Contents

The Singularity-Bounded Holographic Class 4 Automaton: A Cosmological Model from Consciousness Theory

Matthias Gruber

Independent researcher

ORCID: 0009-0005-9697-1665

Correspondence: matthias@matthiasgruber.com

Abstract

I propose a cosmological model — the Singularity-Bounded Holographic Class 4 Automaton (SB-HC4A) — derived from the convergence of four independently motivated frameworks: a five-class computational taxonomy that refines Wolfram's (2002) classification by separating fractal from random dynamics, the Four-Model Theory of consciousness (FMT; Gruber, 2015, 2026a) which identifies self-referential simulation at criticality as the mechanism of phenomenal experience, a formal model-space framework for continuous modeling density (Gruber, 2026b), and 't Hooft's (1993, 2016) holographic automaton interpretation of quantum mechanics. The model proceeds by elimination: Classes 1–3 cannot sustain the universal computation the universe demonstrably supports; Class 5 (genuine randomness) makes physics fundamentally impossible; therefore the universe operates at Class 4 — the edge of chaos. Combined with the information-theoretic observation that singularities at every physical scale (Planck regime, particle interiors, event horizons, cosmological horizons, temporal endpoints) share the property of information impermeability and Bekenstein saturation, the model proposes that these singularities are structurally identical — scale-invariant instances of the same information boundary. The resulting architecture is a self-referential holographic Class 4 automaton bounded at every scale by singularity surfaces, where the observable interior is the "simulation" and the singularity boundary is the "substrate." This architecture is structurally identical to the FMT consciousness architecture, where implicit models (substrate) are separated from explicit models (simulation) by an information-opaque boundary. Consciousness is thus a local, scale-reduced instance of the same computational pattern the universe implements globally. Five weak points are identified, including the fundamental epistemological objection that Class 4 observers may be constitutionally incapable of determining whether this model describes the universe or merely the ceiling of their own computational capacity.

Keywords: cosmology, cellular automata, holographic principle, consciousness, criticality, singularity, computational complexity, edge of chaos, self-referential closure

---

1. Introduction

1.1 The Problem

Two questions haunt the foundations of physics and philosophy respectively: *Why does the universe exist?* and *Why is there consciousness?* These are typically treated as separate problems — the first belonging to cosmology and fundamental physics, the second to neuroscience and philosophy of mind. This paper argues they are the same problem at different scales.

The argument proceeds not from analogy but from structural identity. If both the universe and consciousness can be shown to instantiate the same computational architecture — with the same formal properties, the same boundary conditions, and the same self-referential closure — then the two questions collapse into one: *Why does this architecture exist?* And the answer, I will argue, is that it is the unique self-consistent configuration for a system that computes its own existence.

1.2 Sources and Scope

The model presented here draws on four bodies of work:

1. **The Five-Class Computational Taxonomy** (Gruber, 2015), which refines Wolfram's (2002) four-class classification of dynamical systems by splitting his Class 3 ("random") into genuine fractals (the new Class 3) and genuine randomness (the new Class 5). This produces a clean monotonic gradient from order to disorder and reveals that Class 4 — the edge of chaos — is the maximum complexity achievable by expressible rules.

2. **The Four-Model Theory of Consciousness** (FMT; Gruber, 2015, 2026a), which proposes that consciousness consists of real-time self-simulation across four nested models — Implicit World Model (IWM), Implicit Self Model (ISM), Explicit World Model (EWM), and Explicit Self Model (ESM) — operating at criticality. The central claim: qualia are virtual properties of the simulation, not properties of the substrate. The implicit-explicit boundary is information-opaque from the simulation side.

3. **The FMT Formalization** (Gruber, 2026b), which replaces the discrete four-model taxonomy with a continuous model-space framework and establishes self-referential closure as a fixed point of self-representation: $\Phi(m) = m$.

4. **The Holographic Automaton Interpretation** ('t Hooft, 1993, 2016), which proposes that quantum mechanics emerges from deterministic dynamics at the Planck scale — that the universe is fundamentally a cellular automaton whose holographic structure produces quantum behavior as an emergent phenomenon.

The model is speculative. It is not a proof. It is a logical argument chain that, if the premises hold, yields a specific cosmological architecture. The paper presents the argument, identifies where it could break, and proposes what would count as evidence for or against it.

1.3 Acknowledgment

The importance of symmetries in physical theory — the methodological principle that drove this work — was impressed upon me by my uncle, Bruno J. Gruber (1936–2020), a theoretical physicist who spent his career working on symmetry groups in quantum mechanics (Gruber, 1968, 1980). The observation that the same computational architecture might appear at cosmological and neurological scales is, at bottom, a symmetry claim. Whatever merit it has owes much to his influence.

2. The Five-Class Framework

2.1 Wolfram's Classification and Its Problem

Wolfram (2002) classified the behavior of cellular automata into four classes based on their long-term dynamics:

Wolfram Class	Behavior	Example
1	Uniform (converges to fixed state)	Rule 0
2	Periodic (settles into repeating loops)	Rule 4
3	Random/chaotic (apparently random)	Rule 30
4	Complex (persistent interacting structures)	Rule 110

This classification was genuinely useful and applied far beyond cellular automata — to fluid dynamics, biological systems, economic models, and neural networks. But Wolfram's Class 3 was a grab-bag containing two structurally different phenomena: fractal systems like Rule 90 (which generates a Sierpinski triangle — computationally *reducible*, meaning you can calculate any cell without running the full simulation) and apparently chaotic systems like Rule 30 (which produces output that *looks* random but is completely deterministic and computationally *irreducible*).

Rowland (2006) independently argued that nested (fractal) patterns deserved separate classification. I argue the distinction goes deeper: fractal and genuinely random systems differ in a way that matters for cosmology.

2.2 The Five Classes

The refined classification, ordered as a monotonic gradient from most ordered to most disordered:

Class 1 — Static. Systems that converge to a fixed state. Period: 1. No computation.

Class 2 — Periodic. Systems that settle into repeating loops. Information is stored but never transformed. Period: finite.

Class 3 — Fractal. Systems that produce self-similar structure at every scale. Computationally *reducible* — you can skip ahead without running every step. Structure without processing power. Period: quasi-infinite with exact or statistical self-similarity.

Class 4 — Complex (edge of chaos). Systems that produce persistent localized structures capable of arbitrary computation. Computationally *irreducible* — no shortcuts. These systems support universal computation: given the right initial conditions, they can simulate any algorithm, including themselves. This is where consciousness lives.

Class 5 — Random. Systems whose output is genuinely random — maximal Kolmogorov complexity, incompressible, non-algorithmic. The generating process exceeds what formal symbolic systems can express.

2.3 Why Deterministic Rules Cannot Produce Randomness

A cellular automaton has a finite rule table and a finite initial condition — together, a fixed finite amount of information. A truly random infinite sequence has maximal Kolmogorov complexity and cannot be compressed to anything shorter than itself. Therefore, no deterministic automaton can produce genuinely random output: the output's complexity is bounded by the rule-set's complexity (Kolmogorov, 1965; Chaitin, 1966).

This is a generalized pigeonhole argument. The only way to generate infinite output from finite information is to reuse structure at different scales. Exact reuse is periodicity (Class 2). Patterned reuse is fractal behavior (Class 3). Even the most complex-looking automata — Rule 30, Rule 110, Conway's Game of Life — produce output whose complexity is bounded by their rule-set.

What Wolfram called "random" automata are better described as *high-complexity fractals*: systems whose self-similar structure operates at scales and in dimensions that make it invisible to casual inspection.

2.4 Class 4 as the Expressibility Ceiling

Classes 1 through 4 are what finite, expressible rules can produce. Class 5 requires rules that cannot be written down — if the rule were expressible, the output would be compressible (to: "apply this rule"), and therefore not truly random.

Class 4 is therefore the *maximum complexity achievable by expressible rules*. It is as complex as mathematics can get. Beyond it lies territory that formal systems, by their own nature, cannot map.

This observation — that expressible computation has a ceiling, and that ceiling is Class 4 — is the first premise of the cosmological argument.

2.5 Class 4 Contains All Lower Classes

A critical property: Class 4 automata can generate Class 1 behavior (stable states), Class 2 behavior (oscillations), and Class 3 behavior (fractal patterns) as subprocesses within their own dynamics. No lower class can do this. The Game of Life, for instance, contains still lifes (Class 1), blinkers (Class 2), and self-similar growth patterns (Class 3) as embedded phenomena within its Class 4 dynamics.

This containment property means that a Class 4 universe does not merely permit lower-class phenomena — it *generates* them. Static matter, periodic orbits, fractal coastlines, and self-similar galaxy distributions are all Class 4 subprocesses.

---

3. The Universe Must Be Class 4

3.1 The Ontological Starting Point

Pure nothingness cannot exist as a state of affairs. "Nothing" is a Platonic abstraction — like a perfect circle or a truly periodic sequence, it is a concept with no physical instantiation. This is not a novel observation; it appears in Leibniz ("Why is there something rather than nothing?"), in Heidegger's reformulation of the fundamental question of metaphysics, and in contemporary discussions by Krauss (2012) and Albert (2012). I accept it as Axiom 1: something exists.

Whatever exists must have some dynamical character — if it had none, it would be indistinguishable from nothing by the Identity of Indiscernibles (Leibniz, 1686). Therefore, whatever exists has dynamics classifiable within the five-class hierarchy.

3.2 The Elimination Argument

Classes 1 and 2 (static, periodic): The universe demonstrably contains consciousness, which requires Class 4 dynamics (Gruber, 2015, 2026a; supported independently by the empirical criticality program: Beggs & Plenz, 2003; Shew & Plenz, 2013; Algom & Shriki, 2026). A Class 1 or 2 universe cannot generate Class 4 subprocesses (lower classes cannot produce higher-class behavior). Eliminated.

Class 3 (fractal): A fractal universe would have rich structure but would be computationally reducible — no universal computation, no consciousness. The universe demonstrably supports universal computation (we build Turing machines; consciousness exists). Eliminated.

Class 5 (random): If the universe's fundamental dynamics are genuinely random — rules inexpressible in any formal system — then physics is not merely incomplete but *fundamentally impossible*. The project of writing down laws that predict observations would be an illusion. Our local experience of lawful behavior would be a temporary, coincidental pocket of apparent regularity within genuine randomness — an infinitesimally improbable fluctuation with no mechanism guaranteeing its continued existence. Not logically impossible, but explanatorily empty and deeply unsatisfying.

Class 4 (complex, edge of chaos): The universe operates at the maximum expressible complexity. It supports universal computation (demonstrated by the existence of consciousness and Turing machines). It self-maintains criticality — Class 4 dynamics are self-organizing in complex systems (Bak, Tang, & Wiesenfeld, 1987; Bak, 1996). It contains all lower classes as subprocesses (the universe demonstrably contains static matter, periodic phenomena, and fractal structure). It is the *only* class that does all of these simultaneously.

Conclusion (Proposition 1): The universe operates at Class 4 — the edge of chaos. This is the most parsimonious classification consistent with observation.

3.3 The Strength of the Elimination

The elimination argument is not deductive proof. Two of the four eliminations are empirical (Classes 1–2: the universe contains consciousness; Class 3: the universe supports universal computation), and one is abductive (Class 5: physics would be fundamentally impossible — not a logical contradiction, but an explanatory catastrophe). Only the affirmative case for Class 4 combines empirical evidence with a theoretical mechanism (self-organized criticality).

The argument is strongest read as: *Class 4 is the unique class consistent with all observations and the only class that provides a self-maintaining mechanism for its own persistence.*

---

4. Quasi-Infinity and Information Horizons

4.1 The Speed Limit and Its Consequences

Information cannot travel faster than c (Einstein, 1905). The universe is expanding, and the expansion is accelerating (Riess et al., 1998; Perlmutter et al., 1999). Together, these facts create a

fundamental horizon: for any observer, there exists a maximum distance in space and maximum duration in time from which information can reach them. Beyond this horizon, the universe is informationally inaccessible.

This makes the universe **quasi-infinite** rather than truly infinite. From within, it appears unbounded. But information can never traverse more than approximately one universe-diameter in any direction (spatial or temporal). This is not merely an observational limitation — it is an ontological boundary, because information that can never interact with anything on this side of the boundary is, by Leibniz's Identity of Indiscernibles, indistinguishable from nonexistence.

4.2 The Concept of Quasi-Infinity

I define **quasi-infinity** as the condition of a system that is effectively unbounded for any internal observer but has a finite information content accessible from any given point. Quasi-infinity is distinct from both mathematical infinity (no bound exists) and finitude (a bound is reachable). A quasi-infinite system has a bound that recedes faster than any observer can approach it.

The universe is quasi-infinite in space (the observable universe is finite; expansion pushes the boundary outward faster than light), in time (the Big Bang is a finite time in the past; heat death or Big Crunch is in the finite or infinite future; but neither endpoint is informationally accessible), and — I will argue — in scale.

4.3 Scale as a Quasi-Infinite Dimension

The universe presents singularities not only at the boundaries of space and time but at the boundaries of scale:

- At the smallest scale ($\sim 10^{-35}$ m, the Planck regime), spacetime itself loses physical meaning. No measurement below this scale is possible, even in principle (Planck, 1899; Wheeler, 1957).
- At the largest scale, the observable universe boundary is set by the expansion history and the speed of light. Beyond it: informationally inaccessible.

Between these extremes, the universe spans approximately 60 orders of magnitude in length scale. Both endpoints are singularities — regions where our physical description breaks down and information transfer becomes impossible. The universe is quasi-infinite in scale in the same sense it is quasi-infinite in space and time: effectively unbounded from within, but bounded by information horizons at both extremes.

---

5. Singularities as Information Boundaries

5.1 The Singularity Inventory

At every physical scale, the universe presents singularities — regions where physical description breaks down and information transfer ceases:

Scale	Singularity	Information Property
Planck ($\sim 10^{-35}$ m)	Planck regime	Below this, spacetime loses meaning. No measurement possible.

Scale	Singularity	Information Property
Subatomic (~10^{-15} m)	"Point-like" particles	Treated as zero-dimensional; actually Planck-sized. Interiors inaccessible.
Stellar/galactic	Black hole event horizons	Information cannot escape. Interior causally disconnected from exterior.
Cosmological (space)	Observable universe boundary	Expansion + c creates impenetrable information horizon.
Cosmological (time, past)	Big Bang	All world-lines converge. No "before" is accessible.
Cosmological (time, future)	Heat death / Big Crunch	Infinite dilution or reconvergence — either way, a terminus of accessible information.

5.2 The Unification Claim

These singularities share a structural property: they are all **information-impermeable boundaries** — surfaces across which no signal passes. I propose they are not merely analogous but *structurally identical*: instances of the same information boundary at different scales.

The argument proceeds in three steps:

Step 1: Bekenstein saturation. The Bekenstein bound (Bekenstein, 1981) establishes that the maximum information content of a region is proportional to its surface area, not its volume. Black hole event horizons saturate this bound — they contain the maximum information density per unit area. The holographic principle ('t Hooft, 1993; Susskind, 1995) generalizes this: the information content of any region is encoded on its boundary.

Step 2: Scale invariance. If the universe is a Class 4 system (Proposition 1), its dynamics are self-similar — structure at one scale recurs at other scales. Class 4 systems contain Class 3 (fractal) behavior as a subprocess, and fractal behavior is defined by scale invariance. The boundary structure of a Class 4 system should itself be scale-invariant.

Step 3: Structural identity. All singularities in the inventory (Section 5.1) share the same three properties: (a) information impermeability, (b) Bekenstein saturation (or its scale-appropriate equivalent), and (c) they bound the computational domain — they define the limits of what can be computed from within. The claim is that these are not six different phenomena that happen to share properties, but one phenomenon — the boundary of the automaton's computational domain — at six different scales.

5.3 Particles as Planck-Scale Singularities

A specific prediction follows: "point-like" elementary particles are not truly zero-dimensional. They are Planck-scale singularities — miniature information boundaries whose interiors are as inaccessible as a black hole's. The standard model treats particles as mathematical points for calculational

convenience, but the SB-HC4A model predicts they have Planck-scale structure that saturates the Bekenstein bound at that scale.

This is consistent with approaches in quantum gravity where the Planck scale provides a natural minimum length (Garay, 1995; Hossenfelder, 2013), though the specific claim that particles *are* singularities of the same type as event horizons is novel.

---

6. The Holographic Class 4 Architecture

6.1 Three Relationships Between Holograms and Automata

In the book manuscript that preceded this paper (Gruber, 2015), I identified three possible relationships between holographic systems and Class 4 cellular automata:

1. **A holographic substrate produces Class 4 dynamics.** This is what the brain does: locally holographic neural networks (Lashley, 1950; Pribram, 1971) operating at criticality produce Class 4 cortical dynamics.

2. **A Class 4 automaton produces holographic output.** Local rules at the edge of chaos generate non-local, distributed information encoding as emergent behavior. This describes quantum entanglement from an information-theoretic perspective.

3. **A Class 4 automaton whose rule structure is itself holographic.** The rules encode higher-dimensional information in lower-dimensional structure. If such a system exists, it does what the holographic principle says the universe does — not by analogy, but by construction.

I proposed (Gruber, 2015) that if all three relationships could coexist in a single system — holographic rules, Class 4 dynamics, holographic output — the result would be a computational fixed point: a system that encodes itself.

6.2 The SB-HC4A Architecture

The Singularity-Bounded Holographic Class 4 Automaton (SB-HC4A) is the system in which all three relationships hold simultaneously, bounded by the singularity structure identified in Section 5.

Definition. An SB-HC4A is a dynamical system U = (S, R, B) where: - S is a state space on a d-dimensional manifold - R is a holographic rule set: R encodes (d+1)-dimensional information in d-dimensional structure - B is a singularity boundary: a scale-invariant surface of maximum information density (Bekenstein-saturated) that bounds the computational domain at every scale - U operates at Class 4 dynamics (branching ratio $\sigma \sim 1$, maximum Lyapunov exponent $\lambda\_\mathrm{max} \sim 0$) - The output of U is holographic: emergent large-scale structure encodes non-local information

6.3 Self-Referential Closure

The SB-HC4A is self-referential: it computes its own structure. This can be stated as a fixed-point condition:

$$\Phi(U) = U$$

where Φ is the "compute the output" operator. The holographic rules (R) encode the full system in compressed form. The Class 4 dynamics decompress this encoding into the observable universe. The holographic output re-encodes the result. The computation and the system are the same thing.

This self-referential closure has a precise formal analogue in the FMT formalization (Gruber, 2026b, Section 6.3), where consciousness is identified with the fixed point of self-representation: $\Phi(m) = m$, where m* is the ESM state at which the model and the modeled coincide.

6.4 Inexpressibility

A consequence of self-referential closure: the SB-HC4A cannot be fully specified by any formal system that is a proper subsystem of itself. By Godel's incompleteness theorems (Godel, 1931), a sufficiently complex self-referential system contains true statements that cannot be proven from within. The holographic rules are compressed — and the decompressed universe IS the computation. No sub-universe entity can write down a description smaller than the universe itself.

The "Weltformel" (world equation) is therefore not an equation. It is a *process* — the automaton itself. It can only be expressed by running it.

---

7. The Consciousness-Cosmology Symmetry

7.1 The Structural Mapping

The SB-HC4A architecture maps onto the FMT consciousness architecture with exact structural correspondence:

SB-HC4A (Universe)	FMT (Consciousness)
Singularity boundary (information-opaque)	Implicit-explicit boundary (phenomenally opaque)
Observable interior (physics)	Explicit models (conscious simulation)
Holographic rule structure	Distributed/holographic implicit knowledge
Class 4 dynamics	Cortical criticality
Self-referential closure: $\Phi(U) = U$	Self-referential closure: $\Phi(m) = m$
Conservation of energy/information across boundary	Conservation of information across implicit-explicit split
Inexpressibility from within (Godel)	Meta-Problem: consciousness cannot explain its own substrate
Singularity boundary at every scale	Implicit-explicit boundary at every level of the model hierarchy

7.2 Not Analogy but Structural Identity

This mapping is not metaphorical. Both systems implement the same formal architecture:

1. **Both are Class 4 dynamical systems** operating at the edge of chaos. The universe maintains criticality through self-organized criticality (Bak et al., 1987). The cortex maintains criticality through homeostatic regulation of excitation-inhibition balance (Hengen et al., 2016; Ma et al., 2019).

2. **Both are bounded by information-opaque surfaces.** The universe's singularity boundaries prevent information transfer (event horizons, Planck regime). The brain's implicit-explicit boundary prevents phenomenal access to substrate operations (Gruber, 2026a, Section 3.6).

3. **Both have holographic structure.** The universe's information content is encoded on boundaries ('t Hooft, 1993; Susskind, 1995). The brain's information storage is holographic — distributed across the substrate, degrading gracefully under damage (Lashley, 1950; Pribram, 1971).

4. **Both exhibit self-referential closure.** The universe computes its own structure (the laws of physics are the universe's dynamics applied to itself). Consciousness models its own modeling (the ESM represents itself).

5. **Both are inexpressible from within.** The universe's complete specification exceeds any internal formal system (Godel). Consciousness cannot fully model its own substrate (the Meta-Problem; Chalmers, 2018; Gruber, 2026a, Section 3.7).

7.3 Consciousness as a Local Instance

The claim is not that consciousness *causes* the universe or that the universe is "conscious" in any experiential sense. The claim is that consciousness and the universe instantiate the *same computational pattern* at different scales:

- The **universe** is a Class 4 holographic automaton bounded by singularities, where the interior (observable physics) is the "simulation" and the boundary (singularity layer) is the "substrate."

- **Consciousness** is a Class 4 holographic automaton (cortical dynamics) bounded by the implicit-explicit boundary, where the explicit models are the "simulation" and the implicit models are the "substrate."

Same architecture. Different scale. The pattern is fractal — which is precisely what a Class 4 system predicts, since Class 4 dynamics contain Class 3 (self-similar) behavior as a subprocess.

---

8. Energy, Information, and Conservation

8.1 The Energy-Information Hypothesis

Landauer's principle (Landauer, 1961) establishes a minimum energy cost for erasing one bit of information: kT ln 2. The Bekenstein bound (Bekenstein, 1981) sets the maximum information content of a region proportional to its boundary surface area — a statement about information density with units of energy. Black hole thermodynamics (Bekenstein, 1973; Hawking, 1975) assigns entropy, temperature, and information content to black holes through their surface area.

These results converge on a hypothesis: energy and information are not merely correlated but identical — two descriptions of the same quantity. This is not proven, and I flag it as a weak point (Section 9.1). But if E = I, the conservation laws become information conservation laws, and the SB-HC4A architecture acquires a natural conservation principle.

8.2 Conservation Across the Singularity Boundary

If energy/information is conserved, singularities do not destroy information — they *transform* it. This is the resolution of the black hole information paradox that modern physics is converging on (Almheiri et al., 2020; Penington, 2020; Raju, 2022). Information that enters a black hole is not lost; it is encoded on the event horizon (holographic encoding) and eventually re-emitted through Hawking radiation.

In the SB-HC4A model, this generalizes: the singularity boundary at every scale conserves total information while transforming it between compressed (boundary) and decompressed (interior) forms:

- **Compressed form** (on the boundary): Maximum information density, Bekenstein-saturated, inaccessible from within. This is the "substrate."
- **Decompressed form** (in the interior): Lower density, organized, accessible. This is the "simulation" — the observable universe, or the conscious experience.

8.3 The FMT Parallel

This is precisely the implicit/explicit split in consciousness:

- **Implicit models** (substrate): Hold all the information — synaptic weights, structural knowledge, the full learned model of the world and self. Maximum information density. Not phenomenally accessible.
- **Explicit models** (simulation): A lower-bandwidth, organized, phenomenal projection of selected information. Accessible. This is experience.

The permeability function in FMT (Gruber, 2026a, Section 3.6; Gruber, 2026b, Section 3) determines how much information transfers from implicit to explicit — how much of the compressed substrate becomes part of the conscious simulation. The laws of physics play the analogous role at the cosmological scale: they determine what information from the singularity-layer substrate enters the observable interior.

---

9. Where This Could Break

A theory that claims no weaknesses is not a theory. Here are the five places where the SB-HC4A model is most vulnerable.

9.1 Weak Point 1: Energy = Information Is Not Proven

The energy-information identity is strongly suggested by Landauer's principle, the Bekenstein bound, and the holographic principle, but it is not established as a fundamental law. Landauer's principle has been experimentally confirmed (Berut et al., 2012), but only for erasure — it does not prove that all energy is information. If energy and information are merely correlated rather than identical, the conservation argument (Section 8) weakens, and the mapping between singularity boundaries and information boundaries becomes looser.

What would resolve it: A derivation of energy conservation from information conservation (or vice versa) within a well-defined physical framework, or an experimental demonstration that information has gravitational effects proportional to its energy-equivalent.

9.2 Weak Point 2: Class 4 Universe Is an Empirical Claim, Not a Theorem

Self-organized criticality is well-documented in sandpiles (Bak et al., 1987), neural networks (Beggs & Plenz, 2003), earthquakes (Gutenberg & Richter, 1956), and other systems. But these are all *subsystems within* the universe. Whether the universe *itself* — at its most fundamental level — operates at Class 4 is a much stronger claim. The universe contains Class 4 subsystems; this does

not entail that the universe is itself Class 4. A Class 5 universe could contain Class 4 pockets, just as a Class 4 automaton contains Class 2 substructures.

The elimination argument (Section 3.2) is the strongest available response: Class 4 is the only class consistent with all observations. But the elimination of Class 5 rests on abduction (physics would be impossible), not deduction.

What would resolve it: Evidence that the universe's large-scale dynamics exhibit criticality signatures — power-law distributions in the cosmic microwave background, scale-free structure in galaxy distributions, or branching ratio analysis of cosmological dynamics. Some of this evidence exists (e.g., the scale-free distribution of galaxy clusters), but it has not been framed as a criticality argument.

9.3 Weak Point 3: Singularity Unification Is Speculative

Whether Planck-scale structure, black hole event horizons, and cosmological horizons are truly "the same thing" at different scales requires a theory of quantum gravity to verify. Loop quantum gravity (Rovelli, 2004) and string theory (Polchinski, 1998) both suggest structures consistent with this claim — in loop quantum gravity, spacetime is discrete at the Planck scale with area quantization that connects to black hole entropy; in string theory, the holographic principle is central (Maldacena, 1998). But neither theory confirms the specific claim that all singularities are instances of the same information boundary.

What would resolve it: A derivation, within a quantum gravity framework, showing that the information-theoretic properties of Planck-scale boundaries and macroscopic event horizons are formally identical (same entropy scaling, same information capacity per unit area, same causal disconnection properties).

9.4 Weak Point 4: Testability Is Limited

The model predicts:

(a) **Singularity universality**: All singularities across scales share informational properties. The resolution of the black hole information paradox should confirm that event horizons conserve information — and the same conservation principle should apply at the Planck scale and the cosmological horizon. Partially testable through black hole information research.

(b) **Cosmological criticality**: The universe's large-scale dynamics are at the edge of chaos. In principle testable through statistical analysis of the CMB, large-scale structure, or the distribution of galaxy cluster sizes.

(c) **Consciousness as local instance**: Consciousness implements the same architecture as the universe. Testable through the FMT predictions (Gruber, 2026a, Section 8): nine specific predictions about psychedelic phenomenology, anesthetic mechanisms, and neural criticality.

But the core claim — that the universe *is* an SB-HC4A — may be unfalsifiable from within, for exactly the Godelian reason the model itself predicts (Section 6.4). The system cannot be fully specified by any subsystem. This is either a devastating weakness or a structural prediction of the model, depending on one's philosophical commitments.

9.5 Weak Point 5: The Cognitive Ceiling Problem

This is the deepest objection, and it was identified during the model's initial formulation.

If we are Class 4 automatons operating at criticality, then the SB-HC4A model may simply be the most complex concept our Class 4 brains can produce. We cannot think in Class 5. We cannot conceive of structures beyond our own computational class. The pattern we see — Class 4 everywhere, self-similar at every scale, holographic and self-referential — might be the *signature of our own cognitive architecture projected onto the cosmos*, not a feature of the cosmos itself.

This objection connects to the evolutionary biology of cognition. The human brain evolved under selection pressure for symmetry detection — faces of predators and prey are the most symmetric, and therefore most survival-relevant, patterns in a hunter-gatherer's environment. We are, at the deepest level, symmetry-detection machines. The SB-HC4A model is fundamentally a symmetry claim (the same architecture at every scale). We might find this symmetry not because it exists in the universe but because our brains are optimized to find symmetry wherever they look.

This is structurally identical to the Meta-Problem at cosmic scale: the ESM cannot see its own substrate, so it cannot distinguish between "the universe has this structure" and "my brain can only model the universe with this structure." The cosmological model predicts its own potential unfalsifiability — which is either the strongest possible confirmation of the consciousness-cosmology symmetry (the model predicts this exact epistemological limitation) or the strongest possible objection to it (the model is an artifact of the observer, not a feature of the observed).

A Class 4 system can simulate anything up to and including Class 4 complexity. But it cannot verify whether the universe exceeds that. If the universe is Class 5 but locally appears Class 4 to Class 4 observers (because Class 4 is the maximum pattern we can detect), we would construct exactly this model — and be wrong.

I do not know how to resolve this objection from within. I am not sure it can be resolved from within. I include it because intellectual honesty requires it.

---

10. The Necessity Argument

10.1 Why This Architecture Must Exist

Notwithstanding the weak points above, the SB-HC4A is arguably the unique structure consistent with all axioms simultaneously:

Axiom 1 (Ontological Necessity): Something exists. Nothing is not a possible state of affairs.

Axiom 2 (Computational Character): Whatever exists has dynamical behavior classifiable within the five-class hierarchy.

Axiom 3 (Criticality Stability): Class 4 is the unique computational class that (a) supports universal computation, (b) self-maintains through self-organized criticality, and (c) contains Classes 1–3 as subprocesses.

Axiom 4 (Information Bound): Information propagation speed is bounded by c. Information density is bounded by the Bekenstein bound. Information is conserved.

Axiom 5 (Holographic Encoding): The boundary of a d-dimensional region encodes all information in the region's interior on its (d–1)-dimensional surface.

From these:

- A1 + A2 → Something with dynamical behavior exists.
- A3 → The most stable interesting configuration is Class 4 (the only self-maintaining, universal class).
- A4 → The system is bounded by information horizons at every scale (singularity structure).
- A5 → The boundaries encode the interior (holographic architecture).
- A3 + A5 → The system is a holographic Class 4 automaton.
- A4 + A5 → The boundaries are Bekenstein-saturated and scale-invariant.
- Self-referential closure: A holographic Class 4 system with holographic output is a fixed point: $\Phi(U) = U$.

10.2 Uniqueness

The SB-HC4A is the unique structure satisfying all five axioms. Remove any axiom and the uniqueness fails:

- Without A1: Nothing is possible — no universe required.
- Without A2: The existent thing need not have computational character — mysterian position.
- Without A3: Any class is permissible — no prediction about the universe's dynamics.
- Without A4: No information horizons — no singularity structure — no boundaries.
- Without A5: No holographic encoding — no boundary-interior relationship.

The argument is that these axioms are not arbitrary assumptions but well-supported physical and logical principles, and that their conjunction yields a unique cosmological architecture.

---

11. Discussion

11.1 Relationship to Existing Cosmological Proposals

The SB-HC4A model intersects with several existing research programs:

't Hooft's Cellular Automaton Interpretation ('t Hooft, 1993, 2016): The model is a direct extension of 't Hooft's proposal that quantum mechanics emerges from deterministic automaton dynamics at the Planck scale. The SB-HC4A adds the Class 4 classification (specifying which class of automaton), the singularity boundary structure, and the consciousness-cosmology symmetry.

The Holographic Principle ('t Hooft, 1993; Susskind, 1995; Bousso, 2002): The model presupposes the holographic principle and extends it by proposing that holographic encoding is not just a property of boundaries but of the automaton's rule structure itself (Relationship 3 of Section 6.1).

Self-Organized Criticality (Bak et al., 1987; Bak, 1996): The model relies on SOC as the mechanism by which the universe maintains Class 4 dynamics. The extension is the claim that this self-organization operates at the cosmological level, not just in subsystems.

Digital Physics (Fredkin, 2003; Zuse, 1969; Wolfram, 2002): The model shares the premise that the universe is fundamentally computational but adds specific structure: the five-class hierarchy, the holographic property, the singularity boundary, and the consciousness parallel.

Wheeler's "It from Bit" (Wheeler, 1990): The model is consistent with Wheeler's proposal that information is ontologically fundamental. The E = I hypothesis (Section 8.1) is a strong version of "it from bit."

11.2 Relationship to FMT

The SB-HC4A model does not require the FMT to be correct about consciousness. The cosmological argument (Sections 2–6) stands independently. However, if both the SB-HC4A and FMT are correct, then:

1. Consciousness is a local instance of a universal pattern, not a cosmic accident.
2. The emergence of consciousness in a Class 4 universe is not merely possible but structurally guaranteed — any sufficiently complex Class 4 subsystem with self-referential closure will instantiate the pattern.
3. The Meta-Problem of consciousness (Chalmers, 2018) is the local version of the cosmological inexpressibility problem — both arise because self-referential systems cannot fully specify themselves from within.

11.3 What This Is Not

This paper does not claim:

- That the universe is "conscious" in any experiential sense. The structural identity is architectural, not phenomenal. A building's blueprint is not a building.
- That consciousness creates reality (idealism). The model is physicalist: the substrate is physical; the simulation is a physical process; the correspondence is structural.
- That quantum mechanics is "explained by" consciousness (Penrose, Stapp). The model makes no such claim. Quantum mechanics is an emergent property of the automaton's dynamics — consciousness is also emergent, at a higher scale.
- That this is proven. It is an argument. It could be wrong (Section 9).

---

12. Conclusion

The SB-HC4A model proposes that the universe is a Class 4 holographic automaton bounded at every scale by singularity surfaces — information-impermeable boundaries of maximum density — where the observable interior is a decompressed projection of the boundary-encoded information, and the system is self-referentially closed: it computes its own structure.

This architecture is structurally identical to the architecture of consciousness as described by the Four-Model Theory: a self-referential simulation at criticality, bounded by an information-opaque boundary, with the simulation as the experienced world and the substrate as the informationally inaccessible foundation.

The model rests on five axioms (ontological necessity, computational character, criticality stability, information bounds, holographic encoding), proceeds by elimination (the universe must be Class 4), and yields a unique self-consistent architecture. Five specific weak points have been identified, the deepest being the cognitive ceiling problem: we may find this symmetry because our Class 4 brains are constitutionally incapable of seeing anything else.

Whether the SB-HC4A is a description of the universe or a description of the limits of human cognition is, I believe, the most important open question in the philosophy of science. The model predicts that this question cannot be answered from within — and that prediction is either the model's deepest confirmation or its deepest flaw.

References

Albert, D. Z. (2012). On the origin of everything. *The New York Times*. [Review of Krauss, 2012.]

Albantakis, L., et al. (2023). Integrated information theory (IIT) 4.0. *PLOS Computational Biology*, 19(10), e1011465.

Algom, S., & Shriki, O. (2026). The ConCrit framework: Critical brain dynamics as a unifying mechanism for consciousness theories. *Neuroscience & Biobehavioral Reviews*.

Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., & Tajdini, A. (2020). The entropy of Hawking radiation. *Reviews of Modern Physics*, 93(3), 035002.

Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. Springer.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, 59(4), 381–384.

Beggs, J. M., & Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *Journal of Neuroscience*, 23(35), 11167–11177.

Bekenstein, J. D. (1973). Black holes and entropy. *Physical Review D*, 7(8), 2333–2346.

Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D*, 23(2), 287–298.

Berut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., & Lutz, E. (2012). Experimental verification of Landauer's principle linking information and thermodynamics. *Nature*, 483, 187–189.

Bousso, R. (2002). The holographic principle. *Reviews of Modern Physics*, 74(3), 825–874.

Chaitin, G. J. (1966). On the length of programs for computing finite binary sequences. *Journal of the ACM*, 13(4), 547–569.

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

Chalmers, D. J. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies*, 25(9-10), 6–61.

Einstein, A. (1905). Zur Elektrodynamik bewegter Korper. *Annalen der Physik*, 322(10), 891–921.

Fredkin, E. (2003). An introduction to digital philosophy. *International Journal of Theoretical Physics*, 42(2), 189–247.

Garay, L. J. (1995). Quantum gravity and minimum length. *International Journal of Modern Physics A*, 10(2), 145–165.

Godel, K. (1931). Uber formal unentscheidbare Satze der Principia Mathematica und verwandter Systeme I. *Monatshefte fur Mathematik und Physik*, 38, 173–198.

Gruber, B. J. (1968). *Topics in Mathematical Physics*. Gordon and Breach.

Gruber, B. J. (1980). Symmetries in science. In *Symmetries in Science* (pp. 1–18). Springer.

Gruber, M. (2015). *Die Emergenz des Bewusstseins*. Self-published.

Gruber, M. (2026a). The Four-Model Theory of Consciousness: A simulation-based framework unifying the Hard Problem, binding, and altered states. *Zenodo* preprint. https://doi.org/10.5281/zenodo.18669891

Gruber, M. (2026b). Toward a mathematical formalization of the Four-Model Theory: A recommended approach. Manuscript.

Gutenberg, B., & Richter, C. F. (1956). Magnitude and energy of earthquakes. *Annali di Geofisica*, 9, 1–15.

Hawking, S. W. (1975). Particle creation by black holes. *Communications in Mathematical Physics*, 43, 199–220.

Hengen, K. B., Torrado Pacheco, A., McGregor, J. N., Van Hooser, S. D., & Bhatt, D. H. (2016). Neuronal firing rate homeostasis is inhibited by sleep and promoted by wake. *Cell*, 165(1), 180–191.

Hossenfelder, S. (2013). Minimal length scale scenarios for quantum gravity. *Living Reviews in Relativity*, 16, 2.

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

Krauss, L. M. (2012). *A Universe from Nothing*. Free Press.

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 183–191.

Lashley, K. S. (1950). In search of the engram. *Symposia of the Society for Experimental Biology*, 4, 454–482.

Leibniz, G. W. (1686). *Discourse on Metaphysics*.

Ma, Z., Turrigiano, G. G., Bhatt, D. H., & Bhatt, W. B. (2019). Cortical circuit dynamics are homeostatically tuned to criticality in vivo. *Neuron*, 104(4), 655–664.

Maldacena, J. (1998). The large-N limit of superconformal field theories and supergravity. *Advances in Theoretical and Mathematical Physics*, 2(2), 231–252.

Penington, G. (2020). Entanglement wedge reconstruction and the information problem. *Journal of High Energy Physics*, 2020, 2.

Perlmutter, S., et al. (1999). Measurements of Ω and Λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2), 565–586.

Planck, M. (1899). Uber irreversible Strahlungsvorgange. *Sitzungsberichte der Koniglich Preussischen Akademie der Wissenschaften zu Berlin*, 5, 440–480.

Polchinski, J. (1998). *String Theory* (Vols. 1–2). Cambridge University Press.

Pribram, K. H. (1971). *Languages of the Brain.* Prentice-Hall.

Raju, S. (2022). Lessons from the information paradox. *Physics Reports*, 943, 1–80.

Riess, A. G., et al. (1998). Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3), 1009–1038.

Rovelli, C. (2004). *Quantum Gravity.* Cambridge University Press.

Rowland, E. (2006). Wolfram's classification and its extensions. *NKS Conference Proceedings.*

Shew, W. L., & Plenz, D. (2013). The functional benefits of criticality in the cortex. *The Neuroscientist*, 19(1), 88–100.

Susskind, L. (1995). The world as a hologram. *Journal of Mathematical Physics*, 36(11), 6377–6396.

't Hooft, G. (1993). Dimensional reduction in quantum gravity. In *Salamfestschrift* (pp. 284–296). World Scientific.

't Hooft, G. (2016). *The Cellular Automaton Interpretation of Quantum Mechanics.* Springer.

Wheeler, J. A. (1957). On the nature of quantum geometrodynamics. *Annals of Physics*, 2(6), 604–614.

Wheeler, J. A. (1990). Information, physics, quantum: The search for links. In *Complexity, Entropy, and the Physics of Information* (pp. 3–28). Addison-Wesley.

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

Zuse, K. (1969). *Rechnender Raum.* Friedrich Vieweg & Sohn.