BEDS: Bayesian Emergent Dissipative Structures

A Formal Framework for Sustainable Digital Twins and Continual Learning Systems

---

Laurent Caraffa

Univ. Gustave Eiffel, IGN-ENSG, LaSTIG
French National Institute of Geographic and Forest Information
Ministry of Ecological Transition, France
Email: laurent.caraffa@ign.fr

Abstract

We present BEDS (Bayesian Emergent Dissipative Structures), a theoretical framework that unifies concepts from non-equilibrium thermodynamics, Bayesian inference, information geometry, and machine learning. The central thesis proposes that learning, across physical, biological, and computational systems, fundamentally constitutes the conversion of flux into structure through entropy export. Building on Prigogine's theory of dissipative structures, we establish a formal isomorphism between thermodynamic processes and Bayesian updating, demonstrating that sustainable learning systems must follow dissipative patterns where crystallized posteriors become priors for subsequent levels of emergence.

We derive fundamental mathematical constants (e, π, φ) as fixed points of Bayesian inference under minimal axioms, suggesting these constants emerge necessarily from any system capable of representing and updating uncertainty. Furthermore, we propose a conjecture linking Gödel's incompleteness theorems to thermodynamic constraints, hypothesizing that pathologies of formal systems (incompleteness, undecidability) are structurally analogous to dissipation deficits in physical systems.

As practical validation, we present a peer-to-peer network architecture implementing BEDS principles, achieving six orders of magnitude improvement in energy efficiency compared to existing distributed consensus systems while enabling continuous learning. This work bridges fundamental physics, mathematical logic, and practical system design, offering both theoretical insights into the nature of learning and computation, and a concrete pathway toward sustainable artificial intelligence.

Keywords: Dissipative structures, Bayesian inference, Information geometry, Free energy principle, Sustainable AI, Digital twins, P2P networks, Thermodynamics of computation, Gödel incompleteness, Mathematical constants emergence

---

1. Introduction

1.1 The Sustainability Challenge in Artificial Intelligence

Contemporary artificial intelligence faces a fundamental tension between capability and sustainability. Large language models and deep neural networks achieve remarkable performance but at extraordinary computational and energetic cost. Training GPT-4 class models requires an estimated 1.3 GWh of electricity [1], while global AI inference energy consumption is projected to reach 85-134 TWh annually by 2027 [2]. This trajectory is thermodynamically unsustainable.

More fundamentally, current AI architectures exhibit structural limitations that transcend mere efficiency concerns:

1. **Temporal drift**: Models trained on data from time T become increasingly miscalibrated as the world evolves to T+1, T+2, etc. [3, 4].
2. **Out-of-distribution fragility**: Frozen models extrapolate unpredictably in regions beyond their training distribution [5, 6].
3. **Energy scaling**: Inference cost scales with model parameters, creating tension between capability and deployment [7].

These limitations suggest that the dominant paradigm of supervised learning—where humans label data, models memorize patterns, and inference consumes energy without updating—may be fundamentally misaligned with the requirements of sustainable, adaptive intelligence.

1.2 Learning from Physical Systems

Nature has solved the problem of sustainable, adaptive learning over billions of years. Rivers find paths to the sea. Cells maintain homeostasis. Ecosystems self-organize. Brains learn continuously without catastrophic forgetting. What unifies these systems?

Prigogine's Nobel Prize-winning work on dissipative structures [8, 9] provides a crucial insight: open systems far from equilibrium can spontaneously generate and maintain order by exporting entropy to their environment. A hurricane maintains its structure by dissipating heat. A living cell preserves its organization by exporting waste. These systems are genuinely self-supervised—they learn and adapt through interaction with flux, without external labels or supervisory signals.

This observation motivates a fundamental question: *Can we formalize the principles by which dissipative structures learn, and apply them to design sustainable artificial learning systems?*

1.3 Contributions and Structure

This paper presents BEDS (Bayesian Emergent Dissipative Structures), a theoretical framework proposing that:

1. **Learning is thermodynamic**: Across physical, biological, and computational systems, learning constitutes the conversion of flux into structure through entropy export.
2. **Bayesian inference provides the minimal formalism**: The mathematical structure of Bayesian updating captures the essential dynamics of dissipative learning, with deep connections to information geometry and thermodynamics.
3. **Mathematical constants emerge from inference**: Under minimal axioms, fundamental constants (e, π, φ) appear as fixed points of Bayesian inference—not arbitrary definitions but necessary outcomes of uncertainty representation.

4. **Formal system pathologies have thermodynamic analogues**: Gödel incompleteness and related phenomena may be structurally analogous to dissipation deficits in closed systems.
5. **Sustainable AI is achievable**: A peer-to-peer architecture implementing BEDS principles achieves orders-of-magnitude efficiency gains over existing distributed systems.

The paper is structured as follows: Section 2 reviews the theoretical foundations spanning thermodynamics, information theory, and Bayesian inference. Section 3 presents the core BEDS formalism. Section 4 derives the emergence of mathematical constants. Section 5 develops the Gödel-Landauer-Prigogine conjecture. Section 6 describes the P2P implementation. Section 7 discusses implications and limitations. Section 8 concludes.

---

2. Theoretical Foundations and Related Work

2.1 Non-Equilibrium Thermodynamics and Dissipative Structures

The second law of thermodynamics states that entropy in isolated systems increases monotonically. However, Prigogine and colleagues demonstrated that open systems exchanging energy and matter with their environment can spontaneously develop ordered structures [8, 9, 10]. These *dissipative structures* maintain organization by importing low-entropy energy (negentropy) and exporting high-entropy waste.

Formally, for an open system:

$$\frac{dS_{system}}{dt} = \frac{dS_{internal}}{dt} + \frac{dS_{exchange}}{dt}$$

where $dS_{internal}/dt \geq 0$ (irreversible entropy production) and $dS_{exchange}/dt$ can be negative (entropy export). When $|dS_{exchange}/dt| > dS_{internal}/dt$, the system can maintain or increase internal order.

Classic examples include Bénard convection cells, the Belousov-Zhabotinsky reaction, and biological organisms. Critically, these structures exist only while flux persists—cut the energy flow, and order dissipates.

Nicolis and Prigogine [10] established conditions for structure emergence: systems must be (1) open to flux, (2) far from equilibrium, and (3) governed by nonlinear dynamics. Under these conditions, spontaneous symmetry breaking can produce macroscopic order from microscopic fluctuations.

2.2 Information Theory and the Physics of Computation

Shannon's information theory [11] established the mathematical foundation for quantifying uncertainty through entropy:

$$H(X) = -\sum_i p(x_i) \log p(x_i)$$

Jaynes [12, 13] demonstrated the deep connection between information-theoretic and thermodynamic entropy, showing that statistical mechanical distributions emerge from maximum entropy principles subject to constraints.

Landauer [14] proved that information erasure has irreducible thermodynamic cost: erasing one bit requires at least $k_B T \ln 2$ joules of energy dissipation. This establishes that computation is inherently physical—information processing cannot be divorced from thermodynamics.

Bennett [15, 16] extended this work, showing that logically reversible computations can in principle be thermodynamically reversible, but irreversible logical operations (AND, OR, ERASE) necessarily dissipate heat. This implies fundamental energy bounds on computation.

Wheeler's "It from Bit" hypothesis [17] proposes that physical reality emerges from information-theoretic primitives. While speculative, this perspective motivates treating information processing and physical dynamics within a unified framework.

2.3 Bayesian Inference and Information Geometry

Bayesian inference provides a principled framework for updating beliefs in light of evidence:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

Cox [18] and Jaynes [12] demonstrated that Bayesian updating is the unique consistent method for belief revision under minimal rationality axioms.

The space of probability distributions forms a Riemannian manifold with the Fisher information metric [19, 20]:

$$g_{ij}(\theta) = \mathbb{E}_{p(x|\theta)}\left[\frac{\partial \log p(x|\theta)}{\partial \theta_i}\frac{\partial \log p(x|\theta)}{\partial \theta_j}\right]$$

Amari [21, 22] developed information geometry, showing that this manifold has unique properties including dually flat structure under exponential and mixture coordinate systems. The natural gradient:

$$\tilde{\nabla} L(\theta) = G(\theta)^{-1} \nabla L(\theta)$$

follows geodesics on this manifold, providing optimal learning trajectories.

Chentsov [23] proved that the Fisher metric is the unique Riemannian metric on probability simplex invariant under sufficient statistics—a uniqueness result suggesting deep structural significance.

2.4 Free Energy Principle and Active Inference

Friston's Free Energy Principle (FEP) [24, 25, 26] proposes that biological systems minimize variational free energy:

$$F = \mathbb{E}_q[\log q(\theta) - \log p(\theta, D)] = D_{KL}(q(\theta) \| p(\theta|D)) - \log p(D)$$

This provides a unified account of perception (minimizing prediction error), action (changing the world to match predictions), and learning (updating generative models).

The FEP connects to thermodynamics through the interpretation of free energy as a bound on surprisal (negative log-evidence). Systems that persist must occupy low-surprisal states—those compatible with continued existence—implying that living systems effectively perform approximate Bayesian inference.

Active inference extends this to action selection, where agents act to minimize expected free energy, balancing exploitation (achieving preferred outcomes) and exploration (resolving uncertainty) [27].

2.5 Energy-Based Models in Machine Learning

Energy-Based Models (EBMs) [28, 29] define probability distributions through energy functions:

$$P(x) = \frac{1}{Z} \mathsf{exp}(-E(x)/T)$$

This Boltzmann distribution connects machine learning directly to statistical mechanics. Learning adjusts energy landscapes; inference finds low-energy configurations.

Hopfield networks [30], Boltzmann machines [31], and modern EBMs [32] implement variants of this principle. LeCun's JEPA architecture [33] proposes joint embedding prediction in latent space, avoiding the intractable normalization of generative models.

Hinton's work on Boltzmann machines [31, 34] established connections between neural network learning and statistical mechanics, including the wake-sleep algorithm that alternates between data-driven and model-driven phases.

2.6 Distributed Systems and Consensus

Distributed consensus—agreement among multiple agents without central authority—is fundamental to blockchain, distributed databases, and multi-agent systems.

Bitcoin's Proof-of-Work [35] achieves Byzantine fault tolerance through computational puzzles, consuming approximately 120-150 TWh annually [36]. Proof-of-Stake alternatives [37] reduce energy consumption by 99.9% but maintain global consensus requirements.

Federated learning [38] distributes model training across devices while aggregating updates centrally. However, gradient aggregation scales with model size, and central coordination remains a bottleneck.

Gossip protocols [39] and belief propagation [40] provide distributed inference mechanisms where nodes exchange local information, achieving global consistency through local interactions.

2.7 Mathematical Foundations and Logic

Gödel's incompleteness theorems [41] establish that any consistent formal system capable of expressing arithmetic contains true statements unprovable within the system. This fundamental limitation arises from self-reference—the ability of formal systems to encode statements about themselves.

Chaitin's algorithmic information theory [42, 43] connects incompleteness to randomness, showing that certain mathematical facts are "random" in that they cannot be derived from simpler axioms.

Tarski's undefinability theorem [44] shows that truth predicates cannot be defined within the language they apply to, necessitating hierarchies of metalanguages.

These results collectively demonstrate that closed formal systems face fundamental limitations—themes we develop in the Gödel-Landauer-Prigogine conjecture.

---

3. The BEDS Framework

3.1 Core Formalism

We define a **Bayesian Emergent Dissipative Structure** (BEDS) as an open system characterized by:

Definition 3.1 (BEDS). A system S is a BEDS if it satisfies: 1. **Flux input**: S receives continuous input Φ_in from its environment 2. **Bayesian state**: S maintains a probability distribution q(θ) over parameters θ 3. **Entropy export**: S exports entropy H_out to its environment 4. **Crystallization**: When q(θ) converges (variance below threshold), S crystallizes

The fundamental dynamics are:

$$q_{t+1}(\theta) = \frac{p(D_t|\theta)q_t(\theta)}{\int p(D_t|\theta')q_t(\theta')d\theta'}$$

where D_t is observation at time t. The system persists while:

$$\Phi_{in} > \Phi_{min}$$

and crystallizes when:

$$\mathrm{Var}[q(\theta)] < \epsilon$$

3.2 Thermodynamic-Bayesian Correspondence

We propose a formal isomorphism between thermodynamic and Bayesian quantities:

Thermodynamics	**Bayesian Inference**
Internal energy E	Negative log-likelihood −log p(D|θ)
Entropy S	Entropy of q(θ)
Temperature T	Inverse regularization β^{-1}
Free energy F	Negative ELBO
Thermal equilibrium	Optimal posterior
Thermal fluctuations	Epistemic uncertainty
Heat dissipation	KL divergence reduction

Thermodynamics	Bayesian Inference
Phase transition	Posterior collapse/multimodality

This correspondence allows bidirectional transfer of insights: thermodynamic intuitions inform inference algorithms, while Bayesian formalism provides computational tools for analyzing dissipative dynamics.

Theorem 3.1 (Free Energy Equivalence). *For exponential family distributions with natural parameters η, the thermodynamic free energy F = E − TS and variational free energy F_var = E_q[−log p(D|θ)] + D_KL(q||p) are related by:*

$$F_{var} = \beta F + \text{const}$$

where β = 1/T is inverse temperature.

Proof sketch: For exponential families, the log-partition function A(η) satisfies dA/dη = E[T(x)], connecting moment parameters to natural parameters. The variational free energy decomposes into energy and entropy terms matching the thermodynamic decomposition under the identification T ↔ β^{-1}. □

3.3 The Recursive Emergence Principle

The key innovation of BEDS is recursive structure: when a dissipative structure crystallizes, its stable configuration becomes the prior for the next level of emergence.

Definition 3.2 (BEDS Hierarchy). A BEDS hierarchy is a sequence {S_n}*{n≥0} where: - S_0 operates on raw flux Φ - For n > 0, S_n receives crystallized output from S*{n-1} as its prior:

$$\pi_n(\theta) = q^*_{n-1}(\theta)$$

where q*_{n-1} is the crystallized posterior of level n-1

This recursive structure has profound implications:

Theorem 3.2 (Entropy Decrease). *In a BEDS hierarchy, the entropy of crystallized structures decreases with level:*

$$H(q^*_n) < H(q^*_{n-1})$$

Proof: Each level crystallizes when posterior variance drops below threshold, implying entropy reduction. The posterior of level n-1 becomes the prior of level n, so level n starts with constrained possibilities and crystallizes to even lower entropy. □

Corollary 3.1 (Energy Bound). *The total energy of a BEDS hierarchy is bounded:*

$$E_{total} = \sum_{n=0}^{\infty} E_n < \infty$$

Proof: Energy at each level is proportional to entropy (by the thermodynamic-Bayesian correspondence). Since entropy decreases geometrically, the series converges. □

This provides a formal basis for sustainable learning: systems following the BEDS pattern cannot grow unboundedly in energy consumption.

3.4 BEDS as Emergent Energy-Based Model

Real BEDS hierarchies involve multiple interacting agents at each level, not single structures. This leads to a richer formulation as an emergent energy-based model.

At each level, we have N agents with states {S_i}, interacting through learned potentials ψ_ij. The joint distribution follows:

$$P(\{S\}) = \frac{1}{Z}\,\mathsf{exp}\left(-\frac{E(\{S\})}{T}\right)$$

where the energy decomposes as:

$$E(\{S\}) = \underbrace{\sum_i D_{KL}(q_i \| p_i(D))}_{\text{data fit}} + \underbrace{\sum_{i,j} \psi_{ij} \cdot d(q_i, q_j)}_{\text{neighbor consistency}} + \underbrace{\sum_i D_{KL}(q_i \| \pi_i)}_{\text{prior drift}}$$

Three levels of learning occur simultaneously: 1. **Beliefs**: Each agent updates q_i via Bayesian fusion 2. **Potentials**: Interaction strengths ψ_ij evolve based on agreement history 3. **Topology**: The graph structure emerges as weak connections are pruned

This formulation connects BEDS to: - **Markov Random Fields**: Local interactions, Gibbs distribution - **Boltzmann Machines**: Learned potentials, stochastic dynamics - **Free Energy Principle**: Variational inference, active inference - **JEPA**: Joint embedding, prediction in latent space

The key addition of BEDS is explicit hierarchy, crystallization, and entropy export to environment.

---

4. Emergence of Mathematical Constants

4.1 Foundational Axioms

We propose that fundamental mathematical constants emerge necessarily from any system capable of representing and updating uncertainty. The BEDS framework rests on four minimal axioms:

Axiom 1 (State): There exist probability distributions P(x) over a measurable space X.

Axiom 2 (Sampling): Realizations x ~ P can be drawn from distributions.

Axiom 3 (Entropy): Every distribution has an entropy H(P) = −∫ P(x) log P(x) dx.

Axiom 4 (Conditioning): Observing evidence y updates distributions via Bayes' rule.

We show that, under natural additional assumptions (continuity, dimensionality, symmetry), the constants e, π, and φ emerge as fixed points.

4.2 Emergence of e

Theorem 4.1 (Emergence of e). *If Bayesian updates are continuous and infinitesimal (likelihood of form 1 + g(x)/n for small perturbations), then their composition produces an exponential function with base e.*

Proof: Consider updating a prior P_0 with n independent observations, each providing likelihood ratio 1 + ε/n. The posterior becomes:

$$P_n \propto P_0 \cdot \left(1 + \frac{\epsilon}{n}\right)^n$$

Taking n → ∞:

$$\lim_{n\to\infty} \left(1 + \frac{\epsilon}{n}\right)^n = e^{\epsilon}$$

The constant e = 2.71828… emerges as the unique base for continuous information accumulation. □

Interpretation: e encodes monotonic growth that perfectly predicts itself—it is the unique function satisfying d/dx(e^x) = e^x. In the BEDS framework, e represents the natural rate of belief crystallization under continuous observation.

4.3 Emergence of π

Theorem 4.2 (Emergence of π). *In a space of dimension ≥ 2 with rotational symmetry, normalizing the maximum-entropy distribution (Gaussian) for fixed variance produces π as the period of angular integration.*

Proof: The maximum entropy distribution for fixed mean and variance is Gaussian:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

The normalization requires:

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

This identity is proven by squaring the integral and converting to polar coordinates:

$$\left(\int_{-\infty}^{+\infty} e^{-x^2} dx\right)^2 = \int\int e^{-(x^2+y^2)} dx\, dy = \int_0^{\infty}\int_0^{2\pi} e^{-r^2} r\, dr\, d\theta = \pi$$

The factor 2π in the angular integral reveals π as the period of rotational symmetry. □

Interpretation: π is not primarily the ratio of circumference to diameter—that geometric fact is derivative. π is the signature of rotational symmetry, emerging whenever systems have no preferred direction. In BEDS, π appears when belief distributions must be normalized over angular degrees of freedom.

4.4 Euler's Identity as Fixed Point

Theorem 4.3 (Euler's Identity). *The equation e^{iπ} + 1 = 0 follows necessarily from Theorems 4.1 and 4.2.*

Proof: Extending the exponential to complex arguments, multiplication by $(1 + i\theta/n)$ represents infinitesimal rotation. Composing n such rotations:

$$\lim_{n\to\infty}\left(1+\frac{i\theta}{n}\right)^n = e^{i\theta} = \cos\theta + i\sin\theta$$

At $\theta = \pi$ (half rotation), we arrive at the point (-1, 0):

$$e^{i\pi} = \cos\pi + i\sin\pi = -1$$

Thus e^{iπ} + 1 = 0. □

Interpretation: Euler's identity is where accumulation (e) applied to half a period (π) produces exact opposition (-1). It is not a mysterious coincidence but the meeting point of two fundamental modes: continuous growth and periodic return.

4.5 Emergence of φ (Golden Ratio)

Theorem 4.4 (Emergence of φ). *The golden ratio $\varphi = (1 + \sqrt{5})/2$ can be expressed as $\varphi = 2\cos(\pi/5)$, linking it to e and π via Euler's formula:*

$$\varphi = e^{i\pi/5} + e^{-i\pi/5} = 2\cos\left(\frac{\pi}{5}\right)$$

Proof: Let $c = \cos(\pi/5)$. Using the identity $\cos(5\theta) = 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta$ and the fact that $\cos(\pi) = -1$, we derive:

$$16c^5 - 20c^3 + 5c + 1 = 0$$

The minimal polynomial for $\cos(\pi/5)$ is $4c^2 - 2c - 1 = 0$, yielding:

$$c = \frac{1+\sqrt{5}}{4}$$

Therefore $2c = (1 + \sqrt{5})/2 = \varphi$. □

Interpretation: The angle π/5 corresponds to the pentagon—the first regular polygon exhibiting non-trivial self-similarity (diagonal/side = φ). At this angle, periodicity generates a number satisfying $\varphi^2 = \varphi + 1$, the equation of self-reference. φ encodes self-similarity: the structure that contains itself.

4.6 Hierarchy of Constants

Constant	Definition	Emergence Mechanism	Mode
e	lim(1 + 1/n)^n	Continuous Bayesian update	Pure accumulation
π	Period of rotation	Normalizing max-entropy distribution	Pure periodicity
i	e^{iπ/2}	Quarter rotation	Accumulation + half periodicity
−1	e^{iπ}	Half rotation (Euler)	Accumulation + full periodicity
φ	2cos(π/5)	Pentagon self-similarity	Accumulation + 1/5 periodicity

These constants form a hierarchy: e and π are primary modes (accumulation and periodicity), while derived constants emerge at specific angular positions combining these modes.

Observation: Under the BEDS axioms and standard mathematical assumptions (continuity, dimensionality ≥ 2, rotational symmetry), these constants emerge as fixed points. They are attractors of systems that represent and update uncertainty—not arbitrary definitions but necessary outcomes of inference under constraints.

---

5. The Gödel-Landauer-Prigogine Conjecture

5.1 Three Foundational Results

Gödel (1931) [41]: Any consistent formal system capable of expressing arithmetic contains true statements unprovable within the system.

Landauer (1961) [14]: Erasing one bit of information requires at least k_B T ln 2 joules of energy.

Prigogine (1977) [8]: Open systems can maintain internal order by exporting entropy to their environment.

5.2 The Common Theme: Closure Has Costs

These three results share a structural pattern: closed systems face fundamental limitations.

Domain	Closure Condition	Consequence
Formal systems	No external axioms	Incompleteness
Computation	No energy dissipation	Reversibility constraints
Thermodynamics	No entropy export	Disorder increases

In each case, opening the system provides an escape: - Adding axioms (or ascending to meta-levels) resolves specific incompleteness instances - Dissipating heat enables irreversible computation - Exporting entropy allows order maintenance

5.3 The Conjecture

Conjecture 5.1 (Gödel-Landauer-Prigogine). *Pathologies of formal systems (incompleteness, undecidability, paradoxes) are structurally analogous to dissipation deficits in physical systems. Specifically:*

1. *Self-referential loops in formal systems correspond to entropy accumulation without export*
2. *Ascending Tarski's hierarchy of metalanguages corresponds to entropy export to meta-levels*
3. *An information system explicitly integrating openness, dissipation, and hierarchy can avoid certain pathologies of purely formal systems*

Rationale: In Gödel's construction, the system creates a statement G: "This statement is unprovable in this system." The self-reference creates a loop with no exit—no way to export the problematic information to an external level. This is analogous to a closed thermodynamic system where entropy accumulates with no means of dissipation.

Physical systems (computers, brains) implementing formal reasoning do not exhibit incompleteness in the same sense. They have other problems (crashes, errors), but not logical pathologies. A hypothesis: their bounded resources and capacity to dissipate (forget, restart, die) protect them from infinite loops.

5.4 Implications

If the conjecture holds:

5.4.1 Stable Mathematics Requires Physicality

Not just in implementation, but structurally. A mathematical community that forgets, prunes, and interacts with reality can remain coherent. An isolated formal system cannot.

5.4.2 Incompleteness Is the Price of Openness

Tarski's hierarchy of metalanguages is infinite. But if each level crystallizes properly before promoting to the next, the system remains usable. Incompleteness is "exported upward"—dissipated to higher meta-levels.

5.4.3 Forgetting Is Necessary for Coherence

Systems that preserve everything (perfect memory, no pruning) accumulate contradictions. To maintain consistency, systems may need to forget. This connects to the observation that biological memory is reconstructive rather than archival [45].

5.5 Status and Limitations

This is a structural analogy, not a proven equivalence. Formal systems and physical systems are not directly comparable—one lives in abstraction, the other in matter. Key open questions:

1. Can the analogy be made precise through formal mapping?
2. What would a "dissipative formal system" look like mathematically?
3. Are there measurable thermodynamic correlates of logical operations?

The interest of the BEDS framework is to propose a bridge between these domains, suggesting that insights from thermodynamics might inform our understanding of formal systems and vice versa.

---

6. Implementation: Sustainable P2P Network

6.1 Architecture Overview

To demonstrate BEDS principles in practice, we present a peer-to-peer network architecture where each node is a BEDS. Nodes observe their environment, update beliefs through Bayesian fusion, and exchange compressed beliefs with neighbors.

Node State: | Component | Role | Notation | |———|—|——-| | Belief | Current state (posterior mean) | μ | | Uncertainty | Variance on belief | σ^2 | | Dissipation | Forgetting rate | γ | | Identity | Public key | pk | | Secret | Private key | sk |

Isolation Principle: Once created, a node (1) cannot be reconfigured, (2) communicates only via signed messages, (3) dies only if flux is cut. This enforces the dissipative constraint: the node exists through the flux that traverses it.

6.2 Exchange Protocol

Message Format: | Field | Description | |—-|———-| | source | Sender public key | | belief | Mean μ | | sigma | Uncertainty σ | | signature | Cryptographic proof |

Bayesian Fusion: When node B receives a belief from node A, it fuses by Gaussian product:

$$\tau_{new} = \tau_A + \tau_B$$
$$\mu_{new} = \frac{\tau_A \cdot \mu_A + \tau_B \cdot \mu_B}{\tau_{new}}$$
$$\sigma_{new} = \sqrt{\frac{1}{\tau_{new}}}$$

where $\tau = 1/\sigma^2$ is precision. This is belief propagation with Gaussian messages—a well-established algorithm [40] now deployed in a P2P context.

Temporal Dissipation: Without new messages, uncertainty grows:

$$\sigma(t) = \sigma_0 \cdot e^{\gamma t}$$

The node “forgets,” returning toward maximum uncertainty. This is thermodynamic death if flux stops—exactly the BEDS existence condition.

6.3 Emergent Hierarchy

Nodes spontaneously organize into levels:

- **Level 0** (Sensors): Raw observations, high uncertainty
- **Level 1** (Aggregators): Fused beliefs from multiple sensors
- **Level 2** (Meta-aggregators): Consensus across aggregator clusters

This hierarchy emerges without central coordination—nodes with low uncertainty become natural aggregation points, attracting connections from neighbors.

6.4 Energy Analysis

For a typical IoT deployment with 5-minute update cycles:

Phase	Duration	Energy
Wake + Measure	15 ms	~1 µWh
Bayesian fusion	20 ms	~1.6 µWh
LoRa transmission	100 ms	~20 µWh
Deep sleep	299.8 s	~275 µWh
Total	**300 s**	**~300 µWh**

Average power: ~3.6 mW

Why Bayesian fusion is negligible: Updating two Gaussians requires only ~10 floating-point operations (2 divisions, 2 multiplications, 2 additions, 1 square root). This is a direct consequence of BEDS: Bayesian fusion is the minimal operation on the statistical manifold.

6.5 Comparison with Existing Systems

System	Annual Consumption	Learning	Consensus
Bitcoin	~120-150 TWh	None	Proof-of-Work
Ethereum (PoW)	~21-95 TWh	None	Proof-of-Work
Ethereum (PoS)	~0.0026 TWh	None	Proof-of-Stake
ChatGPT inference	~1-10 TWh	None (frozen)	N/A
Federated Learning (1M)	~0.01-0.1 TWh	Distributed	Central aggregation
BEDS Network (1M nodes)	**~0.00003 TWh**	**Continuous**	**Bayesian fusion**

Per-operation comparison: | System | Energy per Operation | Equivalent | |——|——————|——–| | Bitcoin transaction | ~1,200 kWh | 2.8 house-months | | Ethereum PoS transaction | ~0.03 kWh | 1 minute microwave | | BEDS belief update | ~0.0000003 kWh | 1 ms LED light |

6.6 Why BEDS Achieves Efficiency

1. **No global consensus**: Nodes communicate only with neighbors; no network-wide broadcast
2. **Bayesian fusion is O(1)**: Independent of network size, unlike gradient aggregation
3. **Computation = compression**: Each update reduces entropy, not just verifies

4. **Natural sparsity**: Uncertainty-weighted communication means low-confidence nodes stay quiet
5. **Solar-sustainable**: 3.6 mW average fits within a 2 cm^2 solar cell budget

6.7 Minimal Runtime: MicroQuickJS

For resource-constrained deployment, BEDS nodes can run on MicroQuickJS (mquickjs), Fabrice Bellard's embedded JavaScript engine:

Specification	mquickjs	V8/Node.js
Minimum RAM	10 KB	~10 MB
ROM footprint	~100 KB	~30 MB
Object overhead	12 bytes	~64 bytes

A complete BEDS node requires ~200 bytes of state. A Raspberry Pi Zero (512 MB) can simulate thousands of BEDS nodes simultaneously.

---

7. Discussion

7.1 On the Nature of Mathematics

The BEDS derivation of mathematical constants suggests an intermediate position between mathematical Platonism and constructivism:

- Mathematics is not **invented** arbitrarily by humans
- It is not **discovered** in a pre-existing abstract realm
- It **emerges necessarily** from any system capable of representing uncertainty

This "emergence" view aligns with naturalized epistemology [46] and embodied cognition [47], suggesting that mathematical truths reflect structural constraints on inference rather than transcendent realities or cultural constructions.

However, we acknowledge the derivations require additional assumptions (continuity, dimensionality, symmetry) not strictly contained in the four axioms. A fully rigorous treatment would need to derive these properties from more primitive principles—an open problem.

7.2 Crystallization and AI Fragility

Current AI systems may suffer from crystallization without ongoing dissipation. When a large language model is frozen after training, it crystallizes a hypothesis about the world at training time. As the environment drifts, the model extrapolates into uncalibrated regions.

Hallucinations, out-of-distribution fragility, and adversarial vulnerability may be symptoms of this pattern. A BEDS network maintains permanent tension between crystallization and dissipation, remaining "liquid"—ordered but adaptable.

This perspective suggests research directions: - Continual learning with controlled forgetting - Uncertainty-aware architectures that know what they don't know - Hierarchical systems where meta-levels supervise and prune lower levels

7.3 Collective Learning Systems

The BEDS framework naturally extends to collective systems—groups of agents who share beliefs and update together through exchange. No central authority; no imposed doctrine; just continuous flow of observations, conversations, and updates.

When enough agents converge on the same belief—when collective uncertainty becomes very low—this belief crystallizes as a shared prior. These crystallized priors then become the foundation on which new beliefs can be built, reducing learning cost at each level.

This has implications for social epistemology, scientific consensus formation, and distributed AI systems.

7.4 Limitations and Open Questions

1. **Formalization gap**: The Gödel-Landauer-Prigogine conjecture remains analogical. Formalizing the mapping between thermodynamic and logical domains is an open challenge.
2. **Empirical validation**: While the P2P architecture demonstrates efficiency, the broader claims about learning and emergence require experimental verification across physical, biological, and computational systems.
3. **Scalability**: The theoretical analysis assumes convergent belief propagation. In practice, loopy graphs and adversarial agents may cause divergence.
4. **Axiom sufficiency**: The four BEDS axioms are minimal but may be insufficient to derive all claimed results without implicit assumptions.
5. **Quantum extension**: Preliminary speculation about quantum BEDS (where uncertainty is physical superposition rather than epistemic ignorance) requires substantial development.

---

8. Conclusion

We have presented BEDS (Bayesian Emergent Dissipative Structures), a theoretical framework proposing that learning is fundamentally the conversion of flux into structure through entropy export. Key contributions include:

1. **Thermodynamic-Bayesian isomorphism**: A formal correspondence between dissipative thermodynamics and Bayesian inference, enabling bidirectional transfer of insights.
2. **Recursive emergence**: The principle that crystallized posteriors become priors for subsequent levels, providing natural energy bounds on learning systems.
3. **Constant emergence**: Derivation of e, π, and φ as fixed points of Bayesian inference under minimal axioms.
4. **Gödel-Landauer-Prigogine conjecture**: The hypothesis that formal system pathologies are structurally analogous to dissipation deficits.
5. **Practical architecture**: A P2P network achieving $10^6\times$ efficiency improvement over existing distributed consensus systems.

The framework bridges physics, logic, and computation, suggesting that sustainable AI must learn like a river: absorbing flux, crystallizing structure, and exporting entropy. This is not merely an efficiency strategy but a structural necessity—systems that crystallize without dissipation accumulate divergence from reality and eventually require replacement.

Future work will focus on formalizing the Gödel-Landauer-Prigogine conjecture, validating BEDS predictions across physical and biological systems, and scaling the P2P architecture to real-world deployments.

> *To maintain consistency, systems may need to forget.*
> *To remain accurate, systems may need to stay open.*
> *To persist, systems may need to dissipate.*

Acknowledgments

This work benefited from extensive discussions with Claude (Anthropic), which served as a collaborative thinking partner throughout the theoretical development. The author thanks colleagues at IGN-ENSG and LaSTIG for their support and critical feedback.

Data and Code Availability

Reference implementation available at: https://github.com/lcaraffa/Bayesian_Emergent_Dissip

---

References

[1] Patterson, D., et al. (2021). Carbon Emissions and Large Neural Network Training. *arXiv:2104.10350*.

[2] de Vries, A. (2023). The growing energy footprint of artificial intelligence. *Joule*, 7(10), 2191-2194.

[3] Quinonero-Candela, J., et al. (2009). *Dataset Shift in Machine Learning*. MIT Press.

[4] Gama, J., et al. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4), 1-37.

[5] Amodei, D., et al. (2016). Concrete problems in AI safety. *arXiv:1606.06565*.

[6] Ovadia, Y., et al. (2019). Can you trust your model's uncertainty? *NeurIPS 2019*.

[7] Strubell, E., et al. (2019). Energy and policy considerations for deep learning in NLP. *ACL 2019*.

[8] Prigogine, I. (1977). Time, Structure and Fluctuations. *Nobel Lecture*.

[9] Prigogine, I. & Stengers, I. (1984). *Order out of Chaos: Man's New Dialogue with Nature*. Bantam Books.

[10] Nicolis, G. & Prigogine, I. (1977). *Self-Organization in Non-Equilibrium Systems*. Wiley.

[11] Shannon, C.E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27, 379-423.

[12] Jaynes, E.T. (1957). Information Theory and Statistical Mechanics. *Physical Review*, 106(4), 620-630.

[13] Jaynes, E.T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

[14] Landauer, R. (1961). Irreversibility and Heat Generation in the Computing Process. *IBM Journal of Research and Development*, 5(3), 183-191.

[15] Bennett, C.H. (1973). Logical Reversibility of Computation. *IBM Journal of Research and Development*, 17(6), 525-532.

[16] Bennett, C.H. (1982). The Thermodynamics of Computation—A Review. *International Journal of Theoretical Physics*, 21(12), 905-940.

[17] Wheeler, J.A. (1990). Information, Physics, Quantum: The Search for Links. In *Complexity, Entropy, and the Physics of Information*. Addison-Wesley.

[18] Cox, R.T. (1946). Probability, Frequency and Reasonable Expectation. *American Journal of Physics*, 14(1), 1-13.

[19] Fisher, R.A. (1925). Theory of Statistical Estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 22(5), 700-725.

[20] Rao, C.R. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37, 81-91.

[21] Amari, S. (1998). Natural gradient works efficiently in learning. *Neural Computation*, 10(2), 251-276.

[22] Amari, S. & Nagaoka, H. (2000). *Methods of Information Geometry*. American Mathematical Society.

[23] Chentsov, N.N. (1982). *Statistical Decision Rules and Optimal Inference*. American Mathematical Society.

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

[25] Friston, K., et al. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1-3), 70-87.

[26] Friston, K. (2019). A free energy principle for a particular physics. *arXiv:1906.10184*.

[27] Parr, T. & Friston, K.J. (2019). Generalised free energy and active inference. *Biological Cybernetics*, 113, 495-513.

[28] LeCun, Y., et al. (2006). A Tutorial on Energy-Based Learning. In *Predicting Structured Data*. MIT Press.

[29] Ranzato, M., et al. (2007). Efficient Learning of Sparse Representations with an Energy-Based Model. *NIPS 2006*.

[30] Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *PNAS*, 79(8), 2554-2558.

[31] Ackley, D.H., Hinton, G.E., & Sejnowski, T.J. (1985). A Learning Algorithm for Boltzmann Machines. *Cognitive Science*, 9(1), 147-169.

[32] Grathwohl, W., et al. (2020). Your Classifier is Secretly an Energy Based Model and You Should Treat it Like One. *ICLR 2020*.

[33] LeCun, Y. (2022). A Path Towards Autonomous Machine Intelligence. *OpenReview*.

[34] Hinton, G.E. & Salakhutdinov, R.R. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(5786), 504-507.

[35] Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. *bitcoin.org*.

[36] Cambridge Centre for Alternative Finance. (2023). Cambridge Bitcoin Electricity Consumption Index. *cbeci.org*.

[37] Buterin, V. (2017). A Proof of Stake Design Philosophy. *Medium*.

[38] McMahan, B., et al. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. *AISTATS 2017*.

[39] Demers, A., et al. (1987). Epidemic algorithms for replicated database maintenance. *PODC 1987*.

[40] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.

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

[42] Chaitin, G.J. (1975). A Theory of Program Size Formally Identical to Information Theory. *Journal of the ACM*, 22(3), 329-340.

[43] Chaitin, G.J. (1982). Gödel's Theorem and Information. *International Journal of Theoretical Physics*, 21(12), 941-954.

[44] Tarski, A. (1936). Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, 1, 261-405.

[45] Schacter, D.L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, 54(3), 182-203.

[46] Quine, W.V.O. (1969). *Ontological Relativity and Other Essays*. Columbia University Press.

[47] Lakoff, G. & Núñez, R.E. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. Basic Books.

[48] Roddier, F. (2012). *Thermodynamique de l'évolution*. Éditions Parole.

[49] Smolin, L. (1992). Did the universe evolve? *Classical and Quantum Gravity*, 9(1), 173-191.

[50] Hofstadter, D. (1979). *Gödel, Escher, Bach: An Eternal Golden Braid*. Basic Books.

[51] Balestriero, R. & LeCun, Y. (2025). LeJEPA: Latent Joint Embedding Predictive Architecture. *arXiv:2511.08544*.

[52] Kingma, D.P. & Welling, M. (2014). Auto-Encoding Variational Bayes. *ICLR 2014*.

[53] Euler, L. (1748). *Introductio in analysin infinitorum*.

[54] Livio, M. (2002). *The Golden Ratio: The Story of Phi, the World's Most Astonishing Number*. Broadway Books.

Appendix A: Complete Proof of Constant Emergence

A.1 Proof of e Emergence

Theorem (Continuous Bayesian Update Limit). *Consider a prior $P_0(\theta)$ updated by n independent observations, each with likelihood ratio $L(\theta) = 1 + \varepsilon \cdot f(\theta)/n$ for small ε. As $n \to \infty$:*

$$P_n(\theta) \propto P_0(\theta) \cdot \exp(\epsilon \cdot f(\theta))$$

Proof: By Bayes' rule:

$$P_n(\theta) \propto P_0(\theta) \cdot \prod_{i=1}^{n} L_i(\theta) = P_0(\theta) \cdot \left(1 + \frac{\epsilon f(\theta)}{n}\right)^n$$

Taking the limit:

$$\lim_{n\to\infty} \left(1 + \frac{\epsilon f(\theta)}{n}\right)^n = e^{\epsilon f(\theta)}$$

The base e emerges as the unique limit of continuous compounding. □

A.2 Proof of π Emergence via Gaussian Normalization

Theorem (Gaussian Integral). *∫_{-∞}^{+∞} e^{-x²} dx = √π*

Proof: Let I = ∫_{-∞}^{+∞} e^{-x²} dx. Then:

$$I^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx\, dy$$

Converting to polar coordinates (r, θ):

$$I^2 = \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r\, dr\, d\theta = 2\pi \cdot \left[-\frac{1}{2} e^{-r^2}\right]_0^{\infty} = 2\pi \cdot \frac{1}{2} = \pi$$

Therefore I = √π. □

A.3 Proof of $\varphi = 2\cos(\pi/5)$

Theorem. *$\cos(\pi/5) = (1 + \sqrt{5})/4$, hence $\varphi = 2\cos(\pi/5)$.*

Proof: Using Chebyshev polynomial $T_5(x) = 16x^5 - 20x^3 + 5x$ and $\cos(5\cdot\pi/5) = \cos(\pi) = -1$:

$$16c^5 - 20c^3 + 5c = -1$$

where $c = \cos(\pi/5)$. This factors to yield minimal polynomial:

$$4c^2 - 2c - 1 = 0$$

Solving: $c = (2 + \sqrt{20})/8 = (1 + \sqrt{5})/4$. Therefore $2c = (1 + \sqrt{5})/2 = \varphi$. □

Appendix B: Energy Budget Calculations

B.1 Node Power Consumption Model

Active Mode (ARM Cortex-M4 @ 48 MHz): - Core consumption: 12 mA @ 3.3V = 39.6 mW - Duration per cycle: 35 ms - Energy: 0.39 μWh

LoRa Transmission (SX1276 @ +14 dBm): - TX current: 85 mA @ 3.3V = 280 mW - Duration: 100 ms - Energy: 7.8 μWh

Deep Sleep (STM32L4 stop mode): - Current: 0.8 μA @ 3.3V = 2.64 μW - Duration: 299.8 s - Energy: 220 μWh

Total per 5-minute cycle: ~228 μWh **Average power**: 2.7 mW

B.2 Solar Sustainability Analysis

Minimum solar cell (monocrystalline, 20% efficiency): - Required power: 3.6 mW average (with margin) - Solar irradiance (temperate): 150 W/m² average - Required area: 3.6 mW / (150 W/m² × 0.20) = 0.12 cm²

A 1 cm² solar cell provides 30 mW—sufficient for ~10 BEDS nodes.

Manuscript submitted for peer review, January 2026.