The Theory of General Intelligence: Intelligence as a Thermodynamically Favored Phenomenon

Mounir Shita
Global Economic Alliance and EraNova Global
Preprint v1 — May 2024

Abstract—We develop a physics-first framework for understanding intelligence. Rather than treating intelligence as a biological accident, we argue that goal-directed causal manipulation is a thermodynamically favored outcome for any physical system that maintains persistent low-entropy structure in an energy-rich, far-from-equilibrium environment. We do not claim a deductive proof that intelligence must arise everywhere; we present a derivation showing why, under stated conditions, systems that model causality, simulate forward consequences, and select interventions are strongly selected on energetic grounds.

Building on a dimensional-insufficiency argument [1] — that purely temporal representations cannot capture causal emergence — we formulate an emergence argument: in spacetime regions with thermodynamic disequilibrium, sufficient material complexity, and ordinary light-cone causal structure, configurations with higher causal-prediction accuracy are energetically favored, so the probability of intelligent causal agency approaches unity as organizational complexity grows. We introduce the Comprehension Factor (CF) as a substrate-independent per-goal measure of how reliably one planned causal chain reaches its target, and argue that intelligence is not the maximization of any single CF but the management of a network of causally linked goals: no known embodied living agent operates with a single isolated goal, so the object intelligence actually operates on is always a goal network, and the defining act is deciding which goals to pursue, delay, or abandon so that local success does not raise the probability of network-wide collapse. The network-level quantity is the Temporal Comprehension Factor (TCF), a functional of the agent's policy — the product of the Comprehension Factors of the selections made along a path through the goal network, evaluated in contexts that each selection reshapes; single-goal CF is its degenerate one-goal case. We are explicit about epistemic standing: the conceptual core is goal-network management, while the falsifiable anchor is the per-goal machinery — we apply CF in an initial worked example, Escherichia coli chemotaxis, where the four predicted operations are realized in classical biochemistry and CF is estimated from published chemotaxis data using an explicit information-efficiency proxy tied to measured drift-speed bounds, and the network-level claim is carried by a stated prediction rather than asserted. We close with falsifiable predictions across biological, artificial, and social systems, each with an explicit failure criterion. The work is presented as a research program whose empirical content is the falsifiable-prediction section.

Keywords: *general intelligence; thermodynamics of computation; causal modeling; non-equilibrium systems; comprehension factor; falsifiability*

1. INTRODUCTION AND THE PHYSICS-FIRST AXIOM

We adopt the position established in [1]: there exists exactly one universe, and it is the relativistic spacetime described by physics. Any entity claimed to be causally efficacious must lie within spacetime and obey its causal structure — light-cone constraints, thermodynamic costs, Lorentz covariance. This is not merely a philosophical preference. Any entity X that influences events must transmit information to them; special relativity forbids superluminal information propagation; therefore X must lie within the past light cone of the affected events. Entities placed outside spacetime are causally disconnected from every physical process and are explanatorily inert.

We begin from a minimal ontology: reality consists of quantum fields and their excitations arranged in four-dimensional spacetime with signature (−,+,+,+). Intelligence, insofar as it is physically real, must be explicable in these terms. The remainder of Part I derives physical definitions of goals and intelligence; Part II states the emergence argument, the Comprehension Factor, and its temporal extension; Part III applies the machinery in a worked molecular example; Part IV states falsifiable predictions and addresses anticipated objections.

Scope of this preprint. The framework presented here is the minimal physical core of a larger body of work [2]. The full Comprehension Factor includes a spatial–temporal penalty and a hierarchical-efficiency (H-Factor) term; both appear in the CF equation below, but only the core product and the multi-agent interference term carry quantitative weight here — the spatial–temporal penalty and the H-Factor parameters are stated structurally and their calibration is left to the fuller treatment and future empirical work. That fuller treatment also develops multi-path simulation ("thinking"), in which an agent evaluates competing futures by their CF values before committing, and the engineering implications for a single, domain-general foundational algorithm; these are not developed here. The present paper isolates the publishable core — goals as low-entropy configurations, intelligence as causal intervention, the Causality Hierarchy as energetic compression, the multi-agent interference term, the Comprehension Factor and its temporal extension (TCF) as operational measures, and falsifiable predictions. The deferred components are higher-order extensions of this core, not abandoned parts of it; pressing their unfitted parameters into this paper would dilute the falsifiable content it is built to expose.

Relation to prior work on prediction and thermodynamics. A connection between predictive modeling and thermodynamic efficiency is already established. Still et al. [27] prove that, for a system responding to a stochastic drive, the nonpredictive part of its memory lower-bounds dissipation, so any memory-bearing system operating at maximal energetic efficiency must be predictive. That result is derived under an explicit assumption of no feedback from the system to the driving signal: the system conditions on an environment it does not act upon, the regime of P(e | c). The present framework concerns the regime that assumption sets aside — systems that intervene on their environment toward a represented low-entropy target, P(e | do(c)). The Comprehension Factor is therefore not a re-derivation of the prediction–dissipation equivalence but a measure defined on the interventional structure those results exclude by construction; Still et al. themselves note adaptation toward environmental function as an open direction, which is precisely where goal-directed intervention enters. A second line of work derives apparently goal-directed behavior from entropy maximization rather than from a represented goal. Wissner-Gross and Freer [28] propose causal entropic forces under which a system is driven to maximize the diversity of its accessible future paths, and report that behaviors resembling tool use and cooperation can emerge in simple systems from this principle alone. The present framework is the sign-inverse of that construction: it concerns directed approach to a specified low-entropy target configuration, not maximization of future-path entropy, and it requires a represented goal against which interventions are selected — an element causal entropic forces do not contain. The two are therefore not variants of one principle but opposite optimizations over different objects. Dissipation-driven adaptation results (England [29]) likewise describe organization favored under nonequilibrium driving without a represented target, and so sit on the same side of this distinction. We note that the causal-entropic-force result itself drew a priority comment [30]; the emergence claim made here is correspondingly stated as favored rather than as a derived inevitability. What originates the intervention — the selection of do(c) itself — is a distinct question the formal results here do not address and on which nothing in this paper depends; it is taken up in separate work.

2. FROM GOAL-DIRECTED BEHAVIOR TO PHYSICAL GOALS

Across psychology, cognitive science, evolutionary biology, and artificial intelligence, intelligence is agreed to involve goal-directed behavior, but the consensus is descriptive. Standard definitions are circular: a goal is a desired outcome, a target, an objective — each presupposing the intentionality we aim to explain. We derive what a goal is from first principles.

2.1. Particle-configuration ontology

Any observable phenomenon reduces to a configuration of particles in spacetime: a bacterium is an arrangement of roughly 10^{11} atoms; a galaxy some 10^{67} atoms in a spiral. Let Ω denote the configuration space of all possible arrangements; for N particles with six degrees of freedom it scales as a continuum raised to $6N$. When we say a bacterium "seeks nutrients," we reference a physical transition from configuration ω_p to a configuration ω_g of elevated local nutrient concentration [5].

2.2. Goals as low-entropy configurations

Most configurations in Ω are high-entropy and thermodynamically typical. Yet the universe contains persistent low-entropy structures — crystals, living cells, cities — each occupying a vanishingly small region of Ω. Schrödinger's observation [6] is that living systems import free energy and export entropy, maintaining local order through continuous work that opposes relaxation toward disorder.

Definition 1 (Physical goal). *A goal is a specific future low-entropy particle configuration $\omega g \in \Omega$ satisfying (i) spatial localization to a region $R \subset \mathbb{R}^3$; (ii) temporal specification within $[t_{min}, t_{max}]$; and (iii) low entropy, $S(\omega g) \ll S(\Omega)$.*

3. INTELLIGENCE AS CAUSAL MANIPULATION

Definition 2 (Intelligence). *Intelligence is the capacity of a physical system to manipulate causal chains in four-dimensional spacetime toward specific low-entropy configurations, through strategic intervention informed by an internal causal model.*

The definition is substrate-independent (function, not material), universal (no biochemical or evolutionary assumptions), graded rather than binary (quantified by the Comprehension Factor below), and physically grounded (causal chains are sequences of events linked by light-cone-respecting causal relations; interventions are physical events; modeling is matter configured to represent causal structure). Intelligence manifests through four operations observed across natural systems: (i) causal modeling — internal representation of cause–effect relations; (ii) forward simulation — prediction of future configurations under candidate interventions; (iii) action selection — choice of interventions that advance ωg while preserving the broader goal network, rather than merely maximizing the isolated probability of reaching a single goal; and (iv) physical execution — introduction of the chosen cause into spacetime.

4. THE CAUSALITY HIERARCHY AS THERMODYNAMIC NECESSITY

4.1. Intractability of reductionism

Reductionism is correct in principle but intractable for finite agents. Predicting that pressing a button yields coffee, at the fundamental level, requires tracking on the order of 10^{26} atoms. By Landauer's principle [11], each irreversible bit operation costs at minimum $kT \ln 2 \approx 3\times10^{-21}$ J at $T = 300$ K. A naive bit-level accounting over such a system across tens of seconds yields a computational energy on the order of 10^5 J — far exceeding the free energy available from generic environmental gradients (10^2–10^3 J). This figure is a deliberate lower bound — it counts a single bit per atom over the interval, whereas tracking full classical microstates would raise the cost by many further orders of magnitude — so the understatement only strengthens the conclusion. We treat this as an order-of-magnitude argument; the conclusion is robust to the precise constants.

Proposition 1 (Abstraction necessity). *Any physical system maintaining persistent low-entropy configurations in an energy-limited environment must construct hierarchical abstractions of causal relationships that skip intermediate microphysics.*

Argument. Suppose, for contradiction, that a system maintains its goal through full reductionist simulation. Its computational energy dwarfs the available free energy, so it cannot persist. Any process — Darwinian or purely thermodynamic — favoring systems that minimize energy per successful intervention therefore favors systems that bypass microphysics via hierarchical causal models, whose energetic advantage over full simulation is enormous. This is a scaling argument, not a formal proof: the constants are order-of-magnitude.

4.2. Structure of the hierarchy

Definition 3 (Causality Hierarchy). *A Causality Hierarchy (CH) is a four-dimensional knowledge structure organizing cause–effect relations by spatial abstraction level s, temporal-displacement level t, causal probability $P(e \mid do(c), \theta)$, and query energy cost $E(s,t)$.*

The 4D structure is required because causal relations are irreducibly temporal: cause precedes effect, and the light-cone structure constrains propagation. Query cost falls with abstraction roughly as $E(s,t) \approx E_0 \exp[-\alpha(s+t)]$. "Press button → coffee" is one top-layer link; the molecular route is $\sim10^{26}$ links — same prediction, vastly different cost.

4.3. Convergence with neuroscience

If hierarchical causal structure is energetically mandatory, biological intelligence should exhibit it, and it does: prefrontal cortex carries the longest behavioral timescales — intrinsic electrophysiological timescales rise along the cortical hierarchy into the hundreds of milliseconds [31], while goal and memory representations extend over seconds to years in the working-memory and learning literature [32] — premotor cortex over hundreds of milliseconds, primary motor cortex over tens of milliseconds, cerebellum at millisecond to tens-of-milliseconds timing precision. Friston's free-energy principle [13] and predictive coding [19] describe brains as hierarchical prediction-error minimizers; hippocampal time cells [18] encode temporal offsets directly. We present this as convergent evidence consistent with the framework, not as proof of it, relying only on the well-supported hierarchical-temporal organization.

4.4. Pearl's ladder as physical necessity

Pearl's levels [3] — association $P(Y|X)$, intervention $P(Y|do(X))$, counterfactuals $P(Y_x|X',Y')$ — are here levels with increasing dimensional and energetic demands. Association needs only a spatial snapshot. Intervention requires temporal structure (act at t_1, observe at $t_2 > t_1$), hence a 4D representation; this is why a purely spatial representation cannot distinguish correlation from causation. Counterfactuals require two divergent 4D histories simultaneously, prohibitive without hierarchical abstraction.

Proposition 2 (Hierarchical necessity of causal levels). *Level 1 requires only 3D spatial structure; Level 2 requires 4D spacetime; Level 3 requires hierarchical abstraction to remain tractable; and energy costs scale $E_1 < E_2 \ll E_3$.*

Proof. Level 1 follows from Bayes' rule on a spatial joint distribution. Level 2 follows because an intervention at t_1 affecting an outcome at t_2 demands temporal ordering. Level 3 follows because simulating two complete particle-level histories is energetically impossible; only hierarchical abstraction over macroscopic variables makes it tractable. The ordering follows from $E_1 \approx$ measurement only, $E_2 \approx$ measurement plus physical action, $E_3 \approx$ two interventions plus a dominating forward simulation. ■

5. THE MULTI-AGENT INTERFERENCE INTEGRAL

We formalize how other goal-directed agents impede an agent's path through configuration space. We use the term causal field for descriptive convenience only; the construction is entirely classical and shares with quantum field theory nothing but mathematical vocabulary. The relevant analogy is the classical electromagnetic field — real and measurable at each point — not the quantum operator field.

A goal ω_g induces a field $g(\omega) = -\nabla\omega\ V(\omega, \Omega g)$, with V an entropy or free-energy distance to the goal. For N agents pursuing distinct goals, the causal density $\rho(x,t)$ measures crowding under a light-cone (Heaviside) constraint and distance decay. Agent A's accumulated interference along path $\mathbb{P}$ is the work integral

$$I(\mathbb{P}) = \int_{\mathbb{P}} \rho(x) \cdot \omega(x) \cdot (1 - \cos\theta(x))\ ds \qquad (1)$$

where θ is the angle between the aggregate other-agent direction and A's intended direction. The factor $(1 - \cos\theta)$ ranges from 0 (alignment, no resistance) through 1 (orthogonal) to 2 (direct opposition). The density and weight kernels are normalized so that $I(\mathbb{P})$ is a dimensionless, non-negative resistance measure; this is what allows it to enter the Comprehension Factor as the dimensionless term $1 + I(\mathbb{P})$.

Proposition 3 (Interference properties). *$I(\mathbb{P}) \geq 0$ for all paths; $I(\mathbb{P}) = 0$ iff $\rho = 0$ along the path or $\theta = 0$ everywhere; for fixed endpoints I is minimized through low-density or high-alignment regions; and I is additive over disjoint segments.*

Proof. Non-negativity and the zero condition follow from $(1 - \cos\theta) \geq 0$ with equality only at $\theta = 0$. The minimization property follows from the variational condition $\delta I/\delta\mathbb{P} = 0$, yielding density-weighted geodesics in configuration space. Additivity follows from linearity of integration. ■

Unlike a quantum amplitude, $I(\mathbb{P})$ is directly observable: in traffic ρ is vehicle count and θ the directional-flow conflict; in markets ρ is competitor count per segment; in ecology ρ is population density. Measurability at every point is what makes the construction empirically testable rather than merely interpretive.

5.1. Quantum effects and classical intelligence

Tegmark [25] estimates decoherence times for the relevant neural and microtubular degrees of freedom in the range of roughly 10^{-13} to 10^{-20} s, whereas intelligent decisions occur on timescales of roughly 10^{-3} to 10^{-1} s (neural spikes ≈ 1 ms; bacterial tumbling ≈ 100 ms). The separation is many orders of magnitude, so by decision time coherence is long destroyed and intelligence operates firmly in the classical regime; the general decoherence framework is Zurek [24]. Goals are macroscopic configurations, and CF measures the probability of achieving one — interpretation-independent across Copenhagen, Many-Worlds, de Broglie–Bohm, and QBism. Quantum computers do not change this; speedup affects efficiency, not the structure of intelligence. We rely on no quantum-mind hypothesis: bacterial chemotaxis achieves goal-directed navigation with classical biochemistry, which is the relevant existence proof.

6. THE EMERGENCE ARGUMENT

We argue intelligent causal agency is a strongly favored — not logically necessary — outcome under generic conditions. We use "argument" and "■" rather than "theorem" and "Q.E.D." where the conclusion is a probabilistic limit; this is a deliberate calibration of claim strength.

Emergence Argument. *Consider a spacetime region satisfying (C1) thermodynamic disequilibrium — a persistent free-energy gradient; (C2) material complexity — matter able to form feedback loops faster than the environment changes; and (C3) causal transparency — information propagates within ordinary light cones. Then intelligent causal agency ($CF > \varepsilon$) arises with probability approaching 1 as organizational complexity C exceeds a finite threshold C_{min}.*

6.1. Step 1: thermodynamic necessity of modeling

Proposition 4. *Sustained local entropy reduction requires energy expenditure whose efficiency equals the system's causal-prediction accuracy.*

Derivation sketch. Maintaining S_{system} below equilibrium requires exporting entropy at minimum power $T\cdot|dS_{system}/dt|$ (the Carnot limit). With intervention success probability p, actual expenditure is the minimum divided by p, so efficiency equals p. Sustained entropy reduction therefore rewards accurate causal modeling — persistence requires efficiency, efficiency requires prediction, prediction requires a model.

6.2. Step 2: geometric necessity of 4D representation

Proposition 5. *Accurate causal modeling of complex goals requires an internal representation preserving four-dimensional spacetime structure.*

Proof. Special relativity restricts causes to the past light cone. Complex goals involve converging causal chains whose intersection is specified by both when and where they meet. The projection discarding spatial structure provably loses spatial-convergence information [1], unrecoverably; representations preserving 4D structure, obtained naturally through embodied sensors and actuators, are required. ■

6.3. Step 3: the optimization imperative

Proposition 6. *Under C1–C3, configurations with higher causal-prediction accuracy are thermodynamically favored.*

Derivation sketch. A system at accuracy 0.8 needs only about 60% of the energy a system at 0.5 needs for equivalent entropy maintenance, so it persists longer or acquires energy faster. This generates an optimization pressure — not necessarily Darwinian, though that is one mechanism. Material complexity (C2) supplies the feedback loops enabling refinement of models before conditions change; that refinement is learning.

6.4. Step 4: genericity

Proposition 7. *Conditions C1–C3 are generic in astrophysical environments, so intelligent agency is a likely outcome rather than a rare accident.*

Argument. C1 holds on essentially any planet receiving stellar radiation. C2 is conjectured to be reachable by ordinary chemistry given energy and geological time, though the prebiotic route to sufficient complexity remains an open problem. C3 holds for all non-exotic spacetimes. Given C1–C3 and complexity above C_{min}, a fluctuation-plus-drift dynamics is conjectured to reach high-CF attractor basins with probability approaching 1 as C grows. We state this as a plausibility argument, not a derivation: the "probability approaching 1" claim is asserted on the analogy that intelligence is expected as heavy-element nucleosynthesis is expected given sufficient stellar conditions, not proven, and is not logically guaranteed in any finite region or finite time. Establishing it rigorously is open work.

The claim's scope is therefore "intelligence is thermodynamically favored and generically expected," not "metaphysically inevitable." The difference is the empirical content: prediction P1 in the falsifiable-prediction section — that systems below C_{min} do not reliably exhibit CF > 0.5 — can fail.

7. THE COMPREHENSION FACTOR

CF operationalizes the Causality Hierarchy. For a causal chain $\Pi = \omega p \rightarrow \cdots \rightarrow \omega g$, each transition probability is retrieved from the CH at an appropriate abstraction level. Without a CH there is nothing to query; without CF the CH has no operational meaning. The transition from present to goal occurs through interventions c_i producing effects in context θ. The core term is the predicted joint success probability

$$\prod_i P(e_{i+1} \mid do(c_i), \theta_i) \quad (2)$$

where do(c) denotes a deliberate intervention [3]. The product form reflects a Markov assumption; conditioning on do(c) distinguishes active intervention from passive observation. This bare product, however, is incomplete: a causal chain can succeed mechanically yet still fail the goal if the outcome lands in the wrong place or time, if competing agents resist the path, or if the agent has descended needlessly into a low, energetically expensive layer of the Causality Hierarchy. The full Comprehension Factor scales the core product by a spatial–temporal alignment term and a hierarchical-efficiency term, and discounts it by multi-agent interference:

$$CF = [\, \prod_i P(e_{i+1} \mid do(c_i), \theta_i) \cdot STP \cdot H \,] \,/\, [\, 1 + I(\mathbb{P}) \,] \quad (3)$$

Here STP ∈ (0,1] is the spatial–temporal penalty (how well the achieved configuration aligns in place and time with the intended goal; defined below), H ∈ (0,1) is the hierarchical-efficiency or H-Factor (defined below), and $I(\mathbb{P})$ is the multi-agent interference integral introduced earlier. $I(\mathbb{P})$ is taken in normalized, dimensionless form so that $1 + I(\mathbb{P})$ is well defined; multiplying the numerator by H and STP, both dimensionless and in (0,1], leaves CF dimensionless and bounded. CF remains in [0,1]; for uniform per-step probability p the core product decays as p^n, which is exactly why complex goals require hierarchical decomposition into reliable sub-goals; CF depends on environmental context, making it an objective measure; and genuinely intelligent agents exhibit predicted CF ≈ actual CF. On the apparent circularity — CF is derived, not definitional: intelligence is defined first (Definition 2) and CF then measures it, as temperature is defined thermodynamically and then measured by a thermometer.

7.1. *The spatial–temporal penalty (STP)*

A causal chain can succeed mechanically and still miss the goal if the effect lands off-target in space or time. STP scales CF by how well the achieved configuration aligns with the intended one, treating misalignment as graded rather than a binary failure — close enough may still be good enough, and how close is "enough" is task-dependent. It is the product of a spatial and a temporal factor.

The spatial factor is a Gaussian centred on the target location, peaking at 1 at the intended coordinates and falling off with distance at a rate set by per-axis tolerances:

$$SP = \exp\left(-\left[(x-\mu_x)^2/(2\sigma_x^2) + (y-\mu_y)^2/(2\sigma_y^2) + (z-\mu_z)^2/(2\sigma_z^2) \right] \right) \quad (4)$$

where (μ_x, μ_y, μ_z) is the target location, (x, y, z) the achieved one, and σ the tolerated spatial scale per axis (the Gaussian standard deviation, not the variance, which would be σ^2) — small σ for precision tasks (a vaccine injection, millimetres), large σ for loose ones (parking near an office, metres). The temporal factor is a window built from two logistic terms. Writing t for the predicted time of effect, [a, b] for the acceptable window (a = earliest, b = latest acceptable time), and τ_m, τ_p for the early- and late-side sharpness, define the two sigmoids

$$L_a = 1 / (1 + \exp[(t - a)/\tau_m]), \quad L_b = 1 / (1 + \exp[(t - b)/\tau_p]) \quad (5)$$

The temporal penalty is their difference, a smooth window that is near 1 inside [a, b] and falls off on either side:

$$TP = L_b - L_a \quad (6)$$

The asymmetry is real and intended — τ_m and τ_p need not be equal (arriving hours early for a flight is fine; minutes late is not). STP is the product SP · TP, bounded in (0,1], equal to 1 only at perfect spatial and temporal alignment. Like the H-Factor, STP is presented here in its structural form; its tolerance parameters (σ, τ_m, τ_p, and the target window) are unfitted in this preprint, stated as the form the full theory requires with calibration left to the fuller treatment [2] and future empirical work. STP is not load-bearing for any quantitative claim made here — the worked chemotaxis example uses only the core product — but it is included because, like the interference and H-Factor terms, it is part of the Comprehension Factor and omitting it would present an incomplete equation as a complete one.

7.2. *The hierarchical-efficiency (H-Factor) term*

The H-Factor term encodes a thermodynamic fact already argued in the Causality Hierarchy section: descending into a lower layer of the Causality Hierarchy costs more energy per prediction. It is defined as

$$H = 1 / (1 + e^{\alpha(B - L - \delta E)}) \quad (7)$$

with H ∈ (0,1), where B is the agent's baseline (competent) causal layer for the domain, L is the layer it is currently working at, E ∈ [0,1] is its experience at that depth, and α, δ are sensitivity constants. The mechanism is the point: when the agent works at or above its baseline (L ≥ B), the exponent is zero or negative, the term approaches 1, and CF is held high — the agent is operating at an efficient, predictable abstraction. When it drops below baseline into unfamiliar depth (L < B with low E), the exponent grows, the denominator inflates, and CF is driven down. The H-Factor does not penalize abstraction; it keeps CF high precisely while the agent stays at the right level of the hierarchy and collapses it when the agent descends unnecessarily. This is the formal counterpart of the energetic argument in Proposition 1 and is consistent with the predictive-coding and conflict-monitoring evidence cited there.

The H-Factor is not an engineering convenience; it follows from three commitments the framework already makes. First, energy conservation. The Causality Hierarchy can be read as an inverted pyramid whose tip is the most fundamental description and whose every higher layer is a configuration produced by more intersecting causal chains integrated over larger spacetime volume. This vertical axis is an emergence coordinate: alongside the three spatial and one temporal coordinate of a configuration, it records the degree of causal organization at that location — operationally, the size of the largest causally integrated system for which the point is an interior member. The coordinate is not a postulated extra spatial dimension; it is read off causal volume and respects locality and the light-cone constraint exactly as the rest of the framework does. By Proposition 1, predicting low on this axis (deep, near the particle tip) costs energy that grows exponentially in the number of micro-causal steps, while predicting at a higher, more compressed emergence layer is cheap. Operating below the layer a goal requires therefore burns energy a thermodynamically open agent cannot afford. The H-Factor is the quantitative expression of that asymmetry: it is low exactly where energy cost is high. Second, cause-effect learning. The baseline B and experience term E are not free settings but learned quantities — they record how well the agent's acquired cause-effect models support efficient prediction at a given emergence layer. An agent that has learned a domain has, in effect, raised its baseline; an agent forced below what it has learned predicts poorly and expensively. H thus couples the static CF machinery to learning: experience is what moves B, and movement in B is what learning is. Third, the Causality Hierarchy itself. The emergence-coordinate pyramid and the Causality Hierarchy are the same structure viewed from two angles — one emphasizing how degree of causal integration rises with abstraction, the other how cause-effect relations are organized for query — so the H-Factor introduces no new ontology; it makes explicit the energetic cost already implicit in operating within the hierarchy the paper has used throughout. (The emergence coordinate is developed at length in [2]; here it is used only in its operational, locality-respecting sense and the paper depends on nothing beyond causal integration and energy.)

The H-Factor must not be read as forbidding descent into lower causal layers. It penalizes only unnecessary descent — operating deep when a reliable higher abstraction already exists. When no such abstraction exists, during novelty, failure, troubleshooting, or innovation, descent is exactly what intelligence is required to do. In that case the temporary drop in H is not an error; it is the correct energetic accounting of exploration. A successful descent discovers causal structure that did not previously have a compressed representation, which raises the agent's experience term and, with repetition, lifts its baseline — the same B-movement identified above as learning. The new abstraction then restores a high H for every future encounter with the same problem. So the H-Factor is not a punishment term but an energy-accounting term: using an existing abstraction conserves energy and keeps H high; building a missing one requires spending energy at temporarily low H, after which the resulting abstraction makes the expenditure non-recurring. This is, in compressed form, the physics of learning and invention — and it is why the same term that rewards efficiency does not thereby penalize discovery.

We include the H-Factor because it closes the CF equation: presenting the core product alone, while retaining the interference term $I(\mathbb{P})$, would show an incomplete definition rather than a simplified one. We do not, however, claim empirical weight for its parameters. The constants α, δ, the baseline B, and the experience term E are unfitted in this preprint — they are stated as the structural form the full theory requires, with calibration and any falsifiable test left to the fuller treatment [2] and future empirical work. The worked chemotaxis example exercises only the core product through a task-level information-efficiency proxy, with no hierarchy depth or multi-agent interference in play; the H-Factor is not load-bearing for any quantitative claim made here.

8. GOAL-NETWORK MANAGEMENT: THE TEMPORAL COMPREHENSION FACTOR

The Comprehension Factor as defined above measures one causal chain reaching one goal. No known embodied living agent, however, operates with a single isolated goal. A bacterium simultaneously seeks nutrients, avoids toxins, conserves energy, and maintains membrane integrity; a person never has "a goal" but a shifting network of them. The single-goal case is therefore the degenerate $n = 1$ instance of the general situation, not the base case from which intelligence is built. The object intelligence actually operates on is a network of causally linked goals distributed across time, and the defining act of intelligence is the management of that network: deciding which goals to pursue, which to delay, and which to abandon, because an action that secures one goal can drive a downstream goal's CF to zero, and when the first step of a future chain fails the whole chain collapses. Per-goal optimization that ignores this is the formal signature of short-termism — winning a battle while losing the war.

The network-level quantity is the Temporal Comprehension Factor (TCF). It is not a passive product of pre-existing goal scores: at each step the agent faces a choice set of available goals, paths, and interventions, and the selected choice rewrites the causal-field context in which every subsequent choice is evaluated. Writing π for a policy

— the sequence of selections π_t made at each step — and θ_t for the context at step t, TCF is a functional of the policy:

$$TCF(\pi) = \prod_{t=1}^{n} CF_{t,\pi t}(\theta_t), \quad \theta_{t+1} = F(\theta_t, \pi_t) \tag{8}$$

Here π_t is the goal, path, or intervention selected from the available set at step t; $CF_{t,\pi t}(\theta_t)$ is the Comprehension Factor of that selection evaluated in the prevailing context; and $\theta_{t+1} = F(\theta_t, \pi_t)$ states that the selection itself transforms the context — the spacetime configuration on which the next step's CF landscape depends. This is the formal content of the claim that a high-CF route to one goal can lower the CF of others: the path is an input to F, not a passive read of a fixed field. Intelligence is the selection of the policy that maximizes network survival:

$$\pi^* = argmax_{\pi}\, TCF(\pi) \tag{9}$$

Single-goal CF is the degenerate n = 1 case of TCF(π) with a singleton choice set. The functional form has a sharp consequence: any selection that drives a linked CF toward zero collapses the whole product, regardless of how high that selection's own CF is. This is what makes goal-dropping a first-class operation rather than an exception — an agent maximizing TCF(π) must, on detecting that retaining a goal lowers it, drop or defer that goal even when it is individually achievable. Intelligence is not the pursuit of goals; it is the selection of a policy through a changing goal network under collapse risk, and the same machinery that scores a single selection (CF in context) drives both which goals are kept and which path is taken to the kept ones.

A worked illustration makes the structure concrete. Suppose an agent's timeline contains the linked goals: obtain coffee, arrive at a mandatory meeting on time, retain employment, receive a paycheck, pay rent. If pursuing the first goal compresses the schedule — CF(coffee) = 1.0 but CF(on time) = 0.2, propagating to CF(keep job) = 0.5, CF(paycheck) = 0.5, CF(rent) = 0.5 — then TCF = 1.0 × 0.2 × 0.5 × 0.5 × 0.5 ≈ 0.025. Skipping the first goal raises the chain to CF(on time) = 0.8, CF(keep job) = 1.0, CF(paycheck) = 1.0, CF(rent) = 0.9, giving TCF ≈ 0.72. The first goal's isolated CF is 1.0 in both cases; it is nonetheless the wrong goal to prioritize, and TCF — not CF — is what reveals this.

Dropping a goal is only the coarsest network-management move. The context recursion $\theta_{t+1} = F(\theta_t, \pi_t)$ above already carries the subtler one: because the path taken to one goal is itself an input to F, a high-CF route to goal A can lower the CF of goals B, C, D by reshaping the context in which they are evaluated, even when A itself succeeds and is retained. The intelligent response is then not to drop A but to reach it by a lower-CF path — to spend tactical efficiency to protect strategic position. This is the division the formalism makes precise: CF is the tactical question of which path most reliably achieves a single goal, while TCF is the strategic one of how that path reshapes the field the rest of the network depends on. A competent agent accordingly accepts a locally suboptimal CF to keep TCF high, and declines the globally cheapest plan when its side effects propagate through θ into the network.

This also explains why goal importance need not be assigned. Conventional planners rely on hand-tuned utility values or predefined reward functions to encode what matters. Under network management, importance is not declared but revealed by causal structure: a goal is critical precisely when failing it cascades into the failure of many downstream goals — its weight is its position in the collapse structure, not an intrinsic number. This is why the framework is adaptive and general where static reward functions are brittle, and it is, in our view, the most consequential difference between this account and utility-based formulations: the quantity an intelligent agent protects is the survival of its goal network, and that quantity supplies its own priorities.

TCF carries its own scope limit, which we state rather than hide. A TCF chain is only as reliable as the confidence values entering it. Projecting long causal chains through an uncertain environment is speculative: a first-year student cannot meaningfully estimate the CF of an action four years downstream, not because the distal goal is invalid but because the near links are being asked to predict more than the environment supports. Competent agents calibrate the length of their TCF chains to the stability of the causal terrain — long chains where it is predictable, short chains where it is not. This calibration is part of the construct, not a patch on it. We are correspondingly explicit about epistemic standing: goal-network management is the conceptual core of this account, but the empirically anchored part is the per-goal machinery — CF and the worked chemotaxis application — together with prediction P5, which tests whether agents that evaluate plans by their network-level TCF outperform per-goal-CF maximizers in environments with delayed consequences. We therefore do not attach an independent falsification criterion to TCF itself; its empirical content is exercised through P5, and the network-management thesis is presented as the organizing claim whose quantitative validation is the open research program, not as an already-measured result.

One consequence of this account deserves to be named, though its development lies outside this paper. If affect is taken not as introspective feeling but as the agent's relation to the state of its own goal network, several states usually treated as ineffable acquire structural definitions. Desperation is the regime in which TCF is falling and no available plan offers a high per-goal CF, yet some action must be taken — the agent commits to low-probability paths because no better one exists. Hopelessness is the limiting case in which TCF approaches zero: no path with non-vanishing product remains. Empathy, structurally, is the extension of the network to include other agents' goals, so that their goal-failures enter one's own network product and reshape one's CF landscape accordingly. The substantive point is one of structure, not analogy: because these are regimes of the network variable rather than additional faculties, any system that genuinely implements goal-network management may exhibit structural analogues of these states, independent of phenomenal experience — so the question raised by machine intelligence is not whether affect-like structure can be added but whether, in a system that manages a goal network at all, it is already implicit. Nor do we develop the affective state space, which spans far more than the three states named. That development — a systematic treatment of the emotions as readouts of network state — is deferred to dedicated future work. It is noted here only because the bridge from the formalism to affect is short and direct, and leaving it entirely unstated would misrepresent how far the network account reaches. We make no claim here about phenomenal experience: the framework is committed to the functional and structural identity of these regimes, not to whether they are accompanied by subjective feeling, and the hard problem of consciousness is deliberately left untouched. Whether phenomenal experience requires constraints beyond those specified here is outside the scope of this paper.

9. A WORKED APPLICATION: BACTERIAL CHEMOTAXIS

We analyze Escherichia coli chemotaxis as a worked, quantitative instance at molecular scale. The goal — reaching higher nutrient concentration — satisfies Definition 1. An E. coli cell (≈2 μm) is too small to read a spatial concentration difference directly and must infer the gradient from temporal measurements while swimming.

Causal modeling: receptor methylation (CheR/CheB on Tar, Tsr, Trg, Tap) implements a molecular memory encoding a recent-history baseline; the difference between current occupancy and baseline encodes d[attractant]/dt. Forward simulation and selection: positive derivative → low CheY-P → counterclockwise rotation → smooth swimming (continue); non-positive → elevated CheY-P → clockwise rotation → tumble (re-sample). Physical execution: the ion-motive-force–powered flagellar motor switches rotational direction within milliseconds (the work per rotation is of order 10^{-17} J as an order-of-magnitude estimate, not a measured per-reversal constant).

We first state how CF is operationalized for this system. The quantity CF is intended to capture how efficiently the chemotaxis system converts acquired causal information into up-gradient progress under shallow-gradient conditions. We do not claim a first-principles derivation that any particular observable equals this probability; instead, for this worked example we adopt an explicit *normalized information-efficiency proxy* — the fraction of the task-relevant information-theoretic performance bound achieved by the observed up-gradient drift — and treat the resulting figure as an estimate, not a measured constant. Empirical chemotaxis performance is gradient-dependent, so there is no single universal drift constant; we therefore anchor CF to a measured information-efficiency rather than a raw motion fraction. Berg and Brown [7] established the run–tumble structure of E. coli chemotaxis. Modern single-cell measurements (Mattingly et al. [26]) report a chemotactic drift coefficient $\chi \approx 4300$ μm²/s with up-gradient drift speed $v\_d = \chi g$, giving $v\_d \approx 0.4$–1.7 μm/s over shallow gradients ($g \approx 0.1$–0.4 mm^{-1}) against a swimming speed $v_0 \approx 22.6$ μm/s — a raw drift/speed ratio of order 10^{-2}. The same study reports a chemotactic information-efficiency of 0.65 ± 0.05: the cell climbs at about 65% of the theoretical bound set by the information its receptor–kinase system actually acquires. This information-efficiency, not the raw drift fraction, is the physically appropriate operationalization of CF for the shallow-gradient chemotaxis task: it measures how efficiently the cell converts acquired causal information into goal-directed drift:

$$CF_1 \approx \eta\_info \approx 0.65 \pm 0.05 \ \ \text{(Mattingly et al. [26])}, \quad (10)$$

i.e. an information-efficiency of order 0.65 for the shallow-gradient chemotaxis task. The qualitatively important point is robust to the exact figure: the cell's molecular machinery converts a shallow chemical gradient into reproducible directional bias at a substantial fraction of the information-theoretic bound, rather than unbiased diffusion. We use this measured efficiency as the chemotaxis benchmark for CF; extending the same information-limited and feedback-control treatment to neural, collective, and plant systems is the natural next step rather than one carried out here. A subsequent study from the same group [33] reports that, against the physical Berg–Purcell limit set by stochastic molecule arrivals, E. coli operates far below that bound; this does not conflict with the figure used here, which is efficiency relative to the information the signaling pathway actually acquires, not relative to the molecule-counting limit. The two bounds answer different questions, and the present argument depends only on the former.

Over a long traverse the naive product of per-step values decays rapidly (e.g. 0.6 raised to ~50 cycles is vanishingly small), which would wrongly predict that bacteria cannot reach distant sources. They do, because the operative per-step goal is local ("continue while improving"), not the distal target; the per-step bias is renewed each cycle rather than multiplied. Berg's tracking shows reliable accumulation at the source over minutes — a concrete instance of the CF decay property established for the Comprehension Factor and its hierarchical remedy, not a refutation of it.

The emergence conditions are met qualitatively: C1, nutrient gradients supply free energy (the cell sustains active metabolism, with ATP consumption of order 10^6–10^7 ATP per cell per second as a benchmark estimate); C2, receptor–methylation feedback adapts on a seconds timescale, faster than typical gradient change; C3, chemical diffusion obeys ordinary causality. The system thus performs goal-directed navigation with zero neurons and no anthropocentric computation — the existence proof the framework needs at molecular scale. Neural, collective-insect, and plant/fungal analyses appear in the book-length treatment [2] and a companion empirical paper in preparation; they are omitted here so every result in this preprint is independently checkable. Modern feedback-control and information-limited treatments of E. coli chemotaxis would strengthen this section and are noted as the natural next citations.

10. EXPERIMENTAL PREDICTIONS

Each prediction states a quantitative expectation and an explicit failure criterion.

P1. No system with material complexity $C < C_{min}$ exhibits $CF > 0.5$; an order-of-magnitude benchmark estimate places C_{min} at 10^2–10^3 components at $T \approx 300$ K (derived in Appendix A as an estimate, not a measured constant). Test: minimal synthetic cells of varying component count in microfluidics. Falsification: a sub-threshold system consistently achieving $CF > 0.5$.

P2. CF for organisms navigating gradients should fall in characteristic taxon-specific ranges correlating with ecological success (chemotaxis of order 0.6, illustrated in the worked application as an initial benchmark). Falsification: simple-model organisms consistently above 0.9, or complex-model organisms stuck below 0.3.

P3. In traffic networks, $I(\mathbb{P})$ computed from density and directional conflict should predict travel time, with a target benchmark of $R^2 > 0.7$ after controlling for path length. Falsification: no meaningful predictive relationship (e.g. $R^2 < 0.3$), or low-interference paths systematically slower.

P4. Under selection for goal achievement, mean population CF should increase across generations with positive slope. Falsification: CF decreasing while task performance increases, or plateauing below 0.5 for competent agents.

P5. For long causal chains ($N > 50$), hierarchical representations should beat flat ones by roughly an order of magnitude (for $p = 0.9$, flat $\approx 0.9^{50} \approx 0.005$ vs hierarchical $\approx 0.9^5 \approx 0.59$). Equivalently, agents that select plans by their policy-level Temporal Comprehension Factor — the product of CFs along the chosen path through the goal network — should outperform agents that maximize per-goal CF, because the latter accept delayed consequences that collapse downstream chains. Falsification: equivalent performance between flat and hierarchical representations for $N > 50$, or no advantage for TCF-evaluating agents over per-goal-CF maximizers in environments with delayed consequences, directly refutes the hierarchical-decomposition argument.

P6. Cross-domain: model-based market participants should outperform model-free ones on a risk-adjusted basis; rostral-prefrontal damage should impair long-horizon planning disproportionately; current language models should show low CF on physical-intervention tasks, improvable with interventional training data.

Broad confirmation would support universality; failure in any domain requires revising the framework or showing the domain violates C1–C3.

11. ANTICIPATED OBJECTIONS

(1) ***Entropy reduction violates the Second Law.*** Intelligent systems are open: they reduce local entropy by exporting more entropy to the environment, so total entropy still rises. Perpetual-motion machines fail precisely because they cannot pay this entropy tax; intelligence pays it.

(2) ***The definition applies to crystals and tornadoes.*** Crystals reduce local entropy but lack feedback control, internal causal modeling, and intervention selection (the C2-type properties); they do not lack causal transparency (C3), which holds for any ordinary physical system. Tornadoes dissipate gradients along least-resistance paths, pursuing no goal and maintaining no model. Crystallization is relaxation; intelligence is feedback-driven defiance of relaxation.

(3) ***You tie intelligence to entropy, but it is substrate-independent.*** Substrate independence is not physics independence. Entropy is universal; a goal as a low-entropy configuration is substrate-independent because it is defined by physics, not the implementing material. Multiple implementations reach the same goal; none escapes thermodynamics.

(4) ***You conflate intelligence with consciousness.*** We explicitly do not. Intelligence here is causal manipulation toward goals; phenomenal experience is untouched. Bacterial chemotaxis shows intelligence with no plausible experience; impaired patients may have experience with little intelligence. The two are treated as orthogonal.

(5) ***Too reductionist — it misses emergence.*** The Causality Hierarchy is the account of emergence: higher layers are genuine, causally efficacious abstractions arising from energy minimization, not redescriptions. Emergence here is efficient compression of causal structure, not irreducibility in principle.

(6) ***Dimensional Insufficiency is unfamiliar — can the framework stand without it?*** The practical claims rest on relativity and thermodynamics alone. Dimensional Insufficiency strengthens the geometric argument but is not load-bearing for the falsifiable predictions; a skeptic may treat Proposition 5 as assuming temporal structure is fundamental, which relativity already grants.

(7) ***Bacteria are just chemistry.*** Chemistry structured to model causality, predict, and intervene toward a goal — which is exactly the claim. One must still explain why that chemistry reliably converts a shallow gradient into reproducible up-gradient navigation. Dissolving the bacterium/brain distinction is a feature.

(8) ***Goals are anthropomorphic.*** Goals here are thermodynamic, not psychological: a future low-entropy configuration maintained by measurable energy expenditure. A bacterium “wants” nutrients as a ball “wants” to roll downhill — both follow physical gradients — except the bacterium actively models and intervenes.

12. CONCLUSION

We have presented a physics-first framework in which intelligence is a thermodynamically favored mode of organization for matter that maintains low-entropy structure far from equilibrium. Its central thesis is that intelligence is not the maximization of any single objective but the management of a network of causally linked goals under collapse risk: because no known embodied living agent operates with a single isolated goal, the per-goal Comprehension Factor is the degenerate case, and the defining act is the continuous decision to pursue, defer, or abandon goals so that local success does not destroy the network. The framework derives the four operations of intelligent agency from energy and geometry, quantifies the per-goal measure through the Comprehension Factor and the network-level quantity through its product (TCF), applies the per-goal measure in a fully classical molecular example, and commits to falsifiable predictions with explicit failure criteria. Its status is a research program: the conceptual core is goal-network management, the empirically anchored content is the per-goal machinery and the falsifiable-prediction section, several predictions are testable with existing methods, and the framework is wrong in its current form if sub-threshold systems reliably exhibit high CF, if flat and hierarchical representations perform identically on long chains, if TCF-evaluating agents show no advantage over per-goal-CF maximizers under delayed consequences, or if the interference integral fails to predict path difficulty. That exposure to refutation is the point.

APPENDIX A. COMPLEXITY THRESHOLD

A system maintaining low entropy satisfies $dS_{total}/dt \geq 0$, exporting entropy at minimum power $T\cdot|dS_{system}/dt|$. With success probability p, efficiency equals p. Achieving p above chance needs an internal model requiring at least $k_B \ln N$ of representational entropy for N states, instantiated at a Landauer cost per bit. Combining energy, information, and feedback-speed constraints gives $C_{min} \approx \max\{(T\sigma)/(p\cdot E_{avail}), M \log N, k\cdot\tau_{env}\}$; for typical planetary conditions this is 10^2–10^3 distinct molecular species — the basis for Prediction P1.

APPENDIX B. EMERGENCE PROBABILITY

Configuration space has measure scaling as exp(N). Configurations with CF > ε waste less energy and persist longer, defining a fitness $f(\omega) = CF(\omega)\cdot\tau_{persist}(\omega)$. Under fluctuation-plus-drift dynamics, high-CF configurations form attractor basins whose volume grows with complexity, so $P(CF > \varepsilon \mid C1,C2,C3) \rightarrow 1$ as $C \rightarrow \infty$. This is an argument, not a theorem: it establishes intelligence as the expected attractor, not a finite-time guarantee. The empirical commitment remains Prediction P1.

REFERENCES

[1] M. Shita, Dimensional Insufficiency: Why Logic Cannot Capture Physical Emergence, Global Economic Alliance preprint, 2025.
[2] M. Shita, Science of Intelligence: From Spacetime to Superintelligence, EraNova Publishing, 2024.
[3] J. Pearl, Causality: Models, Reasoning, and Inference, 2nd ed., Cambridge Univ. Press, 2009.
[4] A. Einstein, Zur Elektrodynamik bewegter Körper, Ann. Phys. 322(10), 891–921, 1905.
[5] P. S. Churchland, Brain-Wise: Studies in Neurophilosophy, MIT Press, 2002.
[6] E. Schrödinger, What Is Life?, Cambridge Univ. Press, 1944.
[7] H. C. Berg and D. A. Brown, Chemotaxis in Escherichia coli analysed by three-dimensional tracking, Nature 239, 500–504, 1972.
[8] I. Prigogine, Self-Organization in Nonequilibrium Systems, Wiley, 1977.
[9] J. O'Keefe and J. Dostrovsky, The hippocampus as a spatial map, Brain Res. 34, 171–175, 1971.
[10] D. J. Foster and M. A. Wilson, Reverse replay of behavioural sequences in hippocampal place cells, Nature 440, 680–683, 2006.
[11] R. Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev. 5(3), 183–191, 1961.
[12] J. L. Deneubourg et al., The self-organizing exploratory pattern of the Argentine ant, J. Insect Behav. 3, 159–168, 1990.
[13] K. Friston, The free-energy principle: a unified brain theory?, Nat. Rev. Neurosci. 11(2), 127–138, 2010.
[14] S. A. Kauffman, The Origins of Order, Oxford Univ. Press, 1993.
[15] A. Tero et al., Rules for biologically inspired adaptive network design, Science 327, 439–442, 2010.
[16] E. Koechlin and C. Summerfield, An information theoretical approach to prefrontal executive function, Trends Cogn. Sci. 11(6), 229–235, 2007.
[17] M. M. Churchland and K. V. Shenoy, Temporal complexity of single-neuron activity in premotor and motor cortex, J. Neurophysiol. 97(6), 4235–4257, 2007.
[18] C. J. MacDonald et al., Hippocampal time cells bridge the gap in memory for discontiguous events, Neuron 71(4), 737–749, 2011.
[19] R. P. Rao and D. H. Ballard, Predictive coding in the visual cortex, Nat. Neurosci. 2(1), 79–87, 1999.
[20] M. M. Botvinick, Hierarchical models of behavior and prefrontal function, Trends Cogn. Sci. 12(5), 201–208, 2008.
[21] A. Clark, Whatever next? Predictive brains, situated agents, Behav. Brain Sci. 36(3), 181–204, 2013.
[22] J. B. Tenenbaum et al., How to grow a mind, Science 331(6022), 1279–1285, 2011.
[23] D. Wallace, Taking particle physics seriously, Stud. Hist. Philos. Mod. Phys. 42(2), 116–125, 2011.
[24] W. H. Zurek, Decoherence, einselection, and the quantum origins of the classical, Rev. Mod. Phys. 75(3), 715–775, 2003.
[25] M. Tegmark, Importance of quantum decoherence in brain processes, Phys. Rev. E 61(4), 4194–4206, 2000 (arXiv:quant-ph/9907009).
[26] H. H. Mattingly et al., Escherichia coli chemotaxis is information-limited, Nature Physics 17, 1426–1431, 2021.
[27] S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks, Thermodynamics of prediction, Phys. Rev. Lett. 109, 120604, 2012.
[28] A. D. Wissner-Gross and C. E. Freer, Causal entropic forces, Phys. Rev. Lett. 110, 168702, 2013.
[29] J. L. England, Statistical physics of self-replication, J. Chem. Phys. 139, 121923, 2013; Dissipative adaptation in driven self-assembly, Nat. Nanotechnol. 10, 919–923, 2015.
[30] E. Canessa, Comment on "Causal Entropic Forces", arXiv:1308.4375, 2013 (cf. E. Canessa, Physica A 341, 165, 2004).
[31] J. D. Murray et al., A hierarchy of intrinsic timescales across primate cortex, Nat. Neurosci. 17, 1661–1663, 2014.
[32] J. A. Miller and C. Constantinidis, Timescales of learning in prefrontal cortex, Nat. Rev. Neurosci. 25, 597–610, 2024.
[33] H. H. Mattingly, K. Kamino, J. Ong, R. Kottou, T. Emonet, and B. B. Machta, E. coli chemosensing accuracy is not limited by stochastic molecule arrivals, Nat. Phys. 22(1), 123–130, 2026.