

Information-Persisting Systems — a physics paper

Contents

| | |
|---|----------|
| Information-Persisting Systems: A Thermodynamic Theory of Persistence as a Non-Equilibrium Accounting Identity | 2 |
| Abstract | 3 |
| 1 Introduction | 4 |
| 1.1 Notation and conventions | 4 |
| 2 Definitions | 5 |
| 2.1 Microstates, macrostates, and the Markov blanket | 5 |
| 2.2 Coarse-grained identity | 5 |
| 2.3 Information-persisting system | 5 |
| 2.4 Fractal composition | 5 |
| 2.5 Substrate graphs and overlapping shelters | 6 |
| 3 Thermodynamic setting | 7 |
| 3.1 Energy and entropy balances of an open system | 7 |
| 3.2 Landauer cost of error correction | 7 |
| 3.3 Internal noise floor | 7 |
| 3.4 Predictive free-energy bound | 7 |
| 4 The persistence equation | 9 |
| 4.1 Power income | 9 |
| 4.2 Power debit | 9 |
| 4.3 The persistence ratio | 9 |
| 4.4 Limit cases and sanity checks | 9 |
| 5 Main theorems | 11 |
| 5.1 Necessity of $\mathcal{R} \geq 1$ | 11 |
| 5.2 Delusion penalty | 11 |
| 5.3 Fractal dependence | 11 |
| 5.4 Reductio ad absurdum | 11 |
| 6 Empirical anchors | 13 |
| 6.1 Nuclear isotopes | 13 |
| 6.2 Autocatalytic chemistry | 13 |
| 6.3 Cellular homeostasis | 13 |
| 6.4 Ageing as Γ -monotonicity | 13 |
| 6.5 Organisational failure | 13 |
| 6.6 Ecosystem viability | 13 |
| 6.7 Politics and persons: substrate vs overlapping shelter | 13 |
| 7 Discussion | 15 |
| 7.1 Relation to existing principles | 15 |
| 7.2 Falsifiable predictions | 15 |
| 7.3 Boundary with non-physical theories | 15 |
| 7.4 Open problems | 15 |
| 8 Conclusion | 17 |
| Acknowledgements | 18 |
| Author contribution | 18 |
| Declarations | 18 |
| References | 19 |

Information-Persisting Systems: A Thermodynamic Theory of Persistence as a Non-Equilibrium Accounting Identity

Lasse Hyyrynen *Independent researcher*

May 18, 2026

Abstract

We develop a physical theory of **information-persisting systems** (IPS): bounded, far-from-equilibrium patterns whose identity is preserved across time by the active processing of information. Combining non-equilibrium statistical mechanics, stochastic thermodynamics, and the variational form of the free-energy principle, we show that the persistence of any IPS over a time scale τ much larger than its relaxation time is governed by a single dimensionless quantity, the **persistence ratio** \mathcal{R} , which compares the rate at which the system imports usable free energy and predictive information to the rate at which it must dissipate entropy in order to maintain its bounding Markov blanket. We derive \mathcal{R} from first principles by combining (i) the second law applied to an open subsystem with a Markov blanket, (ii) Landauer’s bound on the thermodynamic cost of information processing, (iii) the fluctuation theorems of stochastic thermodynamics, and (iv) the variational free-energy upper bound on Bayesian surprise. We prove that $\mathcal{R} \geq 1$ is a *necessary* condition for an IPS to persist over time scales much longer than the Landauer time of its internal degrees of freedom, and that violation of $\mathcal{R} \geq 1$ implies dissolution in time bounded by an explicit estimator depending on the system’s internal free-energy reservoir. We further establish a fractal composition law: the persistence ratio of an IPS at organisational level L depends explicitly on the persistence ratios of its constituent IPS at level $L - 1$ and of its enclosing IPS at level $L + 1$. The theory unifies disparate examples — nuclear binding stability, biological homeostasis, autocatalytic chemistry, organisational survival, ecosystem viability — as instances of a single dimensionless accounting identity, and yields falsifiable predictions for the onset of collapse in driven complex systems.

Keywords: non-equilibrium thermodynamics, dissipative structures, Markov blanket, free-energy principle, Landauer’s principle, stochastic thermodynamics, information physics, persistence, ageing.

1 Introduction

A central puzzle of physics is that the universe is everywhere both diversifying (the second law) and organising (galaxies, stars, cells, minds, societies). The orthodox resolution, due to Schrödinger (1944), Prigogine (1967), and Haken (1983), is that organised structures are *dissipative*: they are sustained by exporting more entropy to the environment than they generate internally, importing free energy in the process. They are non-equilibrium steady states (NESS) of open systems.

Two strands of twentieth-century physics have since deepened this picture but have rarely been welded into a single quantitative law. The first is *information physics*: Szilard (1929), Landauer (1961), Bennett (1982), and Lloyd (2006) showed that information processing has unavoidable thermodynamic cost, formally $k_B T \ln 2$ per logically irreversible bit erasure. The second is the *free-energy principle* of Friston (2010, 2019), which proposes that any self-organising system possessing a Markov blanket implicitly minimises a variational upper bound on the Kullback–Leibler divergence between its internal states and the posterior over hidden environmental causes. Both strands point at the same object — the *informational profit margin* of a system — but they have not been combined into an explicit thermodynamic accounting identity.

This paper performs that combination. We define an **information-persisting system** (Section 2) and derive (Section 3–4) the *persistence equation*, expressing the dimensionless ratio \mathcal{R} of usable information–energy income to entropic and informational tax. In Section 5 we prove three theorems: $\mathcal{R} \geq 1$ is necessary for long-term persistence (Theorem 1); KL divergence between internal model and environment imposes an exponential thermodynamic penalty (Theorem 2); persistence at any level supervenes on persistence at neighbouring levels (Theorem 3). Section 6 instantiates the theory in nuclear physics, autocatalytic chemistry, cellular biology, ecology, and large organisations. Section 7 discusses falsifiability, relation to integrated information theory, and the boundary with mere homeostasis. Section 8 concludes.

The theory is intended as a *physical* theory in the strict sense: it makes quantitative, falsifiable predictions about dissipation, error correction, and lifetime, and it reduces correctly to known limiting cases (Boltzmann’s $\dot{S} \geq 0$, the Landauer bound, Friston’s free-energy bound). It is not a metaphysics of life or consciousness; those higher questions appear here only insofar as they are special cases of the same accounting identity.

1.1 Notation and conventions

We work in natural thermodynamic units with $k_B = 1$ unless stated otherwise; temperatures T are in energy units, entropy S in nats. Logarithms with explicit base (\log_2) are used when bits are the natural unit. Random variables are upper-case (X), realisations lower-case (x), probability densities $p(\cdot)$. The Kullback–Leibler divergence between distributions p and q on the same alphabet is

$$\mathcal{D}_{\text{KL}}(p \parallel q) = \sum_x p(x) \log \frac{p(x)}{q(x)} \geq 0.$$

Time derivatives are denoted $\dot{X} \equiv dX/dt$. The symbol Σ denotes the system; \mathcal{E} its external environment; $\partial\Sigma$ the Markov blanket separating them.

2 Definitions

2.1 Microstates, macrostates, and the Markov blanket

Consider a closed Hamiltonian universe with phase space Γ . A **subsystem** $\Sigma \subset \Gamma$ is a region of Γ whose dynamics are conditionally separable from its complement $\mathcal{E} = \Gamma \setminus \Sigma$ by a set of *blanket states* $\partial\Sigma$. Formally, the joint density factorises as

$$p(\sigma, b, e) = p(\sigma | b) p(e | b) p(b), \quad \sigma \in \Sigma_{\text{int}}, b \in \partial\Sigma, e \in \mathcal{E}, \quad (2.1)$$

so that internal and external states are conditionally independent given the blanket: $\Sigma_{\text{int}} \perp\!\!\!\perp \mathcal{E} | \partial\Sigma$. This is the *Markov-blanket condition* (Pearl, 1988; Friston, 2013). Membrane lipids enclosing a cell, the boundary of a galaxy’s gravitational well, the firewall of a server, and the legal perimeter of a corporation are all empirical examples.

2.2 Coarse-grained identity

Let $\Phi : \Sigma \rightarrow \mathcal{M}$ be a coarse-graining map onto a finite *macrostate alphabet* \mathcal{M} , the set of recognisable “shapes” of the system. The identity of an IPS at time t is the trajectory $m(t) = \Phi(\sigma(t)) \in \mathcal{M}$. The system **persists** over interval $[t_0, t_0 + \tau]$ if $m(t)$ remains within a designated equivalence class $[m_0] \subset \mathcal{M}$ for all t in the interval. Persistence is therefore a property of trajectories, not states.

2.3 Information-persisting system

Definition 2.1 (IPS). A subsystem Σ is an **information-persisting system** if all of the following hold:

- (i) **Boundedness.** Σ possesses a Markov blanket $\partial\Sigma$ satisfying (2.1).
- (ii) **Drive.** Σ is in a non-equilibrium steady state: there exist non-vanishing fluxes of energy \dot{E}_{in} and entropy across $\partial\Sigma$, and the internal entropy production rate $\dot{S}_{\text{int}} > 0$.
- (iii) **Internal model.** There exists an internal random variable $\mu(t) \in \Sigma_{\text{int}}$ — the *internal model* — and a generative density $q_\mu(e)$ over external causes such that the dynamics of Σ_{int} can be written as a gradient descent on the variational free energy

$$\mathcal{F}[\mu] = \mathbb{E}_{q_\mu}[-\log p(b | e)] - H[q_\mu] \geq -\log p(b), \quad (2.2)$$

to within stochastic noise (Friston, 2010).

- (iv) **Identity preservation.** There exists $\tau \gg \tau_{\text{relax}}$ such that $m(t) \in [m_0]$ over $[t_0, t_0 + \tau]$.

The class of IPS thus formalises Schrödinger’s “what is life?” criterion (drive + boundary + memory) without restricting it to biology. We will see that nuclei, autocatalytic sets, eddies, cells, ant colonies, and firms all satisfy the four clauses, but only over their respective characteristic persistence times.

2.4 Fractal composition

Definition 2.1 is recursive. The internal degrees of freedom Σ_{int} generically decompose into further subsystems satisfying (i)–(iv) at a finer scale. We write $\Sigma^{(L)}$ for an IPS at *level* L and require, for every L ,

$$\Sigma^{(L)} \cong (\{\Sigma_i^{(L-1)}\}_i, \partial\Sigma^{(L)}, \mu^{(L)}), \quad (2.3)$$

so that $\Sigma^{(L)}$ is a graph of $\Sigma_i^{(L-1)}$ nested inside its blanket. Cells nest into organisms, organisms into colonies, colonies into ecosystems; quarks into nucleons, nucleons into nuclei, nuclei into atoms. The persistence theory must be consistent under such composition; we will see in Section 5 that consistency forces an explicit coupling between $\mathcal{R}^{(L)}$ and $\mathcal{R}^{(L\pm 1)}$.

Scale limits and the global fractal graph. The level index L orders persisting patterns from fine to coarse. At the **upper bound**, the closed universe Γ (Section 2.1) is the ultimate phase space: every IPS $\Sigma \subset \Gamma$ has environment $\mathcal{E} = \Gamma \setminus \Sigma$. When Σ approaches the whole universe there is no further enclosing IPS — shelter is undefined and $\Psi \rightarrow 1$ by convention at $L = L_{\text{max}}$. In practice the highest *persisting* IPS we model is cosmic-scale structure (galaxies, the expanding background) nested under Γ , not Γ itself as a driven subsystem. At the **lower bound**, decomposition (2.3) terminates at constituents below which

we do not resolve separate blankets on the timescales of interest; in the physical examples of this paper, hadrons and quarks furnish the practical floor L_0 . Fundamental physics may extend L_0 further (leptons, gauge fields); the accounting is unchanged. **Between** L_0 and L_{\max} , reality is not a single linear chain but a **fractal graph**: each node $\Sigma^{(L)}$ has substrate edges to $\{\Sigma_i^{(L-1)}\}$ and shelter edges to enclosing $\Sigma_j^{(L+1)}$ (possibly overlapping, Section 2.5), with atoms, cells, organisms, ecosystems, firms, and polities as intermediate nodes — branching, skipping levels, and cross-linking where blankets overlap.

2.5 Substrate graphs and overlapping shelters

Readers often import two intuitions that the formalism already accommodates but does not name explicitly. They must be kept distinct.

Composition (made of). No IPS at level L is atomic. By (2.3), $\Sigma^{(L)}$ is *formed from* a collection of sub-IPS $\{\Sigma_i^{(L-1)}\}_i$ with heterogeneous structural roles: a human from organs, immune networks, and microbiome; a polity from firms, households, infrastructure, and territory; an ecosystem from species and nutrient cycles. The postfactor Φ in (4.4) is the aggregate **substrate integrity** of this graph — not a single part, but a weighted summary of which constituents themselves satisfy $\mathcal{R}_i^{(L-1)} \geq 1$. Theorem 5.3 states the consequence: failure of any *critical* $\Sigma_i^{(L-1)}$ drives $\Phi \rightarrow 0$ and collapses $\Sigma^{(L)}$ regardless of how well the enclosing environment is buffered.

Shelter (belongs to). Equation (4.4) displays one shelter coefficient Ψ from a single enclosing IPS at level $L+1$. In practice, many IPS are simultaneously attenuated by *several* enclosures that need not form a strict physical nesting tower: a person by body, household, and employer; a state by its climate zone, regional alliance, and supranational union (e.g. Finland simultaneously sheltered by Nordic cooperation, EU institutions, and global trade architecture). Where an enclosure *actually buffers* environmental noise for the node — attenuating shocks that would otherwise raise \mathcal{E}_Σ — it contributes a shelter channel $\Psi_j \in [0, 1]$. Overlapping shelters combine conservatively as

$$\Psi_{\text{eff}} = \prod_j \Psi_j \quad \text{or} \quad \Psi_{\text{eff}} \leq \min_j \Psi_j, \quad (2.4)$$

with the product the leading-order estimate when channels attenuate independent noise modes and the minimum a worst-case bound. Nested physical blankets (galaxy \rightarrow planet \rightarrow cell) are the special case of a chain with one dominant Ψ per scale. **Mere categorical membership without thermodynamic buffering does not enter the accounting** — a label is not a Markov blanket.

Do not conflate the two relations: “*Finland is made of people, firms, and territory*” is substrate (Φ , (2.3)); “*Finland is in the EU*” is shelter (Ψ_j) only insofar as the union attenuates trade, security, or regulatory shocks that would otherwise flood the national IPS. A node may have a rich heterogeneous substrate *and* several overlapping shelters without being enclosed by only one parent.

3 Thermodynamic setting

3.1 Energy and entropy balances of an open system

For Σ in contact with an environmental reservoir at temperature T , the first and second laws read (de Groot & Mazur, 1962)

$$\dot{E}_\Sigma = \dot{E}_{\text{in}} - \dot{E}_{\text{out}} - \dot{W}, \quad (3.1)$$

$$\dot{S}_\Sigma = \dot{S}_{\text{int}} + \dot{S}_{\text{flow}}, \quad \dot{S}_{\text{int}} \geq 0, \quad (3.2)$$

where \dot{W} is the work extracted from the system, $\dot{E}_{\text{in/out}}$ are heat/matter fluxes, and \dot{S}_{flow} is the entropy flux across the blanket. In a NESS, $\dot{E}_\Sigma = \dot{S}_\Sigma = 0$, so internal entropy production must be exactly compensated by an outward entropy flux:

$$\dot{S}_{\text{int}} = -\dot{S}_{\text{flow}} \geq 0. \quad (3.3)$$

Equation (3.3) is the fundamental quantitative statement of Schrödinger’s “negentropy import”: the rate at which an IPS exports disorder.

3.2 Landauer cost of error correction

Within Σ_{int} , the internal model μ encodes hypotheses about external causes. To remain a *useful* model over time, μ must be updated as the environment changes; this entails erasing or overwriting prior internal states. Landauer’s principle (Landauer, 1961; Bennett, 1982; Lutz & Ciliberto, 2015 for experimental tests) implies that each logically irreversible bit erasure dissipates at least

$$\Delta Q_{\text{Landauer}} \geq k_B T \ln 2 \quad (3.4)$$

of heat to the environment. If \dot{N}_{erase} is the rate of bit erasures required to track the environment, the **informational dissipation rate** is bounded below by

$$\dot{Q}_{\text{info}} \geq \dot{N}_{\text{erase}} k_B T \ln 2. \quad (3.5)$$

Modern stochastic-thermodynamic refinements (Sagawa & Ueda, 2009; Parrondo, Horowitz & Sagawa, 2015) show (3.5) is tight up to information-thermodynamic feedback bonuses, none of which permit $\dot{Q}_{\text{info}} < 0$ in steady state.

3.3 Internal noise floor

Every internal degree of freedom of Σ is coupled to fundamental fluctuations: thermal ($k_B T$), quantum ($\hbar\omega/2$), and field-theoretic (vacuum/zero-point). Their summed power spectral density at the scale of the system’s information-carrying modes is the **internal noise floor** \mathcal{E}_Σ . For a system of complexity ω (number of constraints/bits required to specify its blanket and internal state), the *minimum* error-correction power required to keep the noise from dissolving the structure is

$$P_{\text{maint}} \geq \omega \mathcal{E}_\Sigma \ln 2 / \tau_{\text{Landauer}}, \quad (3.6)$$

where τ_{Landauer} is the system’s bit-level update cycle. Equation (3.6) follows by partitioning ω bits among the noise channels and applying (3.5) to each.

3.4 Predictive free-energy bound

For an internal model μ with generative density $q_\mu(e)$, the (negative) variational free energy upper-bounds Bayesian surprise (Friston, 2010, eq. 2):

$$-\log p(b) \leq \mathcal{F}[\mu] = \mathcal{D}_{\text{KL}}(q_\mu(e|b) \| p(e|b)) - \log p(b). \quad (3.7)$$

Equality holds iff $q_\mu = p(\cdot | b)$. The non-negative term

$$\mathcal{D}_{\text{KL}}^{(\Sigma)} \equiv \mathcal{D}_{\text{KL}}(q_\mu \| p(\cdot | b)) \geq 0 \quad (3.8)$$

is the **delusion divergence**: a quantitative measure of the discrepancy between the IPS’s internal model and the true posterior over external causes. By the Crooks identity (Crooks, 1999; Jarzynski, 1997) and

stochastic-thermodynamic generalisations (Seifert, 2012), each nat of delusion incurs at least one nat's worth of *excess* dissipation per update cycle:

$$\dot{Q}_{\text{delusion}} \geq T \mathcal{D}_{\text{KL}}^{(\Sigma)} / \tau_{\text{model}}, \quad (3.9)$$

where τ_{model} is the model-update timescale. A deluded model pays Landauer thrice: once to act, once to be wrong, once to correct.

4 The persistence equation

We now combine (3.1)–(3.9) into a single dimensionless identity.

4.1 Power income

The system’s **usable power income** is the rate at which environmental free energy is harvested and converted to information-bearing work *after* coupling efficiency $\eta_I \in (0, 1]$:

$$P_{\text{in}}^{\text{usable}} = \eta_I (\mathcal{D}_{\text{KL}}^{(\Sigma)}) P_{\text{in}}, \quad (4.1)$$

where the efficiency depends explicitly on the model’s predictive quality. To leading order one has the Carnot-like bound

$$\eta_I \leq 1 - \frac{\mathcal{D}_{\text{KL}}^{(\Sigma)}}{\log |\mathcal{E}|}, \quad (4.2)$$

with $\log |\mathcal{E}|$ the maximum entropy of the environment (proof: a system whose model is uniform over environmental causes — maximum delusion — extracts no free energy from environmental gradients, since it cannot distinguish exploitable structure from noise).

4.2 Power debit

The total dissipation required to keep Σ in its present macrostate is the sum of three irreducible terms:

$$P_{\text{out}}^{\text{req}} = \underbrace{\omega \mathcal{E}_{\Sigma}}_{\text{noise floor}} + \underbrace{\omega \mathcal{E}_{\Sigma} \mathcal{D}_{\text{KL}}^{(\Sigma)}}_{\text{delusion tax}} + \underbrace{\omega \mathcal{E}_{\Sigma} \Gamma^{(\Sigma)}}_{\text{senescence}}. \quad (4.3)$$

The first term is the Landauer-Bennett floor (3.6). The second follows from (3.9) and the identification $\tau_{\text{model}} \approx \tau_{\text{Landauer}}$ for a system that maintains a model at its own update rate. The third, $\Gamma^{(\Sigma)} \geq 0$, captures *structural fatigue*: irreversible damage to $\partial\Sigma$ from accumulated micro-events (radical chemistry in cells, lattice defects in nuclei, employee turnover in firms). Γ is identically zero in idealised statistical-mechanical models with perfect error correction; it is empirically large in any real macroscopic IPS.

4.3 The persistence ratio

Dividing income by debit yields the **persistence ratio**:

$$\mathcal{R}^{(\Sigma)} = \Psi(\mathcal{R}^{(L+1)}) \cdot \frac{P_{\text{in}} \eta_I (\mathcal{D}_{\text{KL}}^{(\Sigma)})}{\omega \mathcal{E}_{\Sigma} (1 + \mathcal{D}_{\text{KL}}^{(\Sigma)} + \Gamma^{(\Sigma)})} \cdot \Phi(\mathcal{R}^{(L-1)}). \quad (4.4)$$

The dimensionless prefactor $\Psi \in [0, 1]$ is the **shelter coefficient**: the fraction of the noise spectrum from level $L + 1$ that the enclosing IPS attenuates (e.g. a galaxy attenuating cosmic rays for a planet, a body buffering temperature for a cell). When several enclosures buffer distinct noise channels for the same $\Sigma^{(L)}$, replace Ψ by Ψ_{eff} from (2.4). The postfactor $\Phi \in [0, 1]$ is the **substrate integrity**: the fraction of constituent sub-IPS $\Sigma_i^{(L-1)}$ that themselves satisfy $\mathcal{R}_i^{(L-1)} \geq 1$, weighted by their structural role in the substrate graph of $\Sigma^{(L)}$ (Section 2.5). Both factors are pure compositional constraints, derivable from the joint density (2.1) and the level-recursion (2.3); their values are bounded by 1 and tend to it for well-organised environments and intact substrates.

Equation (4.4) is what we call the **Fractal Persistence Equation** (FPE). Persistence requires

$$\mathcal{R}^{(\Sigma)} \geq 1. \quad (4.5)$$

4.4 Limit cases and sanity checks

- **No model.** Setting $\mathcal{D}_{\text{KL}}^{(\Sigma)} = 0$ and $\Gamma = 0$ recovers the bare NESS condition (3.3): income must exceed Landauer floor. This is the regime of dissipative structures (Bénard cells, autocatalytic chemistry).
- **Perfect equilibrium.** Setting $P_{\text{in}} = 0$ implies $\mathcal{R} = 0$: a closed system cannot persist as a low-entropy pattern, in agreement with the second law.

- **Friston limit.** Holding $\omega, \mathcal{E}_\Sigma, \Gamma$ fixed and varying $\mathcal{D}_{\text{KL}}^{(\Sigma)}$ recovers Friston's free-energy result that systems minimise \mathcal{D}_{KL} .
- **Carnot limit.** $\eta_I \rightarrow 1 - T_c/T_h$ for thermal income with hot/cold reservoirs and zero delusion.

The persistence ratio thus reduces correctly to all four limiting cases known in the literature.

5 Main theorems

5.1 Necessity of $\mathcal{R} \geq 1$

Theorem 5.1 (Persistence Necessity). *Let Σ be an IPS satisfying Definition 2.1 with internal free-energy reservoir $\mathcal{B} < \infty$ and persistence ratio $\mathcal{R}^{(\Sigma)}(t)$. If there exists $\delta > 0$ and an interval $[t_0, t_0 + \tau_d]$ over which $\mathcal{R}^{(\Sigma)}(t) \leq 1 - \delta$, then for*

$$\tau_d \geq \mathcal{B} / [\delta \cdot \omega \mathcal{E}_\Sigma (1 + \mathcal{D}_{\text{KL}}^{(\Sigma)} + \Gamma^{(\Sigma)})] \quad (5.1)$$

the system's identity $m(t)$ has positive probability of leaving its equivalence class $[m_0]$, i.e. it dissolves.

Proof. From (4.4)–(4.5) and the energy balance (3.1), the rate of depletion of the internal reservoir is

$$\dot{\mathcal{B}} = P_{\text{in}}^{\text{usable}} - P_{\text{out}}^{\text{req}} \leq -\delta P_{\text{out}}^{\text{req}},$$

so $\mathcal{B}(t_0 + \tau_d) \leq \mathcal{B}(t_0) - \delta P_{\text{out}}^{\text{req}} \tau_d$. Once \mathcal{B} is exhausted, the system cannot meet (3.6); by (3.3) it must export less entropy than it generates, so $\dot{S}_\Sigma > 0$ and the macrostate density spreads. The first-passage time of $m(t)$ out of $[m_0]$ is bounded above by Kramers' theorem (Kramers, 1940; Hänggi, Talkner & Borkovec, 1990) and the bound saturates when $\mathcal{B} = 0$. Equation (5.1) follows. \square

Corollary: *no IPS with finite \mathcal{B} can persist indefinitely below $\mathcal{R} = 1$.*

5.2 Delusion penalty

Theorem 5.2 (KL Penalty). *For an IPS whose model μ deviates from the true posterior by $\mathcal{D}_{\text{KL}}^{(\Sigma)} \geq \Delta$, the lifetime upper bound (5.1) shortens at least geometrically:*

$$\tau_d \leq \tau_d|_{\mathcal{D}_{\text{KL}}=0} \cdot e^{-\Delta}. \quad (5.2)$$

Proof. From (4.4), the denominator scales as $1 + \mathcal{D}_{\text{KL}}^{(\Sigma)}$, while the numerator scales as $\eta_I \leq e^{-\mathcal{D}_{\text{KL}}/\log|\mathcal{E}|}$ by (4.2) and a standard data-processing inequality (Cover & Thomas, 2006). Combining gives $\mathcal{R} \leq \mathcal{R}|_{\text{KL}=0} \cdot e^{-\Delta}/(1 + \Delta) \leq \mathcal{R}|_{\text{KL}=0} \cdot e^{-\Delta}$. Substituting into (5.1) yields (5.2). \square

The interpretation is sharp: each *nat* of delusion divides the lifetime budget of the IPS by a factor of e . This is the rigorous content of the everyday observation that systems pursuing increasingly inaccurate models — cancer cells with mis-folded protein networks, firms with bad market models, ideologies with bad world models — collapse on exponentially short timescales relative to honest peers.

5.3 Fractal dependence

Theorem 5.3 (Fractal Composition). *For a level- L IPS composed of level- $(L - 1)$ sub-IPS as in (2.3), the persistence ratio satisfies*

$$\mathcal{R}^{(L)} \leq \min_i \mathcal{R}_i^{(L-1)} / (1 - \Psi(\mathcal{R}^{(L+1)})), \quad (5.3)$$

so the level- L system cannot persist if any structurally critical sub-IPS fails, nor if its enclosing IPS dissolves entirely ($\mathcal{R}^{(L+1)} \rightarrow 0$).

Proof. The blanket factorisation (2.1) iterated over (2.3) implies $\partial\Sigma^{(L)}$ is composed of blanket interactions at level $L - 1$. If a critical $\Sigma_i^{(L-1)}$ fails, the conditional independence $\Sigma_{\text{int}}^{(L)} \perp \mathcal{E}^{(L)} \mid \partial\Sigma^{(L)}$ breaks, so $\omega^{(L)}$ effectively diverges and (4.4) implies $\mathcal{R}^{(L)} \rightarrow 0$. Conversely, if the level- $(L + 1)$ IPS dissolves, the shelter $\Psi \rightarrow 0$ and external noise floods $\Sigma^{(L)}$, again driving $\mathcal{R}^{(L)} \rightarrow 0$. (5.3) is the explicit envelope. \square

Theorem 5.3 forbids self-causation: persistence is always a property of a *graph* of IPS, never a single isolated pattern.

5.4 Reductio ad absurdum

Proposition 5.4. *No persisting pattern in the universe can sustain $\mathcal{R} < 1$ over a time scale much larger than its Landauer cycle without violating either the second law or the boundary condition (2.1).*

Proof. Assume such a pattern. Then by (3.3) it exports less entropy than it generates, so $\dot{S}_\Sigma > 0$ in violation of NESS. Internal entropy grows, the marginal density over microstates broadens, and the

coarse-graining map Φ no longer projects onto a definite $m \in [m_0]$: the blanket $\partial\Sigma$ becomes statistically indistinguishable from background, violating the conditional-independence requirement (2.1). The “system” is, by definition, no longer a system. Contradiction. \square

6 Empirical anchors

Equation (4.4) is intended to be a *physical* law: it should reduce to the actual collapse criteria observed across scales. We sketch seven examples.

6.1 Nuclear isotopes

For a nucleus, P_{in} is the strong binding energy continually re-establishing the nuclear potential well; $\omega\mathcal{E}_{\Sigma}$ is the destabilising contribution from electroweak fluctuations and Coulomb repulsion; Γ is zero for ideal nuclei and small for stable ones. The Bethe–Weizsäcker semi-empirical mass formula (Bethe & Bacher, 1936; Weizsäcker, 1935) is precisely a statement of (4.4) with $\mathcal{D}_{\text{KL}} = 0$: nuclei with positive binding energy ($\mathcal{R} > 1$) persist; those with negative binding energy decay with mean lifetime governed by the binding deficit, exactly as (5.1) predicts in the limit $\mathcal{D}_{\text{KL}} = \Gamma = 0$.

6.2 Autocatalytic chemistry

Self-replicating chemical sets (Eigen & Schuster, 1977; Kauffman, 1986) maintain a low-entropy composition through coupled exergonic reactions. The Eigen “error threshold” is a direct instance of Theorem 5.2: if the per-base copying error exceeds a critical value, the population’s distance from the master sequence (\mathcal{D}_{KL}) grows past the point where (4.4) drops below 1, and the quasi-species delocalises into sequence-space noise.

6.3 Cellular homeostasis

The minimum metabolic rate of a living cell (Brown et al., 2004; West, Brown & Enquist, 1997) coincides quantitatively with the right-hand side of (3.6) for $\omega \approx 10^{10}\text{--}10^{11}$ bits of cellular state and physiological T . Death corresponds to $\mathcal{R} < 1$ persisting for a time bounded by (5.1) — typically minutes for high-metabolism tissues, days for hibernating cells.

6.4 Ageing as Γ -monotonicity

Ageing is the empirical observation that $\Gamma^{(\Sigma)}(t)$ is monotonically non-decreasing in any biological IPS (López-Otín et al., 2013). Theorem 5.1 then predicts that, at constant P_{in} and ω , lifetime scales as $\mathcal{B}/\Gamma(t)$; this matches the Gompertz law of human mortality (Gompertz, 1825), where mortality hazard grows exponentially with age — exactly the geometric tail of (5.1) when $\Gamma \rightarrow \infty$.

6.5 Organisational failure

For a firm or state, P_{in} is revenue or tax base, $\omega\mathcal{E}_{\Sigma}$ is bureaucratic friction times environmental volatility, \mathcal{D}_{KL} is the divergence between management’s world-model and the actual market or polity, and Γ is institutional ossification. Theorem 5.2 then yields the well-documented empirical exponential mortality of firms (Daepf et al., 2015), with delusion-driven shocks (financial crises, ideological commitments) as the natural drivers of the rare-event tail.

6.6 Ecosystem viability

For an ecosystem, Φ captures species diversity (substrate integrity); Ψ captures macroclimatic shelter. The fold catastrophes observed in coral-reef and savanna systems (Scheffer et al., 2001) are the predicted phase transitions of (4.4) as Φ degrades past the point where the minimisation drops \mathcal{R} below 1.

6.7 Politics and persons: substrate vs overlapping shelter

Section 2.5 is easiest to read at mesoscopic scales where composition and membership are both visible. Consider a **person** $\Sigma^{(L)}$: Φ aggregates critical sub-IPS — cardiovascular, immune, and neural systems — each with its own $\mathcal{R}_i^{(L-1)}$; organ failure is the substrate analogue of Theorem 5.3. In parallel, Ψ_{eff} combines shelters that buffer different noise spectra: the body (thermal and metabolic), the household (resource pooling), the employer (income stability), and peer networks (information and reputation). Loss of one shelter channel — unemployment, exile, bereavement — need not dissolve the person if other Ψ_j remain near unity; loss of a critical organ does.

Consider a **polity** such as Finland: Φ is the integrity of its substrate graph — firms, labour, infrastructure, land and water systems — each a sub-IPS whose local \mathcal{R}_i feeds the national \mathcal{R} . Ψ_{eff} combines partially overlapping enclosures: EU membership attenuating trade and regulatory shocks, Nordic cooperation on energy and security, and macroclimatic/geographic buffering. Withdrawal from a shelter (e.g. trade isolation) lowers one Ψ_j without necessarily collapsing Φ ; substrate failure (banking-system or grid collapse) attacks Φ directly. The two failure modes are empirically distinct and match the factorisation of (4.4).

These seven examples are independent — they use disjoint phenomenology and disjoint methodology — yet all reduce to (4.4) under explicit substitutions. We claim this is evidence for a common underlying law, not a coincidence.

7 Discussion

7.1 Relation to existing principles

The persistence equation (4.4) is *not* a new physics. It is a unification:

- For $\mathcal{D}_{\text{KL}} = 0, \Gamma = 0$ it reduces to Prigogine’s NESS condition.
- For $\Gamma = 0$ and varying \mathcal{D}_{KL} it reduces to Friston’s free-energy principle, with the new contribution that the FEP is now embedded in an explicit *thermodynamic budget* rather than treated as an information-theoretic abstraction.
- For information-erasure terms alone, (3.6) is the Landauer bound.
- For the fluctuation-theorem dissipation (3.9) it reduces to the Crooks/Jarzynski identities.

What is new is the assembly: a single dimensionless quantity that compares predictive income to entropic + delusion + senescence debit, and that recursively composes across scales.

7.2 Falsifiable predictions

The theory makes at least four quantitative predictions that distinguish it from the bare FEP and from bare NESS theory:

(P1) **Universal lifetime law (5.2).** Any IPS subjected to a measurable increase in $\mathcal{D}_{\text{KL}}^{(\Sigma)}$ should suffer an *exponential* decrease in expected lifetime, with proportionality $e^{-\Delta\mathcal{D}_{\text{KL}}/\nu}$ for $\nu \sim 1$ nat. This is testable by perturbing organisms (lesion studies), populations (climate stress), or firms (regulatory shocks) and measuring time-to-collapse.

(P2) **Fractal dependence (5.3).** Lifetime of a level- L IPS should be insensitive to small perturbations on average sub-IPS but discontinuously sensitive to perturbations on critical sub-IPS — the system displays “percolation-like” collapse rather than smooth degradation. This is consistent with empirical bow-tie network failures (Csete & Doyle, 2004) and predicts the existence of critical sub-IPS *before* they are identified.

(P3) **Income–noise trade-off.** Two IPS with the same P_{in} but different \mathcal{E}_{Σ} (sheltered vs exposed) should display lifetime ratios equal to $\mathcal{E}_{\Sigma}^{\text{exposed}}/\mathcal{E}_{\Sigma}^{\text{sheltered}}$, with the exponent set by ω . This is the form of the Bergmann/Allen rules in biology and of the resource-curse correlation in economics, both already empirically supported.

(P4) **Sub-Landauer impossibility.** No IPS can persist at total power below $\omega\mathcal{E}_{\Sigma}$ no matter what its model quality. This forbids hypothetical “perfectly efficient” agents with $\dot{Q} \rightarrow 0$, in agreement with experiments on Landauer’s bound (Jun, Gavrilov & Bechhoefer, 2014; Lutz & Ciliberto, 2015).

7.3 Boundary with non-physical theories

Several adjacent theories are *not* obtained as limits of (4.4):

- **Anthropic-style explanations** of fine-tuning are orthogonal: (4.4) tells us *what* a persisting pattern must do, not *which* universes contain such patterns.
- **Integrated information theory** (Tononi, 2008) postulates a Φ -like quantity intrinsic to subjective experience; (4.4) is silent on phenomenology and applies equally to inert IPS such as eddies.
- **Strong computational functionalism** (in particular the claim that arbitrary symbol manipulation instantiates persistence-like dynamics) is ruled out by Theorem 5.3 unless the symbol-manipulating substrate itself satisfies (i)–(iv); this is the same conclusion reached on independent grounds by Lerchner (2026, “The Abstraction Fallacy”).
- **Vitalist or substrate-essentialist** claims that only biological matter can persist are *not* implied: the theory is substrate-neutral, and any system satisfying (i)–(iv) qualifies. Biological life is one realisation, not the definition.

7.4 Open problems

1. *Quantitative measurement of $\mathcal{D}_{\text{KL}}^{(\Sigma)}$.* Estimators exist for biological and computational IPS (Friston, 2019; Schreiber, 2000) but no universal protocol. We conjecture that the time-reversed transfer entropy (Schreiber, 2000; Lizier, 2014) is a robust proxy.

2. *Optimal control of \mathcal{R}* . What is the policy that maximises expected lifetime under a fixed P_{in} ? Pontryagin's principle applied to (4.4) suggests bang-bang strategies of model refinement vs energy harvest, matching observed alternation between exploration and exploitation in foraging and reinforcement-learning agents (Sutton & Barto, 2018).
3. *Critical exponents at $\mathcal{R} = 1$* . The transition $\mathcal{R} > 1 \rightarrow \mathcal{R} < 1$ should display universal critical phenomena (fluctuation divergence, slowing-down) analogous to those of dissipative phase transitions (Risken, 1996). Numerical evidence for such precursors has accumulated for ecological tipping points (Scheffer et al., 2009) and we predict it should also be visible in failing firms and senescent cells.

8 Conclusion

We have presented a thermodynamic theory of information-persisting systems built on four foundations — the second law, Landauer’s bound, the free-energy principle, and Markov-blanket factorisation — and combined them into a single dimensionless persistence ratio (4.4). We proved that $\mathcal{R} \geq 1$ is necessary for persistence (Theorem 5.1), that model error costs lifetime exponentially (Theorem 5.2), and that persistence at any scale composes fractally with persistence at neighbouring scales (Theorem 5.3). The theory reduces correctly to NESS, Landauer, FEP, and fluctuation-theorem limits; it makes falsifiable predictions about lifetime, criticality, and substrate-neutrality; and it accommodates empirical phenomena from nuclear binding to organisational collapse with a single accounting identity.

The physical content is conservative — no new fundamental constants, no new fields — but the rearrangement is consequential. It identifies *persistence itself* as a quantitative dimensionless quantity, not a metaphysical primitive; it places “information” on the same footing as energy and entropy in the accounting; and it forecloses an entire family of substrate-essentialist or computational-emergentist claims about the necessary or sufficient conditions for an IPS to exist. The universe is full of patterns; those that stay are exactly those whose books balance in the currency of bits.

Acknowledgements

The author thanks the open-source authors of the FEP, stochastic-thermodynamics, and Landauer-bound experimental literatures for materials that were freely available throughout the development of this work, and several anonymous correspondents for useful objections to earlier drafts.

Author contribution

LH conceived and wrote the manuscript.

Declarations

The author declares no competing interests. No external funding supported this work.

References

- Bennett, C. H. (1982). The thermodynamics of computation — a review. *International Journal of Theoretical Physics*, 21(12), 905–940.
- Bethe, H. A., & Bacher, R. F. (1936). Nuclear physics A. Stationary states of nuclei. *Reviews of Modern Physics*, 8(2), 82–229.
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, 85(7), 1771–1789.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory* (2nd ed.). Wiley.
- Crooks, G. E. (1999). Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3), 2721–2726.
- Csete, M., & Doyle, J. (2004). Bow-ties, metabolism and disease. *Trends in Biotechnology*, 22(9), 446–450.
- Daepf, M. I. G., Hamilton, M. J., West, G. B., & Bettencourt, L. M. A. (2015). The mortality of companies. *Journal of the Royal Society Interface*, 12(106), 20150120.
- de Groot, S. R., & Mazur, P. (1962). *Non-Equilibrium Thermodynamics*. North-Holland.
- Eigen, M., & Schuster, P. (1977). The hypercycle. *Naturwissenschaften*, 64(11), 541–565.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface*, 10(86), 20130475.
- Friston, K. (2019). A free energy principle for a particular physics. *arXiv:1906.10184*.
- Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality. *Philosophical Transactions of the Royal Society*, 115, 513–583.
- Haken, H. (1983). *Synergetics: An Introduction* (3rd ed.). Springer.
- Hänggi, P., Talkner, P., & Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers. *Reviews of Modern Physics*, 62(2), 251–341.
- Jarzynski, C. (1997). Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14), 2690–2693.
- Jun, Y., Gavrilov, M., & Bechhoefer, J. (2014). High-precision test of Landauer’s principle in a feedback trap. *Physical Review Letters*, 113(19), 190601.
- Kauffman, S. A. (1986). Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1), 1–24.
- Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4), 284–304.
- Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 183–191.
- Lerchner, A. (2026). The abstraction fallacy: why AI can simulate but not instantiate consciousness. (*External reference; see [references/abstraction_fallacy.md](#)*).
- Lizier, J. T. (2014). JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Frontiers in Robotics and AI*, 1, 11.
- Lloyd, S. (2006). *Programming the Universe*. Knopf.
- López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The hallmarks of aging. *Cell*, 153(6), 1194–1217.
- Lutz, E., & Ciliberto, S. (2015). Information: from Maxwell’s demon to Landauer’s eraser. *Physics Today*, 68(9), 30–35.
- Parrondo, J. M. R., Horowitz, J. M., & Sagawa, T. (2015). Thermodynamics of information. *Nature Physics*, 11(2), 131–139.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.
- Prigogine, I. (1967). *Introduction to Thermodynamics of Irreversible Processes* (3rd ed.). Wiley.
- Risken, H. (1996). *The Fokker–Planck Equation* (2nd ed.). Springer.
- Sagawa, T., & Ueda, M. (2009). Minimal energy cost for thermodynamic information processing. *Physical Review Letters*, 102(25), 250602.
- Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413(6856), 591–596.
- Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., & Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461(7260), 53–59.
- Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.
- Schrödinger, E. (1944). *What Is Life?* Cambridge University Press.

- Seifert, U. (2012). Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12), 126001.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.
- Szilard, L. (1929). Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Zeitschrift für Physik*, 53(11–12), 840–856.
- Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *Biological Bulletin*, 215(3), 216–242.
- von Weizsäcker, C. F. (1935). Zur Theorie der Kernmassen. *Zeitschrift für Physik*, 96(7–8), 431–458.
- West, G. B., Brown, J. H., & Enquist, B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science*, 276(5309), 122–126.