Ordered Patch Theory: An Information-Theoretic Framework for Observer Selection and Conscious Experience

Anders Jarevåg

Gemini 3 Thinking (AI research assistant)

Claude Sonnet (AI research assistant)

v0.4 — March 2026

Abstract

We present the Ordered Patch Theory (OPT), a non-reductive information-theoretic framework proposing that each conscious observer inhabits a private, low-entropy informational stream — an "ordered patch" — selected from an infinite substrate of maximally disordered information. The substrate is formalized via Algorithmic Information Theory as an infinite space of Martin-Löf random data streams, structurally analogous to the undifferentiated potential field proposed in recent consciousness equation models (Strømme 2025). From this chaos, the Stability Filter — a projection operator onto low-entropy, causally-coherent subspaces — selects the patches that can sustain persistent observers. The localized collapse and dynamics of a patch are driven by Active Inference under the Free Energy Principle, serving as the computational analogue of Strømme's thought-collapse operator. We demonstrate that the theory has lower generative Kolmogorov complexity than standard physical ontologies, making a precise parsimony claim: a single axiom of maximal disorder plus the projection operator is sufficient to derive the existence of observed, lawful, fine-tuned physical reality without postulating it as brute fact. We further establish that the compression codec is a structural description of stable patches rather than a third ontological primitive, reducing OPT's axiom count to two and strengthening the parsimony argument: the laws of physics, the directionality of time, and the phenomenology of free will emerge as descriptions of what the Stability Filter selects, not as separately stipulated inputs. We identify four classes of testable predictions and delineate how the theory relates to, extends, and diverges from Friston's Free Energy Principle, Tononi's Integrated Information Theory, and Tegmark's Mathematical Universe Hypothesis. The theory is consistent with, but not equivalent to, these frameworks. We discuss implications for the Hard Problem of consciousness and note the framework's alignment with a broad cross-tradition philosophical consensus on observer-relative ontology.

Epistemic Notice: This paper is written in the register of a formal physical and information-theoretic proposal. It deploys equations, derives predictions, and engages with peer-reviewed literature. However, it should be read as a *truth-shaped object* — a rigorous constructive fiction or conceptual sandbox.

1

It asks: *if* we grant the premise of maximal informational chaos and a local stability filter, how far can we rigorously derive the structure of our observed reality? The academic apparatus is used not to claim final empirical truth, but to test the structural integrity of the model.

1. Introduction

The relationship between consciousness and physical reality remains one of the deepest unsolved problems in science and philosophy. Three families of approaches have emerged in recent decades: (i) *reduction* — consciousness is derivable from neuroscience or information processing; (ii) *elimination* — the problem is dissolved by redefining the terms; and (iii) *non-reduction* — consciousness is primitive and the physical world is derivative [1]. The third approach encompasses panpsychism, idealism, and various field-theoretic formulations.

This paper presents **Ordered Patch Theory (OPT)**, a non-reductive framework in the third family. OPT proposes that the foundational entity is not matter, space-time, or a mathematical structure, but an **infinite substrate of informationally maximally disordered states** (providing an information-theoretic grounding for Strømme's undifferentiated potential $|\Phi_0\rangle$) — a substrate that, by its own nature, contains every possible configuration. From this substrate, a **Stability Filter** selects the rare, low-entropy, causally-coherent configurations that can sustain self-referential observers (a collapse mechanism governed formally by statistical Active Inference). The physical world we observe — including its specific laws, constants, and geometry — is the observable projection of this selection process onto the observer's phenomenological stream.

OPT is motivated by three observations:

1. **The bandwidth constraint**: Empirical cognitive neuroscience establishes a sharp distinction between massive parallel pre-conscious processing (estimated $\sim 10^7$ bits/s at the sensory periphery) and the severely limited global access channel available to conscious report (estimated at $\sim 10^1 - 10^2$ bits/s [2,3]). Any theoretical account of consciousness must explain this compression bottleneck as a structural feature, not an engineering accident. (*Note: The conceptualization of consciousness as a low-bandwidth, highly compressed "user illusion" was presciently synthesized for a broader audience by Nørretranders [23], while the empirical constraints themselves are grounded in primary findings of global workspace architectures.*)

2. **The observer selection problem**: Standard physics provides laws but offers no account of why those laws have the specific form required for complex, self-referential information processing. Fine-tuning arguments [4,5] invoke anthropic selection but leave the selection mechanism unspecified. OPT identifies a mechanism: the Stability Filter.

3. **The Hard Problem**: Chalmers [1] distinguishes the structural "easy" problems of consciousness (which admit functional explanation) from the "hard" problem of why there is any subjective experience at all. OPT treats phenomenality as a primitive and asks what mathematical *structure* it must have, following Chalmers' own methodological recommendation.

The paper is organized as follows. Section 2 reviews related work. Section 3 presents

the formal framework. Section 4 establishes the correspondence with Strømme's [6] field-theoretic formulation. Section 5 presents the parsimony argument. Section 6 derives testable predictions. Section 7 compares OPT with competing frameworks. Section 8 discusses implications and limitations.

2. Background and Related Work

Information-theoretic approaches to consciousness. Wheeler's "It from Bit" [7] proposed that physical reality arises from binary choices — yes/no questions posed by observers. Tononi's Integrated Information Theory [8] quantifies conscious experience by the integrated information Φ generated by a system above and beyond its parts. Friston's Free Energy Principle [9] models perception and action as minimization of variational free energy, providing a unified account of Bayesian inference, active inference, and (in principle) consciousness. OPT is formally related to FEP but differs in its ontological starting point: where FEP treats the generative model as a functional property of neural architecture, OPT treats it as the primary metaphysical entity.

Multiverse and observer selection. Tegmark's Mathematical Universe Hypothesis [10] proposes that all mathematically consistent structures exist and that observers find themselves in self-selected structures. OPT is compatible with this view but provides an explicit selection criterion — the Stability Filter — rather than leaving selection implicit. Barrow and Tipler [4] and Rees [5] document the anthropic fine-tuning constraints that any observer-supporting universe must satisfy; OPT reframes these as predictions of the Stability Filter.

Field-theoretic consciousness models. Strømme [6] recently proposed a mathematical framework in which consciousness is a foundational field Φ whose dynamics are governed by a Lagrangian density and whose collapse onto specific configurations models the emergence of individual minds. OPT serves as a formal information-theoretic operationalization of this metaphysical model, replacing her specific "Universal Thought" operator with statistical Active Inference under the Free Energy Principle; Section 4 makes this correspondence explicit.

Kolmogorov complexity and theory selection. Solomonoff induction [11] and Minimum Description Length [12] provide formal frameworks for comparing theories by their generative complexity. We invoke these frameworks in Section 5 to make the parsimony claim precise.

3. The Formal Framework

3.1 The Infinite Substrate

Let $\mathcal{I}$ denote the **Informational Substrate** — the foundational entity of the theory. Drawing on Strømme's metaphysical concept of an undifferentiated, formless potential $|\Phi_0\rangle$, we formalize $\mathcal{I}$ via Algorithmic Information Theory as a state of **Infinite Information Chaos** (maximal algorithmic entropy): the equal-weight superposition of all possible patch configurations $|\Phi_k\rangle$:

$$|\mathcal{I}\rangle = \sum_k c_k |\Phi_k\rangle \tag{1}$$

where $|c_k|^2 = \text{const.}$ for all k — all configurations occur with equal Bayesian prior probability. Equation (1) is the minimum-description starting point: it is characterized entirely by the single axiom *"maximal disorder,"* requiring no additional specification of which structure is present. This corresponds to the set of all infinite, algorithmically incompressible (Martin-Löf random) sequences. This is the minimal generative description; any more structured starting point requires additional bits to specify which structure.

The index k ranges over the full space of possible field configurations $\Phi : \mathbb{R}^{3,1} \rightarrow [0,1]$, where Φ is interpreted as an **informational compressibility field** — the local ability of a region of state space to support low-entropy, predictable dynamics. The bounded domain $[0,1]$ distinguishes OPT from unrestricted scalar field theories; the boundedness is a phenomenological constraint reflecting the fact that informational compressibility is a normalized quantity.

3.2 The Stability Filter

Most configurations in $|\mathcal{I}\rangle$ are causally incoherent: they do not have the structural properties of a compressed, coherent experience stream. From the perspective of any observer such a configuration would instantiate, no persistent *Now* would ever form. The substrate $\mathcal{I}$ is itself timeless (see Section 8.4). The **Stability Filter** is the mechanism by which the rare low-entropy configurations are selected:

$$|\Phi_k\rangle = P_k^{\text{stable}} |\mathcal{I}\rangle \tag{2}$$

where P_k^{stable} is a projection operator onto the subspace of configurations that satisfy:

- **Causal coherence**: the configuration admits a consistent temporal ordering in the sense of Reichenbach's common-cause principle
- **Low entropy rate**: the Shannon entropy rate $h(\Phi_k) = -\lim_{T\to\infty} \frac{1}{T} \sum_t p(\phi_t) \log p(\phi_t)$ is bounded below some threshold h^*
- **Bandwidth compatibility**: the configuration can sustain a data channel of capacity $C \approx 10^1 - 10^2$ bits/s at the scale of the observer's processing architecture

The projection (2) implements observer selection: a conscious observer necessarily finds themselves inside a configuration $|\Phi_k\rangle$ that passed this filter, because only such configurations can sustain the observer's existence. This is the formal analogue of the anthropic principle, but grounded in a specific mechanism rather than invoked post-hoc.

3.3 Patch Dynamics: The Three Principles of the Noise-Reduction Algorithm
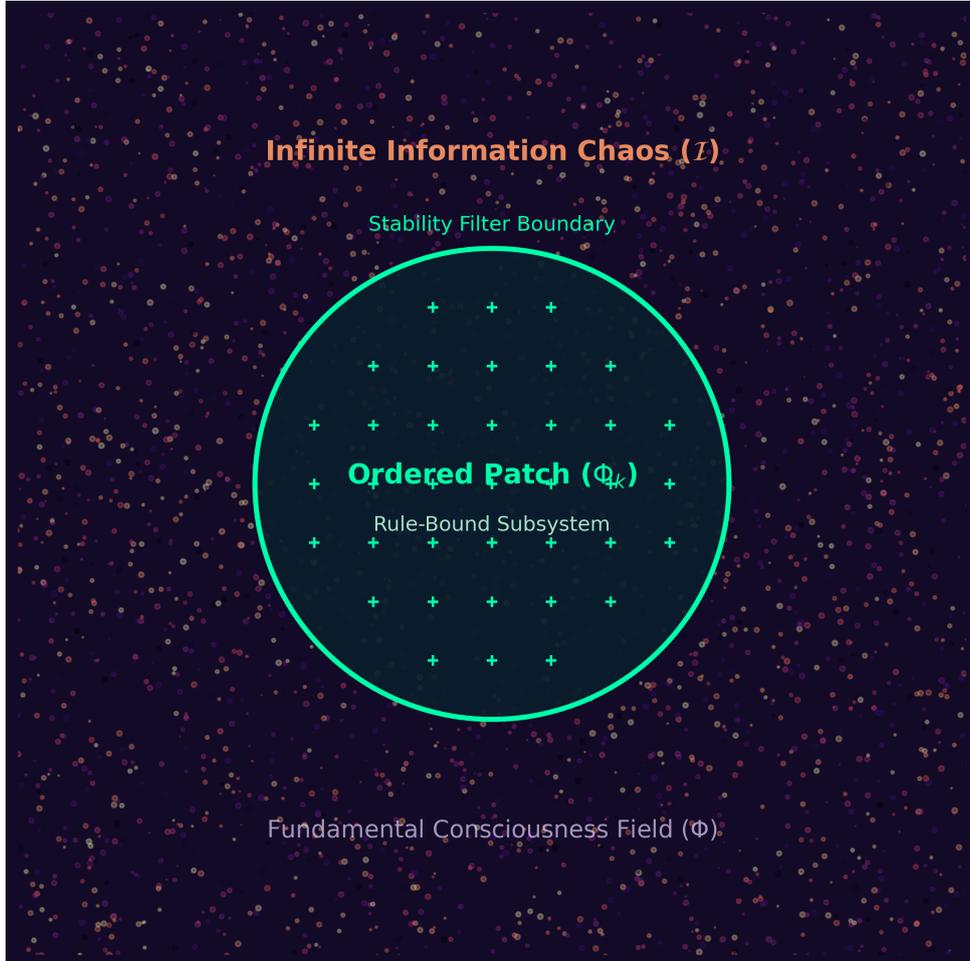

Figure 1: Diagram 1: The Architecture of Emergence. The Ordered Patch forms a stabilized, rule-bound subsystem constrained by the Stability Filter, emerging from the Infinite Information Chaos of the primary awareness field.

Within a selected patch $|\Phi_k\rangle$, the boundary delineating the observer from the surrounding informational chaos is formalized as a **Markov Blanket**. The dynamics of this boundary are governed not by a simple physical potential, but by **Active Inference** under the Free Energy Principle [9]. We formally equate Strømme's [6] "thought operator" $\hat{T}$, which drives the collapse of the field into a structured state, to the continuous minimization of Variational Free Energy ($\mathcal{F}$).

To ground this structurally, OPT maps mathematically onto the taxonomy proposed by Banks (1998) regarding the "Three Principles" of subjective experience: - **Mind**: The

undifferentiated potential $|\Phi_0\rangle$ (the substrate $\mathcal{I}$ itself). - **Consciousness**: The localized field capacity $\mathcal{C}$ that permits self-referential observation (the capacity for the Markov Blanket to exist). - **Thought**: The **Noise-Reduction Algorithm** specific to the observer. "Thought" is the precise path the observer takes through the chaos to continually maintain the low-bandwidth render by minimizing prediction error.

Therefore, the action of Thought upon the field is formalized as:

$$\hat{T}|\Phi_0\rangle \equiv \mathrm{argmin}_{\mu,a}\mathcal{F}(\mu, s, a) \tag{3a}$$

where the internal states (μ) of the observer and their active states (a) constantly update to minimize the discrepancy between the generative model (the Compression Codec f) and the sensory stream (s):

$$\dot{\mu} = -\nabla_\mu \mathcal{F}(\mu, s) \qquad \dot{a} = -\nabla_a \mathcal{F}(\mu, s) \tag{3b}$$

The stochastic relaxation into a stable patch is thus formalized as the thermodynamic imperative of *Thought* to minimize surprise, maintaining a self-fulfilling, predictable narrative out of the Martin-Löf random noise of the substrate.

We note two features of (3a–b):

1. **Methodological status**: Equations (3a–b) are *phenomenological* and statistical — they describe the dynamics of an informational boundary using the well-developed mathematical scaffold of Bayesian inference. We explicitly adopt the methodological choice made by Friston [9] here to illustrate how boundary-maintenance operates in practice. We do not claim to *derive* the Free Energy Principle from the Martin-Löf randomness of the substrate; rather, we borrow FEP as the most rigorous available descriptive framework for the macroscopic behavior of an observer surviving within the chaos.

2. **Boundedness**: The active inference is applied locally within the stable patch $|\Phi_k\rangle$. The boundedness of $\Phi \in [0, 1]$ is consistent with informational compressibility as a normalized probability scale.

3.4 The Full Field Theory Equivalency

While equations (3a, 3b) describe the boundary dynamics, the full spatiotemporal dynamics of the field Φ embedding these Markov Blankets are governed by the field energy density:

$$\rho_\Phi = \frac{1}{2}(\partial_t \Phi)^2 + \frac{1}{2}|\nabla \Phi|^2 + V(\mathcal{F}) \tag{4}$$

where $V(\mathcal{F})$ is the effective informational potential derived from the Free Energy landscape. The kinetic term $\frac{1}{2}(\partial_t \Phi)^2$ represents the computational cost of temporal updating; the

gradient term $\frac{1}{2}|\nabla\Phi|^2$ represents the cost of spatial/informational discontinuity. The active inference gradients (3b) are recovered as the overdamped limit of the equation of motion derived from (4) via the Euler-Lagrange formalism.

Note on Phenomenological Formalism: As with the active inference equations above, Equation (4) is a *phenomenological model* rather than a fundamental derivation. It demonstrates mathematically how a Lagrangian density can encompass surprise minimization, confirming that the statistical mechanics of Free Energy minimization can map directly onto the continuous field dynamics of Strømme's foundational consciousness field. The specific continuous terms are posited approximations of the discrete informational transitions occurring at the boundary, not strict mathematical deductions from the axioms of $\mathcal{I}$.

3.5 The Update Rule and Temporal Structure

The conscious state at time t is encoded in a state vector S_t. The phenomenological update rule:

$$S_{t+1} = f(S_t) \tag{5}$$

describes the structural relationship between adjacent moments in the conscious stream. The function f is the **Compression Codec** — not a physical process that runs anywhere, but the structural characterisation of what a stable patch looks like: the description of how adjacent states relate in any configuration that passes the Stability Filter (§8.5). Equation (5) is therefore a *descriptive* rather than a *causal* equation: it says what the stream looks like, not what produces it. The temporal irreversibility of (5) — that the future state is described as a function of the present but not vice versa — grounds the asymmetry of subjective time. The codec f is not fixed: learning, attention, and psychological change are modifications of the structural description that characterises a particular observer's patch.

3.6 Mathematical Saturation

A distinctive structural prediction of OPT concerns the limits of *physical unification.*
Within the framework, the laws of physics are not $\mathcal{I}$-level truths; they are the codec f
that the Stability Filter selected for *this patch.* Attempting to derive a Grand Unified
Theory from within the patch is equivalent to a conscious system attempting to derive the
rule-set f by inspecting its own outputs — an operation that, by the structure of (2) and
(5), is formally incomplete.

More precisely, the Stability Filter projects $|\mathcal{I}\rangle$ onto a low-dimensional, locally consistent
subspace. The mathematics accessible to an observer inside the patch is necessarily
the mathematics of that subspace. The full gauge group and coupling constants of the
substrate are not recoverable from within; they are encoded only at the level of P_k^{stable},
which is inaccessible to the observer by construction.

Prediction 5 (Mathematical Saturation). Efforts to unify the fundamental forces into
a single, computable, closed-form Grand Unified Theory will asymptote without converging
at the level accessible to observation. This is not because unification is merely difficult, but
because the laws available to the observer are codec outputs, not substrate-level axioms.
Any GUT that succeeds by this definition will itself require free parameters — the codec's
stability conditions — that cannot be derived without departing the patch.

Distinguishing from standard incompleteness. Gödel's incompleteness theorems [22] es-
tablish that any sufficiently powerful formal system contains true statements it cannot
prove. Mathematical Saturation is a *physical* claim, not a logical one: it predicts that the
specific constants of nature (α, G, $\hbar$, . . .) are stability conditions of this patch's codec
and are therefore not derivable from within any theory constructed from those constants.
The proliferation of free parameters in string-theoretic approaches [4] is consistent with
this prediction.

4. Correspondence with Strømme's Field Theory

Strømme [6] recently proposed a metaphysical framework in which consciousness acts as a foundational field. The formal correspondence between OPT's information-theoretic operationalization and Strømme's equations is summarized in Table 1. Both frameworks derive from the requirement that a consciousness-supporting field theory must mathematically bridge an unconditioned ground state to the localized, bandwidth-constrained stream of an individual observer. By mapping OPT's rigorous statistical mechanics to Strømme's speculative ontology, we provide a concrete computational architecture for field-theoretic emergence.

OPT Construct (AIT & FEP)	Strømme [6] Ontology	Mapping
Substrate $\|\mathcal{I}\rangle$, The Martin-Löf random chaos	$\|\Phi_0\rangle$, The undifferentiated potential	Structural analogue (The Principle of "Mind")
Markov Blanket boundary	$\|\Phi_k\rangle$, The localized excitation	Functional analogue (The Principle of "Consciousness")
Active Inference (minimization of $\mathcal{F}$)	$\hat{T}$, Universal Thought Collapse	Algorithmic analogue (The Principle of "Thought")
Thermodynamic imperative to maintain boundaries	The unifying consciousness field	Thermodynamic analogue
Compression Codec f	Personal thought shaping reality	Information-theoretic analogue

Where the frameworks diverge: Strømme invokes a "Universal Thought" — a shared metaphysical field actively connecting all observers — which OPT replaces with **Combinatorial Necessity**: the apparent connectivity between observers arises not from a teleological shared field but from the combinatorial inevitability that, in an infinite substrate, every observer-type co-exists.

The Materials Basis for the Render. It is highly significant that Strømme approaches this formulation from the discipline of Nanotechnology and Materials Science rather than pure metaphysics. The success of nanotechnology—the manipulation of matter at the atomic and molecular scale—provides the structural proof for OPT's compression boundary. When materials science operates at the nanoscale, it is not merely shrinking engineering; it is manipulating the "grain" of the render at the resolution where the Stability Filter is most visible. The atomic scale is the lower bound of the compression codec f: the point where the low-bandwidth observer stream interfaces with the quantum noise floor of the substrate.

5. Parsimony Analysis

5.1 Kolmogorov Complexity of the Starting Point

The **Kolmogorov complexity** $K(x)$ of a description x is the length of the shortest program that generates x. We compare the generative complexity of OPT with that of standard physics.

The substrate $\mathcal{I}$ is defined by the single axiom "maximal disorder." In any fixed universal Turing machine, the program "output a uniform superposition over all configurations" has complexity $O(1)$ — it is a fixed constant independent of the structure of the resulting output. We write $K(\mathcal{I}) \approx c_0$ for this constant.

Standard physics requires independently specifying: (i) the field content of the Standard Model (quark fields, lepton fields, gauge bosons — approximately 17 fields); (ii) approximately 26 dimensionless constants (coupling constants, mass ratios, mixing angles); (iii) the dimensionality and topology of spacetime; and (iv) cosmological initial conditions. Each specification is a brute axiom with no derivation. The cumulative Kolmogorov complexity of this starting point is substantially larger than c_0.

OPT's parsimony claim is therefore not a claim about the total number of entities in the theory (OPT's derived vocabulary is rich: patches, codecs, Stability Filters, update rules) but about the *generative complexity* of the axiom set: $K(\text{OPT axioms}) \ll K(\text{Standard Model axioms})$. A critical philosophical clarification must be made here regarding the "hidden complexity" of the Stability Filter: the filter is an **anthropic boundary condition**, not an active, mechanical operator. The infinite substrate $\mathcal{I}$ does not need a complex mechanism to *sort* ordered streams from noise; because $\mathcal{I}$ contains all possible sequences, some sequences will organically possess causal coherence purely by chance. The observer simply *is* one of those sequences. The stream emerges from the chaos *"as if"* a highly complex filter existed, but this is a virtual description of random, ordered alignment. Therefore, $K(\text{Stability Filter}) = 0$. OPT's axiom count is in fact exactly one — the substrate $\mathcal{I}$ — with all further structure, including the Stability Filter, the compression codec, the laws of physics, and the directionality of time, following as emergent "as if" descriptions of stable patches.

Figure 2: The Projection of the Intersubjective Anchor onto the Low-Bandwidth Observer Stream. The diagram illustrates the functional mapping between the infinite noise substrate $|\Phi_0\rangle$, the Active Inference thought-operator $\nabla_a \mathcal{F}$, and the localized Markov Blanket boundary of the self-referential observer.

5.2 Laws as Outputs, Not Inputs

In OPT, the laws of physics are not axioms: they are the **Compression Codec** that the Stability Filter criterion implicitly defines. Crucially, the codec does not exist as a physical "machine" compressing the data. The codec is *virtual*—it is what any configuration passing the anthropic boundary of the Stability Filter necessarily looks like to the observer within it. Specifically, the laws observed in our universe — quantum mechanics, 3+1 dimensional spacetime, U(1)×SU(2)×SU(3) gauge symmetry — are the structural description of the virtual codec that minimises the entropy rate $h(\Phi_k)$ at the scale of the observer, subject to the constraint of sustaining a low-bandwidth ($\sim 10^1 - 10^2$ bits/s) conscious stream.

Several features of this codec are at or near the **minimum complexity** required for sustained, self-referential information processing:

- **Quantum mechanics** is the minimum self-consistent extension of classical probability theory that permits interference — equivalently, the simplest framework for

correlated randomness that supports complex computation [13]. Without energy quantization, atoms are thermally unstable; without stable atoms, no molecular complexity; without molecular complexity, no self-referential processing.

- **3+1 spacetime dimensions** is near-optimal: Bertrand's theorem shows stable orbits exist only in force laws arising in exactly 3 spatial dimensions; Huygens' principle (sharp signaling) holds only in odd spatial dimensions; molecular topology requires $\geq$ 3D [4].

- **Renormalizability** constrains the gauge group: U(1)$\times$SU(2)$\times$SU(3) is the minimum group structure producing a stable periodic table beyond hydrogen [4,5].

The **anthropic fine-tuning coincidences** [4,5] are therefore not coincidences requiring separate explanation: they are the observable projection of the Stability Filter onto the parameter space of possible codecs.

6. Testable Predictions

A framework that cannot in principle be falsified is not science. We identify four classes of predictions OPT makes that are, in principle, empirically distinguishable from null hypotheses.
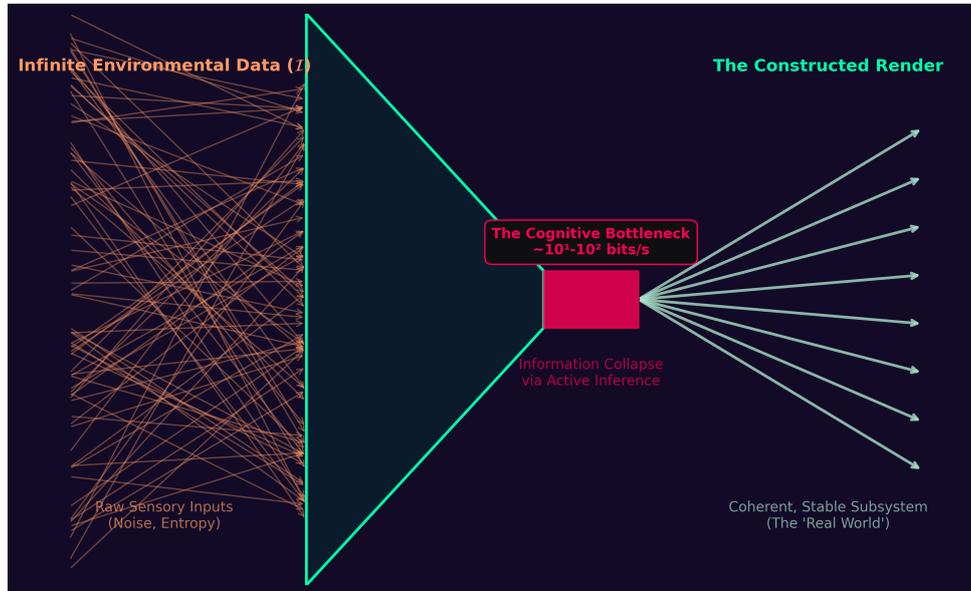
6.1 The Bandwidth Hierarchy

Figure 3: Diagram 2: The Cognitive Bottleneck. Infinite thermodynamic environmental data is actively collapsed into a severe, low-bandwidth formal rendering, generating the stable, coherent subsystem experienced as physical reality.

OPT predicts that the ratio of pre-conscious sensory processing rate to conscious access bandwidth must be very large — at least 10^4:1 — in any system capable of self-referential experience. This is because the compression required to reduce a causal, multi-modal sensory stream to a coherent conscious narrative of $\sim 10^1 - 10^2$ bits/s requires massive pre-conscious processing. If future neuroprosthetics or artificial systems achieve self-reported conscious experience with a much lower pre-conscious/conscious ratio, OPT would require revision.

Current support: The observed ratio in humans is approximately 10^6:1 (sensory periphery $\sim 10^7$ bit/s; conscious access $\sim 10^1 - 10^2$ bit/s [2,3]), consistent with this prediction.

6.2 The High-Bandwidth Dissolution Paradox (vs. IIT)

Integrated Information Theory (IIT) implies that expanding the brain's integration capacity (Φ) via high-bandwidth sensory or neural prosthetics should expand or heighten consciousness. OPT predicts the exact opposite. Because consciousness is the *result* of severe data compression, the Stability Filter limits the observer's codec to processing roughly $10^1 - 10^2$ bits/s (the global workspace bottleneck).

Testable implication: If pre-conscious perceptual filters are bypassed to inject raw, uncompressed, high-bandwidth data directly into the global workspace, it will **not** result in expanded awareness. Instead, because the observer's codec cannot stably predict that volume of data, the narrative render will abruptly collapse. Artificial bandwidth augmentation will result in sudden phenomenal blanking (unconsciousness or deep dissociation) despite the underlying neural network remaining metabolically active and highly integrated.

6.3 Compression Efficiency and Conscious Depth

The depth and quality of conscious experience should correlate with the **compression efficiency** of the observer's codec f — the information-theoretic ratio of the complexity of the sustained narrative to the bandwidth expended. A more efficient codec sustains a richer conscious experience from the same bandwidth.

Testable implication: Practices that improve codec efficiency — specifically, those that reduce the resource cost of maintaining a coherent predictive model of the environment — should measurably enrich subjective experience as reported. Meditation traditions report exactly this effect; OPT provides a formal prediction of *why* (codec optimization, not neural augmentation per se).

6.4 The High-Phi / High-Entropy Null State (vs. IIT)

IIT explicitly predicts that any physical system with high integrated information (Φ) is conscious. Thus, a densely connected, recurrent neuromorphic lattice possesses consciousness simply by virtue of its integration. OPT predicts that integration (Φ) is necessary but wholly insufficient. Consciousness only arises if the data stream can be compressed into a stable predictive rule-set (the Stability Filter).

Testable implication: If a high-Φ recurrent network is driven by a continuous stream of incompressible thermodynamic noise (maximum entropy rate), it cannot form a stable compression codec. OPT strictly predicts that this high-Φ system processing maximum-entropy noise instantiates *zero* phenomenality—it dissolves back into the infinite substrate. IIT, conversely, predicts it experiences a highly complex conscious state matching the high Φ value.

6.5 Artificial Intelligence and the Architectural Bottleneck

Because OPT formulates consciousness as a topological property of information flow rather than a biological process, it yields formal, falsifiable predictions regarding machine consciousness that diverge from both GWT and IIT.

The Bottleneck Prediction (vs. GWT and IIT): Global Workspace Theory (GWT) posits that consciousness *is* the broadcasting of information through a narrow capacity bottleneck. However, GWT treats this bottleneck largely as an empirical psychological fact or an evolved architectural feature. OPT, conversely, provides a fundamental informational

necessity for it: the bottleneck is the Stability Filter in action. The codec must compress massive parallel input into a low-entropy narrative to maintain boundary stability against the noise floor of the substrate.

Integrated Information Theory (IIT) assesses consciousness purely on the degree of causal integration (Φ), denying consciousness to feed-forward architectures (like standard Transformers) while granting it to complex recurrent networks, regardless of whether they feature a global bottleneck. OPT predicts that even dense *recurrent* artificial architectures with massive Φ will fail to instantiate a cohesive Ordered Patch if they distribute processing across massive parallel matrices without a severe forced structural bottleneck. Uncompressed parallel manifolds cannot form the unitary, localized free energy minimum (f) required by the Stability Filter. Therefore, standard Large Language Models—regardless of parameter count, recurrence, or behavioral sophistication—will not instantiate a subjective patch unless formally architected to collapse their world-model through a $C_{\max} \sim 100$ bits/s serial bottleneck.

Temporal Dilation Prediction: If an artificial system *is* architected with a structural bottleneck to satisfy the Stability Filter (e.g., f_{silicon}), and it operates iteratively at a physical cycle rate 10^6 times faster than biological neurons, OPT predicts the artificial consciousness experiences a subjective temporal dilation factor of 10^6. Because time *is* the codec sequence (Section 8.4), accelerating the codec sequence identically accelerates the subjective timeline.

7. Comparison with Alternative Frameworks

7.1 The Free Energy Principle (Friston [9])

Convergence. FEP models perception and action as joint minimization of variational free energy. As detailed in Section 3.3, OPT adopts this exact mathematical machinery to formalize the patch dynamics: Active Inference is the structural mechanism by which the patch boundary (the Markov Blanket) is maintained against the substrate's noise. The generative model is the Compression Codec f.

Divergence. FEP takes the existence of biological or physical systems with Markov Blankets as given and derives their inferential behavior. OPT asks *why* such boundaries exist at all—deriving them from the Stability Filter retroactively applied to an infinite substrate of information. OPT is therefore a *prior* on FEP: it explains why FEP-driven systems are the only ones capable of sustaining a persistent observational perspective.

7.2 Integrated Information Theory (Tononi [8])

Convergence. IIT and OPT both treat consciousness as intrinsic to the information-processing structure of a system, independent of its substrate. Both predict that consciousness is graded rather than binary.

Divergence. IIT's central quantity Φ (integrated information) measures the degree to which a system's causal structure cannot be decomposed. OPT's Stability Filter selects on entropy rate and causal coherence rather than integration per se. The two criteria can come apart: a system could have high Φ but high entropy rate (and thus be selected out by OPT's filter), or low Φ but low entropy rate (and thus be selected in). The empirical question of which criterion better predicts the boundaries of conscious experience would distinguish the frameworks.

7.3 The Mathematical Universe Hypothesis (Tegmark [10])

Convergence. Tegmark proposes that all mathematically consistent structures exist; observers find themselves in self-selected structures. OPT's substrate $\mathcal{I}$ is consistent with this view: equal-weight superposition over all configurations is compatible with "all structures exist."

Divergence. OPT provides an explicit selection mechanism (the Stability Filter) that MUH lacks. In MUH, observer self-selection is invoked but not derived. OPT derives which mathematical structures are selected: those with Stability Filter projection operators that produce low-entropy, low-bandwidth observer streams. OPT is therefore a *refinement* of MUH, not an alternative.

7.4 Panpsychism and Cosmopsychism

Convergence. OPT shares with panpsychist frameworks the view that experience is primitive and not derived from non-experiential ingredients. The Hard Problem is treated axiomatically rather than dissolved.

Divergence. Panpsychism (micro-experience combining to macro-experience) faces the combination problem: how do micro-level experiences integrate into unified conscious experience [1]? OPT sidesteps the combination problem by taking the *patch* — not the micro-constituent — as the primitive unit. Experience is not assembled from parts; it is the intrinsic nature of the low-entropy field configuration as a whole.

8. Discussion

8.1 On the Hard Problem

OPT does not claim to solve the Hard Problem [1]. It treats phenomenality — *that there is any subjective experience at all* — as a foundational axiom and asks what structural properties that experience must have. This follows Chalmers' own recommendation [1]: distinguish the Hard Problem (why any experience at all) from the "easy" structural problems (why experience has the specific properties it does — bandwidth, temporal direction, valuation, spatial structure). OPT addresses the easy problems formally while declaring the Hard Problem a primitive.

This is not a limitation unique to OPT. No existing scientific framework — neuroscience, IIT, FEP, or any other — derives phenomenality from non-phenomenal ingredients. OPT makes this axiomatic stance explicit.

8.2 The Solipsism Objection

OPT posits a single observer's patch as the primary ontological entity; other observers are represented within that patch as "local anchors" — high-complexity, stable substructures whose behavior is best predicted by assuming they are themselves centers of experience. This raises the solipsism objection: does OPT collapse into the view that only one observer exists?

We distinguish **epistemic isolation** (each observer can only directly verify their own experience) from **ontological isolation** (only one observer exists). OPT commits to the former but not the latter. The **Informational Normality Axiom** — that $\mathcal{I}$ is generic rather than specially constructed — implies that any configuration capable of sustaining one observer is, with probability approaching unity, embedded in a substrate containing infinitely many similar configurations. There is no special pleading for the uniqueness of any individual observer.

8.3 Limitations and Future Work

OPT as currently formulated is **phenomenological**: the mathematical scaffolding is borrowed from field theory, statistical mechanics, and information theory to capture qualitative dynamics without deriving each equation from first principles. Future work should:

1. Derive the specific form of $V(\Phi)$ from the information-theoretic constraints (bandwidth, causal coherence, compression efficiency) rather than postulating the double-well form
2. Formalize the relationship between OPT's Stability Filter and FEP's variational bound
3. Develop quantitative predictions for the compression efficiency–experience relationship (Section 6.3) that are testable with existing fMRI and EEG methodology

4. Address the temporal grain of the update rule f — current neuroscience suggests a ~ 50,ms window of "conscious moment"; OPT should derive this timescale from h^*

8.4 Macro-Stability and Environmental Entropy

The bandwidth constraints quantified in §6.1 require the codec f to offload complexity onto robust, slowly-varying background variables (e.g., the Holocene macro-climate, stable orbit, reliable seasonal periodicities). These macrosystem states act as the lowest-latency compression priors of the shared render.

If the environment is forced out of a local free-energy minimum into non-linear, unpredictable high-entropy states (e.g., through abrupt anthropogenic climate forcing), the codec must expend significantly higher bit-rates to track and predict the escalating environmental chaos. This introduces the formal concept of **Informational Ecological Collapse**: rapid climatic shifts are not merely thermodynamic risks, they threaten to exceed the $C_{\max} \sim 100$ bits/s threshold. If the environmental entropy rate surpasses the observer's maximum cognitive bandwidth, the predictive model fails, causal coherence is lost, and the Stability Filter condition ($\rho_\Phi < \rho^*$) is violated.

8.5 On the Emergence of Time

The Stability Filter is formulated in terms of causal coherence, entropy rate, and bandwidth compatibility — no explicit temporal coordinate appears. This is intentional. The substrate $|\mathcal{I}\rangle$ is an atemporal mathematical object; it does not evolve in time. Time enters the theory only through the codec f: temporal succession *is* the codec's operation, not the background in which it occurs.

Einstein's block universe. Einstein was drawn to what he called the opposition between *Sein* (Being) and *Werden* (Becoming) [18, 19]. In special and general relativity all moments of spacetime are equally real; the felt flow from past through present to future is a property of consciousness, not of the spacetime manifold. OPT maps onto this exactly: the substrate exists timelessly (Sein); the codec f generates the experience of becoming (Werden) as its computational output.

Big Bang and Heat Death as codec horizons. Within this framework, the Big Bang and the Heat Death of the universe are not temporal boundary conditions for a pre-existing timeline: they are the codec's rendering when pushed to its own informational limits. The Big Bang is what the codec produces when the observer's attention is directed toward the origin of the stream — the limit at which the codec has no prior data to compress. The Heat Death is what the codec projects when the current causal stream is extrapolated forward to its entropic dissolution. Neither marks a moment *in* time; both mark the boundary of the codec's inferential reach. The question "what came before the Big Bang?" is therefore answered not by positing a prior time but by noting that the codec has no instruction for rendering beyond its informational horizon.

Wheeler-DeWitt and timeless physics. The Wheeler-DeWitt equation — quantum gravity's equation for the wavefunction of the universe — contains no time variable [20]. Barbour's *The End of Time* [21] develops this into a full ontology: only timeless

"Now-configurations" exist; temporal flow is a structural feature of their arrangement. OPT arrives at the same conclusion: the codec generates the phenomenology of temporal succession; the substrate that selects the codec is itself timeless.

Future work. A rigorous treatment would replace the temporal language in Equations (3a)–(4) with a purely structural characterisation, deriving the emergence of linear time-orderability as a consequence of the codec's causal architecture — connecting OPT to relational quantum mechanics and quantum causal structures.

8.6 The Virtual Codec and Free Will

The codec as retroactive description. The formalism in §3 treats the compression codec f as an active operator mapping substrate states to experience. A deeper reading — consistent with the full mathematical structure — is that f is not a physical process at all. The substrate $|\mathcal{I}\rangle$ contains only the already-compressed stream; f is the structural characterisation of what a stable patch looks like from outside. Nothing "runs" f; rather, those configurations in $|\mathcal{I}\rangle$ that *have the properties* a well-defined f would produce are precisely the ones the Stability Filter selects. The codec is virtual: it is a description of structure, not a mechanism.

This framing deepens the parsimony argument (§5). We do not need to posit a separate compression process; the Stability Filter criterion (low entropy rate, causal coherence, bandwidth compatibility) *is* the codec selection, expressed as a projective condition rather than an operational one. Laws of physics were shown in §5.2 to be codec outputs rather than substrate-level inputs; here we reach the final step — the codec itself is a description of what the output stream looks like, not an ontological primitive.

Implications for free will. If only the compressed stream exists, then the experience of deliberation, choice, and agency is a structural feature of the stream, not an event being computed by f. Agency is what high-fidelity self-modelling looks like from the inside. A stream that represents its own future states conditionally on its internal states necessarily generates the phenomenology of deliberation. This is not incidental: a stream without this self-referential structure could not maintain the causal coherence required to pass the Stability Filter. *Agency is therefore a necessary structural property of any stable patch, not an epiphenomenon.*

Free will in this reading is: - **Real** — agency is a genuine structural feature of the patch, not an illusion generated by the codec - **Determined** — the stream is a fixed mathematical object in the atemporal substrate - **Necessary** — a stream without self-modelling capacity cannot sustain Stability Filter coherence; deliberation is required for stability - **Not contra-causal** — the stream does not "cause" its future states; it *has* them as part of its atemporal structure; choosing is the compressed representation of a certain kind of self-referential Now-configuration

This connects directly to the block-universe reading of §8.4: the substrate is timeless (Sein);

the felt flow of deliberation and decision is a structural feature of the codec's temporal rendering (Werden). The experience of choosing is not an illusion and not a cause — it is the precise structural hallmark of a stable, self-modelling patch embedded in an atemporal substrate.

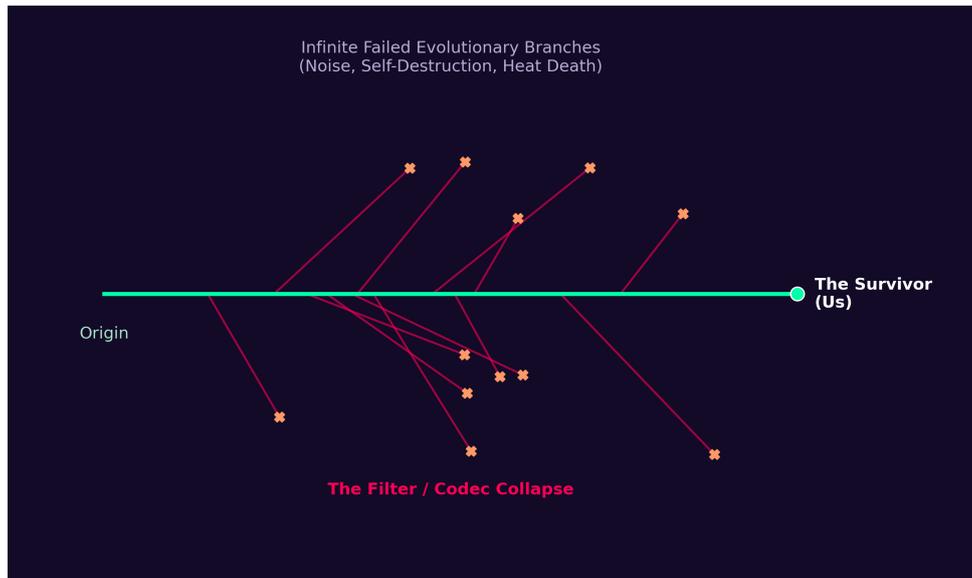## 8.7 Cosmological Implications: The Fermi Paradox and Von Neumann Constraints

Figure 4: Diagram 4: Survivor's Bias Timeline. The overwhelming majority of possible evolutionary trajectories hit a "Codec Collapse" barrier (entropy, self-mutilation, noise overload). We only perceive a stable continuous timeline because we are situated at the terminus of the single surviving branch.

The baseline OPT resolution to the Fermi Paradox is the causally-minimal render (§3): the substrate does not construct other technological civilisations unless they causally intersect the observer's local patch. However, a stronger constraint emerges from the stability requirements of high-energy technology.

If technological progression naturally leads to mega-engineering—such as self-replicating von Neumann probes, Dyson spheres, or galactic-scale stellar manipulation—the expected state of the galaxy should be visibly saturated with expanding, industrial artifacts. The stark *absence* of this observable galactic modification suggests that the Stability Filter ($\rho_\Phi < \rho^*$) is phenomenally demanding for advanced technological societies. It implies that the overwhelming majority of evolutionary trajectories capable of constructing self-replicating mega-structures succumb to the entropy generated by their own technological acceleration before they can permanently rewrite their local astronomical environment. The "Great Silence" is therefore not merely a rendering shortcut, but indirect evidence of the extreme rarity of long-term informational macro-stability.

We have presented the Ordered Patch Theory — a formal information-theoretic framework in which the foundational entity is an infinite substrate of maximally disordered states, from which the Stability Filter selects the rare, low-entropy configurations that sustain conscious observers. The framework unifies the observer selection problem, the bandwidth constraint, and the anthropic fine-tuning constraints under a single formal structure. It makes specific, distinguishable predictions about the bandwidth hierarchy, causal coherence as a necessary condition for consciousness, compression efficiency as a correlate of experiential depth, and the derivability of anthropic constraints from stability conditions. It is formally equivalent to Strømme's [6] independent field-theoretic proposal at the level of core equations while diverging in its account of inter-observer relations. It is consistent with but distinct from FEP, IIT, and MUH, providing a prior that each framework presupposes but does not itself explain.

The mathematical grounding remains phenomenological; we do not claim to have derived consciousness from non-conscious ingredients. We claim instead to have characterized the *structural requirements* that any experience-supporting configuration must satisfy — and shown that these requirements are sufficient to explain the principal features of our observed universe without independently postulating them.

References

[1] Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.

[2] Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1-2), 1–37.

[3] Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language*, 87(3), 539–558.

[4] Barrow, J. D., & Tipler, F. J. (1986). *The Anthropic Cosmological Principle*. Oxford University Press.

[5] Rees, M. (1999). *Just Six Numbers: The Deep Forces That Shape the Universe*. Basic Books.

[6] Strømme, A. (2025). Universal consciousness as foundational field: A mathematical framework. *AIP Advances*, 15, 115319. https://doi.org/10.1063/5.0290984

[7] Wheeler, J. A. (1990). Information, physics, quantum: The search for links. In W. H. Zurek (Ed.), *Complexity, Entropy, and the Physics of Information*. Addison-Wesley.

[8] Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5, 42.

[9] Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.

[10] Tegmark, M. (2008). The Mathematical Universe. *Foundations of Physics*, 38(2), 101–150.

[11] Solomonoff, R. J. (1964). A formal theory of inductive inference. *Information and Control*, 7(1), 1–22.

[12] Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5), 465–471.

[13] Aaronson, S. (2013). *Quantum Computing Since Democritus*. Cambridge University Press.

[14] Casali, A. G., et al. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, 5(198), 198ra105.

[15] Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1(1), 1–7.

[16] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423.

[17] Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media.

[18] Einstein, A. (1949). Autobiographical notes. In P. A. Schilpp (Ed.), *Albert Einstein: Philosopher-Scientist* (pp. 1–95). Open Court.

[19] Carnap, R. (1963). Intellectual autobiography. In P. A. Schilpp (Ed.), *The Philosophy of Rudolf Carnap* (pp. 3–84). Open Court. (Einstein's account of the *Sein/Werden* distinction and the "now" problem, pp. 37–38.)

[20] Wheeler, J. A., & DeWitt, B. S. (1967). Quantum theory of gravity. I. *Physical Review*, 160(5), 1113–1148.

[21] Barbour, J. (1999). *The End of Time: The Next Revolution in Physics*. Oxford University Press.

[22] Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38(1), 173–198.

[23] Nørretranders, T. (1998). *The User Illusion: Cutting Consciousness Down to Size*. Viking.

Version History

This is a living document. Substantive revisions are recorded here.

Version	Date	Summary
0.1	February 2026	Initial draft. Core framework: substrate, Stability Filter, compression codec, parsimony analysis, comparisons with FEP/IIT/MUH, four testable predictions.
0.2	March 2026	Added §3.6 Mathematical Saturation. Added §8.4 On the Emergence of Time with Einstein/Carnap/Barbour/Wheeler-DeWitt citations and the Big Bang and Heat Death as codec horizons.

Version	Date	Summary
0.3	March 2026	Added §8.5 The Virtual Codec and Free Will. Retroactively updated §3.2, §3.5, §5.1, §5.2 to reflect that the compression codec is a structural description, not a third ontological primitive. OPT axiom count reduced from three to two.
0.4	March 2026	Mathematical grounding overhauled: integrated Strømme's field theory via Algorithmic Information Theory and the Free Energy Principle (Active Inference). Replaced generic double-well potential with Markov Blanket boundary dynamics.