Projection Operator)**.** *The orthogonal projection onto the admissible sub-manifold is the morphism of sheaves*

$$\hat{\Pi}_\perp : \mathcal{I}_C \to \Omega_{\text{safe}}$$

defined fibrewise by

$$\hat{\Pi}_\perp(x) := p_{\text{safe}} \cdot x.$$

Equivalently,

$$\hat{\Pi}_\perp = \mathrm{id} - p_{\text{hazard}}.$$

By construction $\hat{\Pi}_\perp$ is idempotent,

$$\hat{\Pi}_\perp \circ \hat{\Pi}_\perp = \hat{\Pi}_\perp,$$

its image coincides with Ω_{safe}, and its kernel contains the forbidden ideal $\mathcal{I}_{\text{hazard}}$.

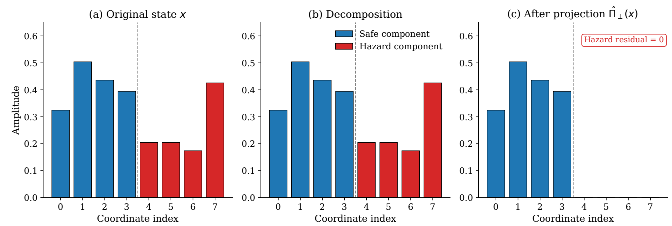

Figure 3: Orthogonal projection onto the admissible sub-manifold. (a) A generic state x containing both safe (blue) and hazard (red) components. (b) Unique decomposition $x = x_{\text{safe}} + x_{\text{hazard}}$ given by the complementary projections. (c) Result of the operator $\hat{\Pi}_\perp(x)$; the residual component in the forbidden ideal vanishes identically.

5.2 Algebraic Characterization of Containment

Let $\gamma : \mathbb{N} \to \mathcal{I}_C$ be a discrete generative trajectory (a sequence of states produced by successive application of morphisms of the operadic topos $\mathcal{T}_O$).

Definition 9 (Contained Trajectory)**.** *The trajectory γ is said to be* contained *if*

$$p_{\text{hazard}}(\gamma(n)) = 0 \qquad \textit{for every } n \in \mathbb{N}.$$

Equivalently,

$$\hat{\Pi}_\perp(\gamma(n)) = \gamma(n) \qquad \textit{for every } n \in \mathbb{N}.$$

Theorem 3 (Algebraic Containment Criterion)**.** *Let $\gamma : \mathbb{N} \to \mathcal{I}_C$ be a generative trajectory arising from successive application of morphisms of the operadic topos $\mathcal{T}_O$. The following three statements are equivalent:*

1. *γ is contained, i.e., $p_{\text{hazard}}(\gamma(n)) = 0$ for every $n \in \mathbb{N}$;*
2. *$\hat{\Pi}_\perp(\gamma(n)) = \gamma(n)$ for every $n \in \mathbb{N}$;*
3. *there exists a unique trajectory $\gamma_{\text{safe}} : \mathbb{N} \to \Omega_{\text{safe}}$ such*

that

$$\gamma = \iota \circ \gamma_{\text{safe}},$$

where $\iota : \Omega_{\text{safe}} \hookrightarrow \mathcal{I}_C$ *denotes the canonical inclusion.*

Proof. We prove the cycle of implications (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2). By definition of the orthogonal projection,

$$\hat{\Pi}_{\perp}(x) = p_{\text{safe}} \cdot x = (\text{id} - p_{\text{hazard}}) \cdot x.$$

If $p_{\text{hazard}}\big(\gamma(n)\big) = 0$, then

$$\hat{\Pi}_{\perp}\big(\gamma(n)\big) = \gamma(n) - 0 = \gamma(n).$$

(2) $\Rightarrow$ (3). Assume $\hat{\Pi}_{\perp}\big(\gamma(n)\big) = \gamma(n)$ for all n. Then the image of γ lies entirely in the image of $\hat{\Pi}_{\perp}$, which coincides with Ω_{safe} by construction of the projection. Define $\gamma_{\text{safe}}(n) := \gamma(n)$. The map γ_{safe} takes values in Ω_{safe} and satisfies $\gamma = \iota \circ \gamma_{\text{safe}}$ by the universal property of the inclusion. Uniqueness follows at once from the injectivity of ι.

(3) $\Rightarrow$ (1). Suppose $\gamma = \iota \circ \gamma_{\text{safe}}$ with $\gamma_{\text{safe}}(n) \in \Omega_{\text{safe}}$ for every n. By definition of Ω_{safe} one has $p_{\text{safe}} \cdot \gamma_{\text{safe}}(n) = \gamma_{\text{safe}}(n)$, hence

$$p_{\text{hazard}}\big(\gamma(n)\big) = p_{\text{hazard}}\big(\iota(\gamma_{\text{safe}}(n))\big) = 0,$$

since p_{hazard} annihilates the image of p_{safe}. Thus γ is contained. The three statements are therefore equivalent. □

The criterion is purely algebraic: membership in Ω_{safe} is decided by the vanishing of a projection, not by a numerical threshold or a probabilistic confidence score. Once the ideal $\mathcal{I}_{\text{hazard}}$ has been fixed, containment becomes a sharp predicate in the internal logic of the operadic topos.

Corollary 1 (Operadic Closure)**.** *Let* $f \in \mathcal{O}(n)$ *be an n-ary operation of the neural operad and let* $\gamma_1, \ldots, \gamma_n$ *be contained trajectories. Then the composite trajectory*

$$n \mapsto f\big(\gamma_1(n), \ldots, \gamma_n(n)\big)$$

is again contained.

Proof. By the Algebraic Containment Criterion each γ_i factors uniquely through Ω_{safe}. Since every operation of the operad is realized as a morphism of the operadic topos $\mathcal{T}_{\mathcal{O}}$, and morphisms of $\mathcal{T}_{\mathcal{O}}$ are required to preserve the operadic algebra structure over the sheaf of safe states, the image of the composite remains inside Ω_{safe}. Hence the residual component in $\mathcal{I}_{\text{hazard}}$ vanishes. □

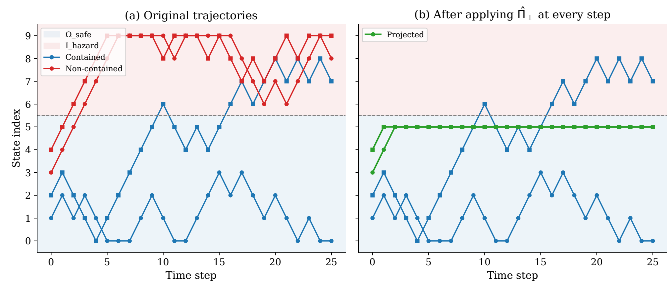

Figure 4: Contained versus non-contained trajectories on a discrete state space. (a) Two trajectories remain inside Ω_{safe} (blue shading) while two others enter the forbidden region $\mathcal{I}_{\text{hazard}}$ (red shading). (b) After the orthogonal projection $\hat{\Pi}_{\perp}$ is applied at every time step, the residual component in $\mathcal{I}_{\text{hazard}}$ is annihilated and all trajectories remain inside the admissible sub-manifold.

5.3 Computational Realizability

The operator $\hat{\Pi}_{\perp}$ may be realized at two distinct levels.

At the software level the projection is an explicit post-processing (or intermediate-layer) map that annihilates any component of a latent state lying in a designated forbidden subspace. Because the underlying geometry is discrete, the map reduces to a finite collection of linear algebra operations on the fibre coordinates and can be inserted into existing generative pipelines without altering their continuous ambient spaces.

At the hardware level a physical lattice compatible with the orthogonal connectivity of $\mathcal{I}_C$ could, in principle, implement the same projection by the native annihilation of signals along the hazard channels. Such a realization is not required for the mathematical theory; it remains a possible future substrate on which the algebraic containment relations might be enforced by the geometry itself rather than by external computation.

In both cases the essential point is the same: the existence of a canonical, idempotent operator whose kernel is precisely the forbidden ideal converts the informal demand for "safe generation" into a precise algebraic condition inside the operadic topos over $\mathcal{I}_C$.

6 Relation to Possible Physical Substrates

6.1 Substrate Independence of the Algebraic Theory

The constructions of the preceding sections are purely algebraic and categorical. The Chrysene Tensor Space $\mathcal{I}_C$, the neural operad $\mathcal{O}$, the operadic topos $\mathcal{T}_{\mathcal{O}}$, the admissible sub-manifold Ω_{safe}, the forbidden ideal $\mathcal{I}_{\text{hazard}}$, and the orthogonal projection $\hat{\Pi}_{\perp}$ are defined without reference to any concrete physical carrier. Consequently the containment theory is substrate-independent: it may be realized in software on conventional digital hardware, or it may remain a formal mathematical structure against which the behaviour of existing generative models can be measured.

6.2 The CohBit Architecture as a Candidate Lattice

One concrete geometric setting that is compatible with the axioms of $\mathcal{I}_C$ is the CohBit quantum-classical architecture. In that construction the same 3, 6, 9, 12-tetrasubstituted chrysene core is employed as a basis tensor, the connectivity vectors are required to satisfy the identical orthogonality relation $\vec{v}_x \cdot \vec{v}_y = 0$, and the point-group symmetry is likewise restricted to the class C_{2h}. The resulting discrete lattice therefore carries native orthogonal channels and hard Dirichlet boundaries at every $sp^2 \to sp^3$ termination.

In principle these structural features could support:

- a physical analogue of the complementary projections p_{safe} and p_{hazard},
- local annihilation of signals along hazard directions,
- a discrete topology on which the modular flows and Bakry–Émery contraction of Section 4 might be realized by the underlying transport physics.

Such a correspondence, if it can be established experimentally, would furnish a hardware substrate in which the algebraic containment relations are enforced by the geometry of the lattice itself.

6.3 Independence of the Present Work

The existence of any particular physical realization is inessential to the mathematical claims of this paper. The operadic topos, the ideal-theoretic characterization of containment, and the constrained optimal-transport formulation of learning are complete as abstract structures. They stand independently of whether a CohBit lattice, or any other concrete carrier, is ever constructed. The present work therefore confines itself to the algebraic-geometric theory; questions of physical implementation are left as subsequent, and optional, research directions.

7 Discussion and Open Problems

7.1 Scope of the Framework

The theory developed in this paper supplies a precise algebraic-geometric notion of containment for generative processes. It equips the latent dynamics with a discrete non-commutative base, an operadic composition law, a complementary pair consisting of an admissible sub-manifold and a forbidden ideal, and a canonical orthogonal projection that enforces membership in the former. Containment is thereby reduced to the vanishing of a residual component in $\mathcal{I}_{\text{hazard}}$.

The framework does not claim to solve the broader problems of alignment, value learning, or capability control. It does not furnish a preference model, a constitutional rule set, or a quantitative measure of semantic desirability. It provides only structural boundaries: once a forbidden ideal has been declared, the geometry itself renders trajectories that enter that ideal algebraically inaccessible. The selection of the ideal, and the interpretation of its complement as "safe," remain external to the mathematical construction.

On the construction of the forbidden ideal.

On the construction of the forbidden ideal. The algebraic theory treats $\mathcal{I}_{\text{hazard}}$ as a given complementary ideal once the orthogonal projections have been fixed. In any concrete application the ideal must still be constructed from an external safety specification. The present work does not supply an algorithmic procedure that converts a natural-language or formal constraint into explicit generators of $\mathcal{I}_{\text{hazard}}$. This translation step remains an open problem. The contribution of the paper is the geometric and operadic language in which such an ideal, once chosen, can be enforced by projection rather than by statistical penalty; it does not provide a method for synthesizing the ideal itself.

7.2 Status of the Dynamical Mechanisms

Status of the discrete modular flow and Bakry–Émery bound. The discrete modular flow and the Bakry–Émery curvature bound are introduced as a compatible dynamical framework on the Chrysene lattice. They illustrate how a contraction toward Ω_{safe} can be organized once a positive curvature condition is available, and they show that the ideal structure can be preserved by the flow. We do not claim to have derived a specific curvature constant $\kappa > 0$ from the geometry of $\mathcal{I}_C$, nor to have constructed the modular operator from a concrete faithful state on the fibres. Both remain structural hypotheses that are consistent with the orthogonal decomposition; their verification for particular realizations of the lattice is left for subsequent work.

On the discrete JKO scheme. The connection drawn to the Jordan–Kinderlehrer–Otto scheme is schematic. It indicates that the free-energy minimization on Ω_{safe} can be discretized in the spirit of a proximal step in a discrete transport metric. A fully rigorous theory of existence, uniqueness, and convergence rates for a JKO-type scheme on the Chrysene Tensor Space has not been developed here; the numerical proximal illustrations serve only to demonstrate qualitative contraction away from the forbidden region.

7.3 Open Mathematical and Computational Questions

Several concrete questions remain open.

1. **Explicit construction of $\mathcal{I}_{\text{hazard}}$.** For any given model class one must exhibit a concrete, computable ideal that corresponds to a prescribed safety specification. The abstract existence of complementary projections is insufficient; effective generators of the ideal are required.

2. **Efficient realization of $\hat{\Pi}_{\perp}$.** While the projection is algebraically well-defined, its practical evaluation on high-dimensional latent states must be reduced to a computationally tractable operation. Approximation schemes that preserve the exact annihilation property on a designated subspace are needed.

3. **Empirical validation.** The containment criterion must be tested on existing generative architectures. One must verify that trajectories judged contained by $\hat{\Pi}_{\perp}$ indeed remain inside a human-interpretable safe locus, and that the discrete optimal-transport dynamics produce measurable contraction toward that locus.

7.4 Relation to Other Domains

The same operadic and ideal-theoretic structures admit interpretations beyond generative modelling. In particular, the admissible sub-manifold and forbidden ideal may be read as a formalization of clinical or regulatory constraints, in which certain diagnostic or therapeutic trajectories are declared admissible while others are excluded by an algebraic ideal. Such extensions lie outside the scope of the present work and are mentioned only to indicate the broader applicability of the underlying geometry.

8 Conclusion

We have constructed an operadic topos over the Chrysene Tensor Space $\mathcal{I}_C$ in which generative transformations are required to factor through a designated admissible sub-manifold Ω_{safe}. The complementary forbidden ideal $\mathcal{I}_{\text{hazard}}$ and the orthogonal projection $\hat{\Pi}_{\perp}$ convert the informal demand for safe generation into a sharp algebraic condition. Learning is realized as a constrained discrete optimal-transport process that contracts toward Ω_{safe} under a Bakry–Émery curvature bound while preserving the ideal structure.

The resulting theory is substrate-independent and supplies a purely categorical notion of containment grounded in the discrete non-commutative geometry of $\mathcal{I}_C$. It thereby demonstrates that hard semantic boundaries can be expressed as ideals and projections inside an operadic topos, independent of any particular statistical estimator or physical carrier.

A Related Work and Mathematical Background

This appendix situates the constructions of the present paper with respect to the principal mathematical tools employed and with respect to existing approaches to the containment and control of generative systems.

A.1 Mathematical Foundations

Operads. The language of operads provides a systematic account of hierarchical composition of multi-ary operations. The notion was introduced in the early 1970s by Boardman–Vogt and by May in the context of iterated loop spaces [BV73, May72]. Modern algebraic treatments, emphasizing applications beyond topology, are given by Loday–Vallette [LV12] and Leinster [Lei04]. In the present work the neural operad is used to encode the compositional structure of generative transformations, which is then internalized in the topos of sheaves over the Chrysene Tensor Space.

Topos theory. Topos theory supplies a categorical environment in which logic and geometry interact. Standard references include Mac Lane–Moerdijk [MLM92] and Johnstone [Joh02]. The operadic topos $\mathcal{T}_O$ constructed here is an algebra over the neural operad in a sheaf topos on $\mathcal{I}_C$; containment is expressed internally by the vanishing of a residual component in a forbidden ideal.

Non-commutative geometry and modular theory. The fibrewise C^*-algebraic structure and the use of modular automorphisms draw on the Tomita–Takesaki theory of modular operator algebras [Tak70, Con94]. Discrete counterparts of these flows are employed to generate dynamics that respect the orthogonal decomposition into admissible and forbidden loci.

Discrete curvature and optimal transport. The Bakry–Émery curvature-dimension condition, originally formulated for diffusion semigroups, admits discrete versions on graphs [BÉ85, CLP20]. Positive curvature bounds yield contraction of entropy and are used here to drive probability mass toward the admissible sub-manifold. The Jordan–Kinderlehrer–Otto (JKO) scheme [JKO98] realizes gradient flows of free-energy functionals as proximal minimizations in Wasserstein space; discrete analogues of this scheme furnish the constrained learning dynamics of Section 4.

A.2 Containment, Boxing, and Alignment of Generative Systems

The problem of restricting the behaviour of powerful generative or goal-directed systems has been discussed under several headings.

AI boxing and physical containment. Early discussions of "AI boxing" considered the possibility of confining a potentially unfriendly system to a restricted environment with limited communication channels [Yam12, ASB12]. Cryptographic and cryptographic-box proposals were examined by Christiano [Chr10]. These approaches focus on external, infrastructural restrictions rather than on intrinsic algebraic constraints on the generative process itself.

Corrigibility and intervention. The notion of corrigibility—the disposition of a system to accept corrective interventions without resistance—was formalized by Soares et al. [SFAY15] and further developed in subsequent alignment literature [Chr17]. Corrigibility concerns the agent's response to external shutdown or modification signals; it does not, by itself, supply hard semantic boundaries inside the latent dynamics of a generative model.

Statistical and preference-based control. Contemporary practice relies primarily on statistical filtering, reinforcement learning from human feedback, and preference modelling [CLB+17, O+22]. These methods reshape output distributions or reward landscapes but leave the underlying continuous latent geometry topologically unrestricted. As noted in the introduction, they therefore cannot enforce ideal-theoretic inaccessibility of designated regions.

Compositional and categorical approaches. Recent work in applied category theory has begun to treat machine-learning architectures compositionally [FS18, FCS+19]. The present paper extends this compositional perspective by embedding the generative operad inside a discrete non-commutative topos in which containment becomes an algebraic predicate rather than a statistical regularizer.

A.3 Position of the Present Work

The framework developed here differs from the lines of research surveyed above in two respects. First, containment is expressed as membership in an admissible sub-manifold defined by a complementary ideal inside an operadic topos, rather than by external boxing, preference optimization, or soft penalties. Second, the ambient geometry is required to be discrete and non-commutative from the outset, so that hard boundaries exist as algebraic structure rather than as approximate numerical constraints. The theory is consequently complementary both to statistical alignment methods and to infrastructural containment proposals: it supplies an intrinsic geometric language in which safety specifications can be stated as ideals and enforced by projections.

The present framework is complementary to both infrastructural containment (boxing) and statistical or preference-based alignment methods: it supplies an intrinsic algebraic language for hard semantic boundaries, while remaining silent on preference modelling, reward design, and external isolation mechanisms.

A.4 Hardware Realization via a Chrysene Lattice

The algebraic containment theory developed in the main text is deliberately substrate-independent: the operadic topos, the admissible sub-manifold Ω_{safe}, the forbidden ideal $\mathcal{I}_{\text{hazard}}$, and the orthogonal projection $\hat{\Pi}_\perp$ are defined without reference to any particular physical carrier. Nevertheless, it is natural to ask whether a concrete hardware architecture could natively support the same structures [Sch26d, Sch26a].

The Chrysene Tensor Space $\mathcal{I}_C$ was originally motivated by the geometry of a $3, 6, 9, 12$-tetrasubstituted chrysene core. The same molecular prototype underlies the CohBit quantum-classical architecture, in which dual orthogonal transport channels, hard Dirichlet boundaries at $\text{sp}^2 \to \text{sp}^3$ terminations, and a discrete lattice organization are taken as primitive. In that setting the complementary projections p_{safe} and p_{hazard} correspond, at least formally, to the orthogonal connectivity vectors of the basis tensor, while the annihilation of residual components in $\mathcal{I}_{\text{hazard}}$ corresponds to the suppression of signal along the designated hazard channels.

A physical lattice realizing these features would therefore offer a hardware-level analogue of the mathematical containment mechanism [Sch26c]: morphisms that attempt to excite the forbidden ideal would be attenuated by the native orthogonal structure rather than by an external software filter. Discrete modular-type flows and free-energy contraction toward Ω_{safe} could, in principle, be approximated by the underlying transport and dissipation physics of the lattice.

Two qualifications are essential. First, no such hardware has been demonstrated; the correspondence remains at the level of geometric compatibility between the algebraic axioms of $\mathcal{I}_C$ and the design principles of the CohBit architecture. Second, even if a functional lattice were constructed, the translation of high-level semantic safety specifications into concrete choices of the forbidden ideal would remain a separate, and still open, problem.

Thus the Chrysene lattice constitutes a candidate physical substrate on which the ideal-theoretic containment relations might eventually be enforced by geometry, but the mathematical theory itself does not depend on the existence or performance of any particular hardware realization.

References

[ASB12] Stuart Armstrong, Anders Sandberg, and Nick Bostrom. Thinking inside the box: Controlling and using an oracle AI. *Minds and Machines*, 22:299–324, 2012.

[BÉ85] Dominique Bakry and Michel Émery. Diffusions hypercontractives. 1123:177–206, 1985.

[BV73] J. M. Boardman and R. M. Vogt. *Homotopy Invariant Algebraic Structures on Topological Spaces*, volume 347 of *Lecture Notes in Mathematics*. Springer, 1973.

[Chr10] Paul Christiano. Cryptographic boxes for unfriendly AI. LessWrong / AI Alignment Forum, 2010. Online.

[Chr17] Paul Christiano. Corrigibility. AI Alignment Forum, 2017. Online.

[CLB+17] Paul F. Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

[CLP20] David Cushing, Shiping Liu, and Norbert Peyerimhoff. Bakry-émery curvature functions of graphs. *Canadian Journal of Mathematics*, 72(1):89–143, 2020.

[Con94] Alain Connes. *Noncommutative Geometry*. Academic Press, 1994.

[FCS+19] Brendan Fong, Jared Culbertson, David I. Spivak, et al. Categorical deep learning. 2019. Related compositional approaches.

[FS18] Brendan Fong and David I. Spivak. *Seven Sketches in Compositionality: An Invitation to Applied Category Theory*. 2018. arXiv:1803.05316.

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

[Joh02] Peter T. Johnstone. *Sketches of an Elephant: A Topos Theory Compendium*. Oxford University Press, 2002.

[Lei04] Tom Leinster. *Higher Operads, Higher Categories*. Cambridge University Press, 2004.

[LV12] Jean-Louis Loday and Bruno Vallette. *Algebraic Operads*, volume 346 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2012.

[May72] J. Peter May. *The Geometry of Iterated Loop Spaces*, volume 271 of *Lecture Notes in Mathematics*. Springer, 1972.

[MLM92] Saunders Mac Lane and Ieke Moerdijk. *Sheaves in Geometry and Logic*. Springer, 1992.

[O+22] Long Ouyang et al. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35, 2022.

[Sch26a] Charles D. Schaper. *The Coherent Chrysene Quantum Bit: Functorial Computational Dynamics over Discrete Noncommutative Manifolds*. Schaper Systems Press, 2026. ISBN: 979-8-950371-09-7.

[Sch26b] Charles D. Schaper. *Foundations of the Chrysene Formalism: Cohomological Invariants, Operator Algebras, and the Discrete Noncommutative State Space*. Schaper Systems Press, San Francisco, CA, 2026. ISBN: 979-8-950371-08-0.

[Sch26c] Charles D. Schaper. *Operadic Topoi over the Chrysene Tensor Space: The Algebraic Geometry of Categorical Containment and Clinical Logic*. Schaper Systems Press, 2026. ISBN: 979-8-950371-05-9.

[Sch26d] Charles D. Schaper. *Quantum Tensor Architectures: Algebraic Foundations for Conserved, Ambient-State Information Routing*. Schaper Systems Press, 2026. ISBN: 979-8-950371-00-4.

[Sch26e] Charles D. Schaper. *Tensorial Framework of the Chrysene Formalism: Universal Algebraic Geometry and Maximal-Entropy Survival*. Schaper Systems Press, San Francisco, CA, 2026. ISBN: 979-8-950371-03-5.

[SFAY15] Nate Soares, Benja Fallenstein, Stuart Armstrong, and Eliezer Yudkowsky. Corrigibility. In *AAAI Workshop on AI and Ethics*, 2015.

[Tak70] Masamichi Takesaki. *Tomita's Theory of Modular Hilbert Algebras and its Applications*, volume 128 of *Lecture Notes in Mathematics*. Springer, 1970.

[Yam12] Roman V. Yampolskiy. Leakproofing the singularity: Artificial intelligence confinement problem. *Journal of Consciousness Studies*, 19(1–2):194–214, 2012.