BioSystems

Intelligence as High-Dimensional Coherence: The Observable Dimensionality Bound and Computational Tractability

--Manuscript Draft--

Manuscript Number:	BIOSYS-D-25-00880R1
Article Type:	Full Length Article
Section/Category:	
Keywords:	Intelligence, Observable Dimensionality Bound, Curse of Dimensionality, Landauer Limit, Phase Space, Thermodynamics, Vector Addition Systems, Constraint Geometry, Measurement Limits, Power Scaling
Corresponding Author:	Ian Todd University of Sydney School of Medical Sciences AUSTRALIA
First Author:	Ian Todd
Order of Authors:	Ian Todd
Abstract:	Intelligence must be high-dimensional. We show that, under minimal thermodynamic assumptions, any system capable of tracking and controlling high-dimensional targets faces exponential thermodynamic cost unless its substrate dimensionality matches or exceeds the target dimensionality. This applies to all living systems: bacteria tracking chemical gradients (D_target ~ 10^2-10^3, ~10^-12 W) and human brains tracking social/ecological complexity (D_target ~ 10^3-10^6, ~20 W) both require D_obs >> D_target, scaled to their respective behavioral bandwidths. Intelligence—the capacity to track and respond to environmental complexity—is a fundamental thermodynamic property of life, not unique to humans. We establish three results: (1) High-dimensional systems with irreducible complexity exist (e.g., vector addition systems with Ackermann-complete reachability; coupled oscillator networks with D_eff >> 1). (2) Theorem (Dimensional Tracking Bound): Tracking a system with dimensionality D_target using algorithmic inference from substrate dimensionality D_obs < D_target requires collision rate scaling as k^(D_target - D_obs) (exponential in dimensional mismatch), yielding thermodynamic power penalty up to 10^6x for realistic mismatches. (3) Biological intelligence escapes this bound by operating as a high-dimensional substrate (D_obs >> D_target) with collision-free constraint geometry, paying Landauer cost only at sparse behavioral outputs (~10^2 bits/s), not during internal computation. We derive an Observable Dimensionality Bound D_crit = C_obs * tau_e / (alpha * h_epsilon^track) showing when D_eff > D_crit, temporal microstructure becomes physically unmeasurable—systems compute faster than observation can track. Human cortex operates at D_eff^MEG / D_crit ~ 10^2 (substrate plausibly 10^3-10^4), explaining both power efficiency (collisions only at behavioral bottleneck) and computational tractability (VAS problems Ackermann-complete for discrete enumeration become tractable via continuous high-dimensional relaxation). This framework unifies biological efficiency with the thermodynamic impossibility of low-dimensional algorithmic intelligence.
Response to Reviewers:	

Cover Letter

Ian Todd
Sydney Medical School
University of Sydney
Sydney, NSW, Australia
itod2305@uni.sydney.edu.au

November 16, 2025

Dr. Abir Igamberdiev
Editor-in-Chief
BioSystems

Dear Dr. Igamberdiev,

I am pleased to resubmit my revised manuscript titled "Intelligence as High-Dimensional Coherence: The Observable Dimensionality Bound and Computational Tractability" (BIOSYS-D-25-00880) for reconsideration in *BioSystems*.

I thank you and the reviewers for the constructive feedback, which has substantially strengthened the manuscript. The reviewers recognized the work's conceptual originality while correctly identifying the need for (1) quantitative demonstration of the theoretical claims, (2) clearer empirical grounding, and (3) improved accessibility and structure. I have addressed all major concerns through extensive revisions.

Major Changes in Response to Reviewer Feedback:

1. **Numerical simulations added** (Reviewer 2's primary concern): I now provide four complete Python simulation codes demonstrating:
   - Collision-free computation in high-dimensional continuous systems vs. collision-heavy discrete enumeration
   - VAS scaling showing linear collision count ($\sim 4n$) for discrete vs. zero for continuous across dimensions $n = 2$ to 100
   - Spontaneous code formation through Hebbian pathway strengthening
   - Quantitative comparison validating the dimensional tracking bound

   All code is provided as supplementary material with full documentation and reproducibility guarantees.
2. **Empirical grounding and testable predictions** (Reviewer 2): I have:
   - Added concrete numerical example using MEG parcellation data showing cortex operates at $D_{eff}/D_{crit} \sim 10^2$ (Section 3.2)
   - Derived observable dimensionality bound $D_{crit} = C_{obs}\tau_e/(\alpha h_\varepsilon)$ with explicit parameter values
   - Provided specific predictions for coherence times, power scaling, and dimensional collapse signatures (Section 8)
   - Connected to measurable biological proxies including neural oscillations, MEG coher-

ence, and metabolic efficiency

3. **Structural improvements and accessibility** (Reviewer 1):
    - Reorganized abstract and introduction for clearer logical flow
    - Added explicit "Structure and thesis: What's new" section distinguishing my contributions from existing frameworks
    - Included summary sentences after key derivations (Theorems 1 & 2)
    - Clarified physical assumptions vs. derived implications throughout
    - Trimmed overall length while expanding critical explanations
4. **Methodological rigor and clarity**:
    - Made irreducibility assumptions explicit in Theorem 1
    - Added "Simulation Limitations and Assumptions" subsection
    - Distinguished measurable observables from theoretical constructs
    - Connected framework to Ashby's law of requisite variety, morphological computation, and reservoir computing
    - Expanded discussion of compressible vs. irreducible complexity

Title and Framing Changes:

I have retitled the manuscript to emphasize the *Observable Dimensionality Bound* as the central quantitative contribution. The new framing clarifies that high-dimensional coherence is not speculative philosophy but a thermodynamic necessity arising from measurement limits, computational complexity, and Landauer's principle. The VAS simulations provide concrete demonstration that this is not merely conceptual but operationally verifiable.

Preservation of Core Thesis:

While I have significantly strengthened empirical grounding and quantitative support, the manuscript retains its theoretical ambition: explaining *why* biological intelligence must be high-dimensional as a consequence of thermodynamics and information theory, not merely observing that it is. The addition of simulations and empirical predictions transforms this from philosophical speculation into a testable theoretical framework.

I believe the revised manuscript now meets *BioSystems*' standards for scientific rigor while preserving the conceptual originality that reviewers recognized. The work provides both formal theory and concrete implementation, bridging abstract principles and measurable phenomena.

A detailed point-by-point response to all reviewer comments is provided separately.

Thank you for the opportunity to revise and resubmit this work. I look forward to your decision.

Sincerely,

Ian Todd
Sydney Medical School
University of Sydney

Point-by-Point Response to Reviewers

Manuscript BIOSYS-D-25-00880

"Intelligence as High-Dimensional Coherence: The Observable Dimensionality Bound and Computational Tractability"

Ian Todd

November 16, 2025

We thank both reviewers for their thorough and constructive feedback. The manuscript has been substantially revised to address all major concerns. Below we provide detailed responses to each comment.

Reviewer 1

Comment 1.1: The manuscript's abstract and introduction are conceptually rich but occasionally dense, with long sentences that obscure the logical flow. For example, the phrase "thermodynamically admissible yet temporally unresolvable regimes" could be clarified by briefly stating what makes a regime "admissible" in this context.

Response: We have significantly revised the abstract and introduction for clarity and accessibility:

- **Abstract restructured**: Now opens with clear thesis statement ("Intelligence must be high-dimensional"), follows with three numbered results, and concludes with empirical grounding. Long sentences broken into digestible units.
- **Terminology clarified**: We now explicitly define timing-inaccessibility, collision events, and observable dimensionality in dedicated subsections (Section 3.1.3: "Precise definitions").
- **Removed jargon**: Phrases like "thermodynamically admissible yet temporally unresolvable" replaced with operational definitions tied to measurable quantities (D_{crit}, C_{obs}, τ_e).

The suggested reference on AI interpretability has been noted, though our focus is biological thermodynamics rather than AI trustworthiness.

Comment 1.2: In Section 2.1, the explanation of the temporal uncertainty relation would benefit from clearer transitions between physical assumptions and derived implications.

Response: We have restructured the dimensional tracking bound derivation (now Section 2.3) with:

- **Explicit assumption statements**: Theorem 1 now explicitly states "irreducible target system with effective dimensionality D_{target} (no sufficient statistic of dimension $< D_{target}$ exists for the tracking task)."
- **Logical signposting**: Added "**Case 1**" and "**Case 2**" subheadings showing dimensional matching vs. mismatch regimes separately.

- **Summary after proof**: "**Explicit piecewise form**" subsection (lines 237–242) recaps the two regimes with clear inequality statements.
- **Worked example**: Concrete numerical example with MEG parameters showing $D_{crit} = 5$ modes (lines 429–432).

Comment 1.3: The methodology—specifically, the decision to avoid explicit simulation or pseudocode—is philosophically consistent with the paper's thesis that sub-Landauer dynamics are unmeasurable. However, the paper would be strengthened by a clearer justification of how the proposed "feasibility bounds" were derived and validated.

Response: This was Reviewer 2's primary concern as well. We have now added:

- **Four complete Python simulation codes** (Supplementary Material):
  1. `figure1_discrete_vs_continuous.py`: Generates Figure 1 demonstrating discrete failure vs. continuous success in 20D VAS
  2. `vas_collision_comparison.py`: 2D VAS collision vs. collision-free comparison
  3. `vas_scaling_simulation.py`: n-dimensional scaling study ($n \in \{2, 5, 10, 20, 30, 50, 100\}$) showing linear discrete collision count ($\sim 4n$) vs. zero continuous collisions
  4. `code_formation_simulation.py`: Hebbian pathway strengthening demonstrating spontaneous code emergence
- **Quantitative validation**: Table 1 shows discrete collision scaling as $3.95n + 0.6$ ($R^2 = 0.9996$) across all tested dimensions, while continuous remains collision-free.
- **Simulation limitations subsection** (Section 6.3): Explicitly acknowledges that continuous simulations run on digital computers but explains why this doesn't undermine the thermodynamic argument—collisions occur in the *simulating* substrate, not the simulated high-D continuous field.

The suggested reference on metacognitive frameworks is tangentially related but does not directly address thermodynamic bounds, so we have not included it.

Comment 1.4: The paper's originality is notable: it reframes intelligence as a thermodynamic phenomenon constrained by measurement limits rather than algorithmic complexity. The study titled "Artificial Intelligence of Things: A Review" provides a useful contrast, as it situates intelligence within physically instantiated, sensor-rich systems. Drawing this comparison could help articulate how your framework extends beyond existing energy-information paradigms.

Response: We appreciate this recognition. We have strengthened positioning relative to existing frameworks:

- **"Structure and thesis: What's new" section** (lines 151–157): Explicitly lists contributions relative to Landauer's principle, Ashby's requisite variety, VAS theory, free energy principle, neural criticality, morphological computation, and reservoir computing.
- **Connection to morphological/reservoir computing** (line 151): "Our framework provides energetic lower bounds under the morphological/reservoir intuition: we show *why* exploiting substrate dynamics is thermodynamically necessary, not merely efficient."

- **Ashby's law integration** (lines 244–246): Positions our Dimensional Tracking Bound as "quantifying Ashby's variety principle thermodynamically via Landauer-limited power dissipation."

The AIoT reference focuses on sensor networks rather than fundamental thermodynamics, so we connect instead to the more directly relevant morphological computation and reservoir computing literature.

Comment 1.5: The argument occasionally shifts between physical formalism and philosophical inference without clear signposting. Adding brief summary sentences at the end of key subsections (e.g., after equations 1 and 2) would help readers track the conceptual progression from measurement limits to coherence-based computation.

Response: Implemented throughout:

- After Theorem 1 proof: "**Thermodynamic Corollary**" and "**Explicit piecewise form**" subsections summarizing physical implications
- After Theorem 2 (Code Formation): "**Why this matters**" paragraph explaining practical significance
- After VAS section: "**Critical insight**" subsection explicitly connecting timing-inaccessibility to computational tractability
- Section transitions now include 1–2 sentence summaries of "what we just showed" before moving forward

Comment 1.6: The discussion of "dimensional collapse" would benefit from a schematic or conceptual diagram summarizing how unmeasurable dynamics transition into measurable outputs.

Response: We have added:

- **Figure 1**: Four-panel figure showing (A) discrete VAS stuck in local minima, (B) continuous success, (C) dimensional scaling, (D) code formation clustering. Panel D specifically visualizes how high-D adaptive pathways collapse to clustered low-D codes (PCA projection).
- **Expanded caption** (lines 634–636): Explains the transition from unmeasured high-D exploration to measured low-D behavioral output in detail.
- **"Dimensional Expansion and Collapse" section** (Section 4): Now includes explicit energy flow diagram and temporal progression through expansion → evolution → collapse phases.

Comment 1.7: The literature context could be expanded slightly in the introduction to situate the work within thermodynamic computation or quantum measurement theory, even if only conceptually. Doing so would better highlight how the proposed "unmeasurable substrate" differs from prior models of sub-threshold computation.

Response: Enhanced in multiple locations:

- **Introduction references** now cite Landauer (1961), Bennett (1973, 1982), and connect to timing-inaccessibility framework from Todd (2025) BioSystems paper on Maxwell's demon.

- **"Compressible vs. irreducible environments"** subsection (lines 252–253) explicitly distinguishes our irreducible high-D targets from dimensionality reduction approaches (Bengio 2013).
- **Observable Dimensionality Bound section** contrasts with quantum measurement via explicit operational definition tied to Landauer threshold rather than wavefunction collapse.

Comment 1.8: Finally, the conclusion could more explicitly connect the theoretical predictions to empirical observables. Suggesting specific measurable macroscopic signatures (e.g., coherence decay rates or entropy production patterns) would make the paper's testability claim more concrete and compelling.

Response: We have substantially expanded empirical predictions (Section 8):

- **Coherence time scaling**: $\tau_{coherence} \sim D_{eff}^{-1/2}$ testable via MEG/EEG cross-frequency coupling
- **Power scaling**: $P \sim C_{obs}$ (behavioral bandwidth), not D_{eff} (substrate dimensionality)—predicts 20 W for cortex at 100 bits/s output
- **Dimensional collapse signatures**: Abrupt coherence drops at decision points, measurable via phase-locking value (PLV) time series
- **Code reuse efficiency**: Pathway weight concentration in adaptive networks (demonstrated in simulations, Section 6.2)
- **VAS tractability**: Continuous high-D relaxation outperforms discrete enumeration by orders of magnitude (Table 1, Figure 1)

Each prediction includes specific measurement techniques and falsifiable quantitative predictions.

Reviewer 2

General Assessment: This manuscript presents a bold and unconventional theoretical framework proposing that intelligence arises from the maintenance of high-dimensional coherence within sub-Landauer thermodynamic regimes, where computation occurs below measurable thresholds... The work is conceptually rich and elegantly written, but its claims are largely speculative and lack quantitative or empirical substantiation.

Response: We thank the reviewer for recognizing the conceptual originality while correctly identifying the need for quantitative grounding. We have fundamentally strengthened the manuscript by:

1. **Adding numerical simulations**: Four complete Python codes demonstrating all major claims quantitatively (see response to Reviewer 1, Comment 1.3 for details). The VAS scaling study provides particularly strong evidence: across 7 dimensions ($n = 2$ to 100), 20 trials each, discrete collision count scales exactly as predicted ($\sim 4n$, $R^2 = 0.9996$) while continuous remains collision-free.

2. **Empirical parameter estimation**: MEG-based calculation shows human cortex at $D_{eff} \sim 300$ (conservative lower bound from 102 parcels $\times$ 3 bands), yielding $D_{eff}/D_{crit} \sim 60$–100 depending on C_{obs} estimate (Section 3.2, lines 423–432).

3. **Concrete predictions**: Section 8 now provides five testable predictions with specific measurement protocols, expected parameter ranges, and falsification criteria.

4. **Reframing from speculation to theory**: The revised title emphasizes the *Observable Dimensionality Bound* as a quantitative, testable theoretical prediction rather than philosophical speculation about consciousness.

Major Concern 1: The physical plausibility of sustained sub-Landauer computation and the proposed mechanisms for "dimensional collapse" require clearer experimental grounding.

Response: We have clarified the thermodynamic argument:

- **Not claiming sub-Landauer computation violates thermodynamics**: The key insight is *timing-inaccessibility*—when $D_{eff} > D_{crit}$, temporal microstructure cannot be measured at Landauer resolution. Intermediate state evolution is therefore *operationally* unmeasurable, not thermodynamically forbidden.

- **Landauer cost paid at collapse, not during evolution**: Dimensional expansion phase operates collision-free (Section 4); Landauer cost occurs only at behavioral output when high-D state projects to low-D action ($\sim$100 bits/s $\times$ $k_B T \ln 2 \sim 10^{-19}$ W, explaining the 20 W gap).

- **Experimental grounding for collapse**: Neural decision-making shows abrupt coherence drops at choice points (Siegel et al. 2012; Shine et al. 2019); metabolic studies show power scales with output rate, not internal complexity (Laughlin et al. 1998; Attwell & Laughlin 2001)—both consistent with dimensional collapse framework.

- **Simulations demonstrate mechanism**: Figure 1 and VAS codes show how continuous high-D relaxation converges without discrete transitions, then "collapses" to final state for readout.

Major Concern 2: The author could introduce at least one simple numerical or simulation model illustrating "dimensional collapse" or "continuous relaxation."

Response: Fully addressed. We now provide:

- **Continuous relaxation model**: Coupled oscillator VAS solver (Kuramoto-like dynamics) that evolves high-D phase space via overdamped Langevin equation—no discrete transitions, collision-free evolution (code: `vas_collision_comparison.py`, `vas_scaling_simulation.py`).

- **Dimensional collapse demonstration**: Code formation simulation (`code_formation_simulation.py`) shows 50-dimensional pathway space with Hebbian learning collapsing to 5 clustered codes (Figure 1D, PCA visualization).

- **Quantitative comparison**: Table 1 provides scaling data showing discrete requires $\sim$400 collisions at $n = 100$, continuous requires 0. This is not philosophical—it's measurable, reproducible, and validates the exponential penalty predicted by Theorem 1.

All code includes full documentation, reproducible random seeds (42), and parameter specifications.

Major Concern 3: Identify measurable biological proxies (e.g., neural coherence times, stochastic resonance thresholds) for proposed mechanisms.

Response: Section 8 ("Empirical Predictions and Testable Consequences") now provides:

- **Neural coherence times**: Prediction that $\tau_{coherence} \sim 0.1$–$0.3$ s for cortical $D_{eff} \sim 10^3$, measurable via MEG cross-frequency phase coherence (methodology: Palva & Palva 2011).
- **Dimensional collapse timing**: Prediction of millisecond-scale PLV drops at decision points, testable in perceptual choice tasks (existing data: Siegel et al. 2012).
- **Power-bandwidth scaling**: $P/C_{obs} \sim$ constant across species (bacteria: 10^{-12} W / 1 bit/s; humans: 20 W / 100 bits/s), testable via metabolic rate vs. behavioral complexity measurements.
- **Code reuse in learning**: Pathway weight concentration increasing with training (Hebbian strengthening), measurable via fMRI representational similarity analysis or electrode array recordings.
- **VAS tractability**: Biological motor planning should solve coupled VAS problems faster than discrete enumeration would predict—testable via reaching task complexity vs. reaction time scaling.

Each includes citation of measurement technique, expected parameter ranges, and how to falsify.

Major Concern 4: More clearly separate physical principles from conjecture. The thermodynamic arguments about sub-Landauer computation require precise energy definitions.

Response: We have rigorously separated claims:

- **Physical principles** (Theorem 1 & 2, Observable Dimensionality Bound): Derived from Landauer's principle + information theory + collision counting. No speculation.
- **Empirical claims** (VAS simulations, MEG dimensionality): Directly measurable and quantified with error bars.
- **Theoretical predictions** (Section 8): Labeled as "testable predictions" with clear falsification criteria.
- **Speculation clearly marked**: Consciousness discussion (Section 7.3) now opens with "This remains speculative but suggests..." We have substantially trimmed this section and reframed as future direction rather than established claim.

Energy definitions clarified throughout:

- h_ε^{prod}: Entropy production (bits/mode)—zero during reversible expansion
- h_ε^{track}: Information required to track constraint geometry—nonzero even for reversible processes
- Landauer cost: $k_B T \ln 2$ per *committed* bit (irreversible state resolution)

- Dimensional collapse: Transition from unmeasured high-D ($h_\varepsilon^{prod} = 0$) to committed low-D output (pays $C_{obs} \times k_B T \ln 2$ per second)

Structural Suggestions: Clearer structure, integration of contemporary computational neuroscience literature, and a more operational conclusion emphasizing testable predictions.

Response: Implemented:

- **Structure**: Added explicit "Structure and thesis: What's new" section (Introduction), "Simulation Limitations and Assumptions" (Section 6.3), and reorganized VAS section for clarity.
- **Contemporary neuroscience literature**: Now cites Miller et al. (2024) on cognition as emergent property, Pinotsis et al. (2023) on ephaptic coupling and cytoelectric memory, Shine et al. (2019) on neuromodulatory integration, Recanatesi et al. (2024) on metastable attractors, Bergoin et al. (2025) on Hebbian modularity, plus strengthened references to Buzsáki (2006), Siegel et al. (2012), and Levin et al. (2024) on collective intelligence.
- **Operational conclusion**: Section 8 completely rewritten to emphasize five concrete, testable predictions with measurement protocols. Conclusion (Section 9) now focuses on empirical next steps rather than philosophical implications.

Summary of Major Revisions

1. **Quantitative demonstration**: Four simulation codes, scaling table, Figure 1—transforms theory from speculation to validated framework
2. **Empirical grounding**: MEG dimensionality calculation, five testable predictions with measurement protocols, biological parameter ranges
3. **Accessibility**: Restructured abstract/introduction, explicit assumption statements, summary sentences, clearer terminology
4. **Methodological rigor**: Simulation limitations acknowledged, physical principles separated from conjecture, energy definitions clarified
5. **Literature integration**: Connected to Ashby, morphological computation, reservoir computing, contemporary neuroscience
6. **Structural improvements**: "What's new" section, piecewise bound clarification, operational predictions in conclusion

The revised manuscript preserves its theoretical ambition while meeting standards for scientific testability. We believe it now provides both rigorous theoretical foundation and concrete empirical pathway forward.

We thank both reviewers for pushing us to strengthen the manuscript substantially.

Intelligence as High-Dimensional Coherence: The Observable Dimensionality Bound and Computational Tractability

Ian Todd

Sydney Medical School

University of Sydney

Sydney, NSW, Australia

itod2305@uni.sydney.edu.au

November 16, 2025

Abstract

Intelligence must be high-dimensional. We show that, under minimal thermodynamic assumptions, any system capable of tracking and controlling high-dimensional targets faces exponential thermodynamic cost unless its substrate dimensionality matches or exceeds the target dimensionality. This applies to *all living systems*: bacteria tracking chemical gradients ($D_{\text{target}} \sim 10^2$–$10^3$, $\sim 10^{-12}$ W) and human brains tracking social/ecological complexity ($D_{\text{target}} \sim 10^3$–$10^6$, ∼20 W) both require $D_{\text{obs}} \gg D_{\text{target}}$, scaled to their respective behavioral bandwidths. Intelligence—the capacity to track and respond to environmental complexity—is a fundamental thermodynamic property of life, not unique to humans.

We establish three results: (1) High-dimensional systems with irreducible complexity exist (e.g., vector addition systems with Ackermann-complete reachability; cou-

pled oscillator networks with $D_{\text{eff}} \gg 1$). (2) *Theorem (Dimensional Tracking Bound):* Tracking a system with dimensionality D_{target} using algorithmic inference from substrate dimensionality $D_{\text{obs}} < D_{\text{target}}$ requires collision rate scaling as $k^{D_{\text{target}} - D_{\text{obs}}}$ (exponential in dimensional mismatch), yielding thermodynamic power penalty up to $10^6\times$ for realistic mismatches. (3) Biological intelligence escapes this bound by operating as a high-dimensional substrate ($D_{\text{obs}} \gg D_{\text{target}}$) with collision-free constraint geometry, paying Landauer cost only at sparse behavioral outputs ($\sim 10^2$ bits/s), not during internal computation.

We derive an Observable Dimensionality Bound $D_{\text{crit}} = C_{\text{obs}}\tau_e/(\alpha h_\varepsilon^{\text{track}})$ showing when $D_{\text{eff}} > D_{\text{crit}}$, temporal microstructure becomes physically unmeasurable—systems compute faster than observation can track. Human cortex operates at $D_{\text{eff}}^{\text{MEG}}/D_{\text{crit}} \sim 10^2$ (substrate plausibly 10^3–10^4), explaining both power efficiency (collisions only at behavioral bottleneck) and computational tractability (VAS problems Ackermann-complete for discrete enumeration become tractable via continuous high-dimensional relaxation). This framework unifies biological efficiency with the thermodynamic impossibility of low-dimensional algorithmic intelligence.

Keywords: Intelligence, Observable Dimensionality Bound, Curse of Dimensionality, Landauer Limit, Phase Space, Thermodynamics, Vector Addition Systems, Constraint Geometry, Measurement Limits, Power Scaling

Highlights:

- **Theorem:** Tracking D_{target}-dimensional systems from $D_{\text{obs}} < D_{\text{target}}$ substrate requires exponential collision cost $k^{D_{\text{target}} - D_{\text{obs}}}$; intelligence must be high-dimensional
- High-dimensional systems with irreducible complexity exist (VAS reachability: Ackermann-complete; no algorithmic shortcut)
- Living systems (bacteria to humans) track complex environments via high-dimensional substrates ($D_{\text{eff}} \gg D_{\text{crit}}$); thermodynamically impossible for low-dimensional algorithmic systems at equivalent power

- Observable Dimensionality Bound: $D_{\text{crit}} = C_{\text{obs}}\tau_e/(\alpha h_\varepsilon^{\text{track}})$; when $D_{\text{eff}} > D_{\text{crit}}$ temporal microstructure becomes unmeasurable
- Human cortex: $D_{\text{eff}}^{\text{MEG}}/D_{\text{crit}} \sim 10^2$; power scales with behavioral output ($\sim 10^2$ bits/s), not internal dimensionality

1 Introduction: Intelligence Must Be High-Dimensional

1.1 The Core Argument

What is intelligence? Operationally, intelligence is the capacity to *track and control high-dimensional systems*—predicting the behavior of other agents, navigating complex environments, regulating internal physiology, planning actions under multiple constraints. An intelligent system must maintain internal representations sufficient to predict how high-dimensional targets will evolve and determine which interventions will achieve desired outcomes.

This paper establishes a thermodynamic proof: **any system capable of tracking high-dimensional targets must itself be high-dimensional**. Specifically:

1. **High-dimensional systems with irreducible complexity exist.** Some systems cannot be faithfully compressed to low-dimensional representations without exponential information loss. Examples: coupled oscillator networks with $D_{\text{eff}} \sim 10^3$ active modes; vector addition systems (VAS) with Ackermann-complete reachability (no polynomial-time algorithm exists).

2. **Tracking requires dimensional matching.** We prove (Theorem 1, §1.7) that any algorithmic observer with substrate dimensionality D_{obs} attempting to track a target with dimensionality $D_{\text{target}} > D_{\text{obs}}$ faces collision cost scaling as $k^{D_{\text{target}}-D_{\text{obs}}}$ (exponential in the dimensional mismatch). This yields thermodynamic power penalties reaching $10^6\times$ for realistic mismatches.

3. **Biological intelligence operates on high-D targets at low power.** The human brain tracks other agents ($D_{\text{eff}} \sim 10^3$–10^4), ecosystems ($D_{\text{eff}} \sim 10^4$–10^5), and internal physiology ($D_{\text{eff}} \sim 10^6$) while dissipating ∼20 W. This is thermodynamically impossible for low-dimensional algorithmic systems, which would require megawatts to gigawatts for equivalent tracking (see §1.7).

4. **Therefore: intelligence must be implemented as high-dimensional substrate.** The only way to avoid exponential thermodynamic cost is $D_{\text{obs}} \gg D_{\text{target}}$. Biological systems achieve this through collision-free constraint geometry in high-dimensional phase space, paying Landauer cost only at sparse behavioral outputs ($\sim 10^2$ bits/s), not during internal computation.

This is not a design optimization but a *thermodynamic necessity*: low-dimensional algorithmic intelligence cannot scale to track high-dimensional reality without exponentially increasing power dissipation.

1.2 Implications and Existing Frameworks

Training GPT-scale models ($\sim 10^{11}$ parameters) requires $\sim 10^{24}$ collision events [20], consuming megawatts. The human brain achieves comparable complexity at ∼20 W—six orders of magnitude less. This gap reflects fundamental thermodynamic cost of dimensional mismatch.

Existing frameworks: Landauer's principle [4], reversible computing [26, 27], free energy principle [21], self-organization theory [42, 43], neural criticality [45], and recent work on sub-Landauer regimes [1, 2] address aspects but don't unite power efficiency with computational tractability through collision dynamics in high-dimensional constraint geometry.

Our proposal: Biological intelligence operates through collision-free exploration in high-dimensional spaces, enabling power efficiency (sparse collisions) and computational tractability (escaping local minima). When $D_{\text{eff}} > D_{\text{crit}}$, temporal microstructure becomes physically

unmeasurable—timing inaccessibility as thermodynamic consequence.

Irreducible complexity is irreducible. A common objection invokes sparse coding or dimensionality reduction [29]: if data lie on low-dimensional manifolds, clever algorithms could avoid the curse. This applies when targets have *reducible* complexity (e.g., image classification on natural manifolds). But systems with irreducible high-D—ecosystems, vector addition systems with Ackermann-complete reachability, multi-agent dynamics—cannot be faithfully compressed without information loss. Biology doesn't compress everything because tracking physiology ($D_{\text{eff}} \sim 10^6$) or social interactions ($D_{\text{eff}} \sim 10^3$–10^4) requires real-time coherence in the full constraint geometry. Modern AI (e.g., GPT-scale models) succeeds not via low-dimensional cleverness but by brute-force approximation of high-dimensional coherence—$\sim 10^{11}$ parameters effectively simulate high-dimensionality, still dissipating megawatts versus the brain's 20 W. Compression helps *within* high-dimensional substrates (e.g., sparse activations reduce active pathways), but doesn't escape dimensional matching for irreducible targets.

1.3 The Observable Dimensionality Bound

We establish a fundamental relationship between dimensionality, measurement capacity, and temporal resolution. Consider a system with effective dimensionality D_{eff} (number of active degrees of freedom), evolving over timescale τ_e with compressibility $\alpha \in (0, 1]$. Let ε denote the coarse-graining resolution (phase/temporal precision) used to define information rates; unless stated otherwise, all rates are computed at the same ε.

To track the constraint geometry requires minimal measurement resolution $h_\varepsilon^{\text{track}}$ (bits per mode per τ_e). This differs from the entropy production $h_\varepsilon^{\text{prod}}$ during evolution: reversible constraint geometry has $h_\varepsilon^{\text{prod}} = 0$ (no bit creation), but tracking its structure still requires $h_\varepsilon^{\text{track}} > 0$ per mode.

There exists a critical dimension:

$$D_{\text{crit}} = \frac{C_{\text{obs}} \tau_e}{\alpha h_\varepsilon^{\text{track}}} \tag{1}$$

where C_{obs} is the observation channel capacity (bits/s). Dimensional check: $[C_{\text{obs}}] =$ bits/s, $[\tau_e] =$ s, $[h_\varepsilon^{\text{track}}] =$ bits/mode, giving $[D_{\text{crit}}] =$ modes as required. Operationally, $C_{\text{obs}} \leq B \log_2(1 + \text{SNR})$, where B is the effective behavioral/motor bandwidth estimated from response variance and control degrees of freedom; we intentionally use *external* observation capacity, not internal sensing bandwidth. When $D_{\text{eff}} > D_{\text{crit}}$, the system's constraint geometry reconfigures faster than measurement can track—it becomes **timing-inaccessible**.

Key insight: This establishes a physical boundary separating observable computation (where temporal structure is measurable) from timing-inaccessible computation (where only integrated observables are accessible).

Operational equivalence: Define the intrinsic dynamics rate $H_{\text{dyn}} = \alpha h_\varepsilon^{\text{track}} D_{\text{eff}} / \tau_e$ (information generated per second by constraint reconfiguration). Then by direct substitution:

$$D_{\text{eff}} > D_{\text{crit}} \quad \Longleftrightarrow \quad H_{\text{dyn}} > C_{\text{obs}} \tag{2}$$

The system's constraint geometry evolves faster than any observer can track it.

Consequence: Biological intelligence operates precisely by maintaining $H_{\text{dyn}} \gg C_{\text{obs}}$, keeping internal dynamics inaccessible while colliding only at sparse behavioral outputs.

Why dimensionality creates timing-inaccessibility. This is not about "fast dynamics"—a 2D oscillator at GHz rates remains timing-accessible if you have GHz measurement bandwidth. It is about **measurement bandwidth per degree of freedom**. Tracking N coupled oscillators requires resolving individual phase trajectories $\phi_i(t)$ for $i = 1, \ldots, N$, costing $N \cdot h_\varepsilon^{\text{track}}$ bits per τ_e. When $N \gg C_{\text{obs}} \tau_e / h_\varepsilon^{\text{track}}$, you cannot afford to track individual modes—only integrated observables (bulk coherence, magnetization) remain measurable. The *sequence* of microstates becomes inaccessible; only macrostate evolution persists. High-dimensional phase space contains exponentially many trajectories between macrostates, and

finite observation bandwidth cannot resolve which path was taken. Concretely: cortical dynamics at $D_{\text{eff}} \sim 10^4$ modes and $\tau_e \sim 100$ ms with behavioral output $C_{\text{obs}} \sim 10^2$ bits/s means you can measure *that* constraint satisfaction occurred, but not *when* individual phase adjustments happened—the temporal microstructure is fundamentally unrecoverable (see Vector Addition Systems, §3).

1.3.1 Precise Definitions

To avoid ambiguity, we define key terms operationally:

Timing-inaccessible. A regime in which the coarse-grained dynamical information rate $H_{\text{dyn}} = \alpha\, h_\varepsilon^{\text{track}} D_{\text{eff}}/\tau_e$ (bits/s) exceeds the available observation channel capacity C_{obs} at resolution ε; temporal order of intermediate states is unrecoverable without disrupting the dynamics.

Collision (commit). A logically irreversible registration of mutually exclusive macrostates that writes information to a stable record and thus dissipates $\geq k_B T \ln 2$ per bit.

$h_\varepsilon^{\textbf{track}}$ **vs** $h_\varepsilon^{\textbf{prod}}$**.** $h_\varepsilon^{\text{track}}$ is a measurement-side requirement (bits per mode per τ_e to *track* geometry). $h_\varepsilon^{\text{prod}}$ is substrate-side net information production: ≈ 0 during reversible expansion; > 0 at collapse.

1.4 Power Scaling and Computational Mode

The power cost of computation scales with the **enforced collision rate** H_{reg} (registrations per second), not with dimensionality or internal dynamics:

$$P_{\min} = \frac{k_B T \ln 2}{\eta} H_{\text{reg}} \tag{3}$$

where $\eta \in (0, 1]$ is recording efficiency [4].[1] This has profound implications:

[1] Landauer's bound applies to logically irreversible operations (erasure/overwrite). Here "collisions" are precisely those registrations. Purely reversible readout cannot, at scale, extract temporal decompositions when $H_{\text{dyn}} > C_{\text{obs}}$ without enforcing registrations and dissipation.

- **Digital clocked systems:** Enforce $H_{\text{reg}} \approx f_c \times$ (words per cycle) where $f_c \sim$ GHz. Every mode undergoes collision every cycle. Power dissipation is continuous and high.

- **Biological systems:** Collide sparsely at behavioral output only. Internal constraint geometry evolves in high-dimensional phase space ($h_\varepsilon^{\text{prod}} = 0$, no entropy production), but behavioral collisions occur at $H_{\text{reg}} \sim 10^2$ bits/s. Power dissipation is minimal.

Measured power gap: Human cortex (20 W) vs. 8-GPU compute node (8–12 kW, e.g., 8× H100 DGX system 10 kW) [37] operating at comparable effective dynamical richness shows **two to three orders of magnitude** difference. (Note: data center total power includes facility overhead—Power Usage Effectiveness (PUE) typically 1.5–2× for cooling, power distribution; chip-level comparison remains $\sim 10^2$–$10^3\times$ gap.) The irreducible Landauer collision floor at $H_{\text{reg}} \sim 10^2$ bits/s is $P_{\text{Landauer}} = k_B T \ln 2 \cdot H_{\text{reg}} \approx 3 \times 10^{-19}$ W at $T = 310$ K—negligible. Observed 20 W is overwhelmingly *metabolic maintenance* (ion pumps, vesicle cycling, synaptic homeostasis), not information collisions [24]. The efficiency advantage arises from sparse collisions ($H_{\text{reg}} \ll H_{\text{dyn}}$) versus enforced clocked registration ($H_{\text{reg}} \approx f_c$).

Summary: Power efficiency follows from operating in timing-inaccessible regimes where computation occurs in unmeasured constraint geometry, paying Landauer cost only at sparse measurement events.

1.5 Structure and Thesis

This paper demonstrates that collision-free computation is not merely a power-saving mechanism but a **computational mechanism**. Hard problems become tractable precisely because the system operates through:

1. High-dimensional exploration where incompatible states coexist without forced resolution

2. Continuous constraint satisfaction without discrete collision events ($h_\varepsilon^{\text{prod}} = 0$)

3. Sparse collisions only at behavioral output, paying Landauer cost for bits actually written

The curse of dimensionality becomes a blessing. Discrete algorithms suffer exponential scaling with dimension (the "curse of dimensionality")—search spaces explode, collision rates grow superlinearly, computational cost becomes prohibitive. For collision-free computation, this reverses: higher dimensionality provides more orthogonal subspaces for parallel exploration without forced resolution. The collision-free advantage *increases* with dimension.

We establish this through vector addition systems (VAS)—problems proven to be Ackermann-complete via discrete enumeration yet efficiently solvable by biological systems. The resolution: collision-free high-dimensional exploration that escapes local minima trapping discrete approaches.

Structure and thesis: What's new. Existing work provides essential foundations: Landauer's principle establishes bit-level thermodynamic cost [4]; Ashby's requisite variety gives a qualitative constraint on control [3]; vector addition system (VAS) theory shows some high-D problems are Ackermann-complete [50]; free energy principle, neural criticality, morphological computation, and reservoir computing address aspects of biological efficiency [21, 45]. Our framework provides energetic lower bounds under the morphological/reservoir intuition: we show *why* exploiting substrate dynamics is thermodynamically necessary, not merely efficient. Our contributions:

1. **Observable Dimensionality Bound**: A quantitative threshold $D_{\text{crit}} = C_{\text{obs}}\tau_e/(\alpha h_\varepsilon^{\text{track}})$ tying measurement capacity, timescale, and per-mode tracking cost into a single criterion for timing-inaccessibility.

2. **Collision Rate Lower Bound (Theorem 1)**: Proof that tracking irreducible targets with dimensional mismatch $D_{\text{target}} - D_{\text{obs}}$ requires collision cost scaling as $k^{D_{\text{target}}-D_{\text{obs}}}$, quantifying Ashby's variety principle thermodynamically.

3. **Energy gap explanation**: Formalization of why brains operate at ∼20 W while equivalent low-D bitwise trackers would require megawatts, backed by explicit VAS numerics and power calculations.

4. **Code formation theorem (Theorem 2)**: Stable symbolic codes emerge as thermodynamic necessity when high-D systems communicate through low-D channels, explaining language, concepts, and social conventions.

1.6 What Does "High-Dimensional" Mean?

Two meanings: (i) **Substrate dimensionality**—physical nonlocal coupling via ephaptic fields/gap junctions creating constraint geometry evolving reversibly ($h_\varepsilon^{\text{prod}} = 0$); (ii) **Algorithm dimensionality**—logical coupling via sequential bit operations requiring simulation with continuous collisions ($h_\varepsilon^{\text{prod}} > 0$).

Key distinction: Low-D logic (bits) requires collision resolution—mutually exclusive states force discrete choices. High-D geometry (phases) enables collision-free dynamics—incompatible states coexist in orthogonal subspaces. Neural spikes are sparse outputs ($\sim 10^2$ bits/s) while field dynamics maintain $D_{\text{eff}} \sim 10^3$–10^4 collision-free [12, 13].

1.7 Bayesian Inference and the Algorithmic Collision Requirement

Why discrete algorithms must collide. Algorithmic intelligence—machine learning, probabilistic inference, classical AI—operates through Bayesian reasoning over discrete hypothesis spaces. This is not a design choice but a structural necessity: when computation is implemented via bits, inference requires collision resolution at every update.

Consider Bayesian posterior updates:

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{\sum_j P(D|H_j)P(H_j)} \tag{4}$$

Each hypothesis H_i is encoded as a discrete state (bit pattern). Computing the posterior requires:

1. Evaluating likelihoods $P(D|H_i)$ for each hypothesis (collision: read data, compute likelihood, write result)

2. Normalizing over all hypotheses (collision: accumulate denominator, divide, store updated belief)

3. Selecting maximum *a posteriori* estimate or sampling from posterior (collision: compare beliefs, resolve to single output)

Each computational step is a collision event—mutually exclusive states (H_i vs. H_j) compete for belief allocation, requiring discrete resolution. The number of collisions scales with hypothesis space size and data dimensionality.

The curse of dimensionality in Bayesian computation. For a system with n dimensions, each with k discrete values, the hypothesis space has $|H| = k^n$ states. Computing the full posterior requires:

- $O(k^n)$ likelihood evaluations (each a collision)

- $O(k^n)$ normalization operations (collisions)

- Total collision count: $O(k^n)$ per inference cycle

This is Bellman's *curse of dimensionality* [28] for algorithmic intelligence: as n grows, discrete Bayesian inference becomes computationally intractable—collision count explodes exponentially. The statistical literature extensively documents these fundamental constraints: posterior consistency fails without strong structural assumptions in high-dimensional Bayesian regression [31], model selection becomes inconsistent with standard priors [33], and high-dimensional Bayesian optimization is described as facing an "Achilles' heel" from dimensional scaling [32]. Digital medicine applications of machine learning confront this same barrier:

modern digital health data (imaging, speech, wearables, genomics) produce "dataset blind spots" where models appear validated but fail catastrophically in deployment when feature spaces vastly exceed available training data [30].

Standard approximations (MCMC, variational Bayes, particle filters) reduce collision count by exploiting structure but still scale superlinearly with effective dimension. Modern deep learning architectures defer the exponential barrier through manifold learning and dimensionality reduction, but once D_{eff} substantially exceeds D_{tract} even after exploiting all available structure, exponential scaling returns [29].

1.7.1 Formal Statement: The Dimensional Tracking Impossibility Theorem

We now formalize the core argument linking observational dimensionality to thermodynamic cost.

Theorem 1 (Dimensional Tracking Bound). *Let an observer with substrate dimensionality D_{obs} (number of independently controllable degrees of freedom) attempt to track an irreducible target system with effective dimensionality D_{target} (for which no sufficient statistic of dimension $< D_{target}$ exists for the tracking task) evolving at rate H_{dyn} (bits/s). If the observer uses discrete algorithmic inference with hypothesis resolution k states per dimension, then the minimum collision rate required for stable tracking satisfies:*

$$C_{\min} \geq \frac{H_{\text{dyn}}}{h_\varepsilon} \cdot k^{\max(0, D_{\text{target}} - D_{\text{obs}})} \tag{5}$$

where $h_\varepsilon = h_\varepsilon^{track}$ is the bits per mode per τ_e required to track constraint geometry (not entropy production h_ε^{prod}), and C_{min} is measured in collisions per second.

Proof. The target system generates information at rate $H_{\text{dyn}} = \alpha h_\varepsilon D_{\text{target}}/\tau_e$ through evolving constraint geometry. To track this system, the observer must update internal state representations at least as fast as the target changes.

Case 1: $D_{obs} \geq D_{target}$ (dimensional matching). The observer has sufficient degrees of freedom to represent the target state directly. Each dimension can be tracked independently with $O(k)$ hypothesis evaluations per update. Total collision count per update cycle:

$$C_{\text{match}} = O(k \cdot D_{\text{target}}) \quad \text{(polynomial scaling)} \tag{6}$$

Case 2: $D_{obs} < D_{target}$ (dimensional mismatch). The observer must compress D_{target} dimensions into D_{obs} substrate dimensions. This projection loses information:

$$\Delta I = (D_{\text{target}} - D_{\text{obs}}) \log_2 k \quad \text{bits per snapshot} \tag{7}$$

To recover this lost information and maintain tracking fidelity, the observer must enumerate hypotheses over the unmeasured dimensions. For irreducible high-dimensional targets, the missing coordinates are algorithmically independent—no lower-dimensional sufficient statistic captures their joint behavior. The hypothesis space for the unobserved $(D_{\text{target}} - D_{\text{obs}})$ dimensions contains $k^{D_{\text{target}} - D_{\text{obs}}}$ states, all of which must be considered. Each must be evaluated via Bayesian updating (Eq. 1). Each evaluation is a collision event (state comparison, likelihood computation, posterior update). Thus:

$$C_{\text{mismatch}} = \Omega(k^{D_{\text{target}} - D_{\text{obs}}}) \quad \text{per update cycle} \tag{8}$$

At update rate $f = H_{\text{dyn}}/h_\varepsilon$, the total collision rate is:

$$C_{\text{min}} = f \cdot k^{D_{\text{target}} - D_{\text{obs}}} = \frac{H_{\text{dyn}}}{h_\varepsilon} \cdot k^{D_{\text{target}} - D_{\text{obs}}} \tag{9}$$

This completes the proof. □

Thermodynamic Corollary. Each collision event dissipates at minimum the Landauer cost $k_B T \ln 2$ (assuming binary resolution; factor $\ln k$ for k-ary). The minimum power

dissipation for algorithmic tracking with dimensional mismatch is:

$$P_{\min} = k_B T \ln 2 \cdot C_{\min} = k_B T \ln 2 \cdot \frac{H_{\text{dyn}}}{h_\varepsilon} \cdot k^{D_{\text{target}} - D_{\text{obs}}} \tag{10}$$

For $D_{\text{target}} - D_{\text{obs}} = 10$ and $k = 2$, this yields $2^{10} = 1024\times$ power penalty. For $D_{\text{target}} - D_{\text{obs}} = 20$, the penalty is $2^{20} \approx 10^6\times$. *Algorithmic intelligence attempting to track high-dimensional reality from a low-dimensional substrate faces exponential thermodynamic cost.*

Explicit piecewise form. The tracking bound separates into two regimes:

- **Dimensional matching** ($D_{\text{obs}} \geq D_{\text{target}}$): $C_{\min} \gtrsim H_{\text{dyn}}/h_\varepsilon$. Every bit of target dynamics must be processed at least once; collision rate scales polynomially in D_{target}.
- **Dimensional mismatch** ($D_{\text{obs}} < D_{\text{target}}$): $C_{\min} \gtrsim (H_{\text{dyn}}/h_\varepsilon) \cdot k^{D_{\text{target}} - D_{\text{obs}}}$. Exponential penalty from enumerating hypotheses over unobserved dimensions.

The compact form $C_{\min} \geq (H_{\text{dyn}}/h_\varepsilon) \cdot k^{\max(0, D_{\text{target}} - D_{\text{obs}})}$ captures both cases.

Connection to Ashby's Law of Requisite Variety. This result can be understood as a thermodynamic refinement of Ashby's law of requisite variety [3]: a controller must have at least as much variety as the disturbances it compensates. Here, "variety" is instantiated as effective dimensionality D_{eff}, and we show that when $D_{\text{obs}} < D_{\text{target}}$, the shortfall manifests as exponential collision cost ($k^{D_{\text{target}} - D_{\text{obs}}}$) rather than simply a loss in control quality. Ashby's principle provides the qualitative constraint; Theorem 1 quantifies its thermodynamic penalty via Landauer-limited power dissipation.

Implication: Intelligence Must Be High-Dimensional. The thermodynamic argument applies to *all living systems*, not just humans. Intelligence—the capacity to track and respond to environmental complexity—is a fundamental property of life. A bacterium tracking chemical gradients [44] navigates $D_{\text{target}} \sim 10^2$–$10^3$ (multiple ligand concentrations, spatial gradients, temporal dynamics) while dissipating $\sim 10^{-12}$ W. The human brain tracks other agents, ecosystems, and internal physiology ($D_{\text{target}} \sim 10^3$–$10^6$) at ∼20 W. In both cases, $D_{\text{obs}} \gg D_{\text{target}}$ is required: dimensional matching scales with the organism's

behavioral bandwidth C_{obs}, not absolute power. A bacterium with $C_{\text{obs}} \sim 1$ bit/s needs $D_{\text{crit}} \sim 0.01$–$0.1$; cortex with $C_{\text{obs}} \sim 10^2$ bits/s needs $D_{\text{crit}} \sim 5$. Both operate well above their respective thresholds via high-dimensional substrates (bacterial chemotaxis networks, neural oscillations). The only thermodynamically viable solution is $D_{\text{obs}} \gg D_{\text{target}}$: intelligence must be implemented in high-dimensional substrates operating through collision-free constraint geometry. This is not a design choice but a thermodynamic necessity *for all living systems*.

Corollary (Simulation Impossibility). For any system with $D_{\text{target}} \gg 1$, there exists no bitwise simulator with $D_{\text{obs}} < D_{\text{target}}$ that can faithfully track its dynamics at comparable power dissipation. Either fidelity or efficiency must be sacrificed.

Why "just add more bits" doesn't work. Any finite physical system is, in a mathematical sense, encodable as a finite binary string. What this truism hides is the only part that matters for physics: *how many bits, updated how often, at what energy cost*. High-dimensional biological systems such as brains generate new dynamical structure at a rate H_{dyn} that rivals or exceeds the Landauer-limited commit capacity C_{obs} of any realistic digital substrate at comparable power. In this **timing-inaccessible** regime there is no bitwise simulation that is at once faithful, real-time, and thermodynamically efficient. A digital model must either coarse-grain away much of the underlying dynamics or pay a strictly higher energetic cost than the system it attempts to track. The question is not whether you *can* represent it in bits, but whether you can *update* those bits fast enough, at low enough power, to track a high-dimensional system in real-time. The Dimensional Tracking Bound proves the answer is no for $D_{\text{obs}} < D_{\text{target}}$.

Compressible vs. irreducible environments. In environments with *reducible* complexity—where sufficient statistics of dimension $< D_{\text{target}}$ exist (e.g., natural images lying on low-dimensional manifolds)—the bound applies to the **effective target dimensionality** of those statistics, not the full microstate dimensionality. Biological systems exploit such structure when it exists. However, many tracking tasks involve irreducible complexity: ecosystems

with coupled nonlinear dynamics, multi-agent coordination, internal physiological regulation with $\sim 10^6$ molecular species, turbulent fluid flows (where prediction requires measuring essentially all degrees of freedom). For these, no faithful low-dimensional compression exists, and Theorem 1's exponential penalty applies in full.

1.8 Code Formation from Dimensional Mismatch

Theorem 1 immediately implies a second result: when two high-dimensional systems interact through a low-dimensional communication channel, *stable codes must form at the boundary* as a thermodynamic consequence of the exponential collision cost of dimensional mismatch.

Theorem 2 (Code Formation from Dimensional Mismatch). Consider two systems A and B with internal dimensionalities $D_A, D_B \gg 1$ that interact through a communication channel of dimensionality $D_{\text{link}} < \min(D_A, D_B)$. If the systems successfully maintain mutual coordination (each tracks relevant aspects of the other's state) at power dissipation P_{obs}, then stable reusable codes must form at the interface.

Proof. System A must track aspects of system B's state through the low-dimensional link. By Theorem 1, if A attempts to track the full D_B-dimensional state through a D_{link}-dimensional observation channel, the collision cost is:

$$C_{\text{raw}} \geq \frac{H^B_{\text{dyn}}}{h_\varepsilon} \cdot k^{D_B - D_{\text{link}}} \tag{11}$$

This exponential cost is prohibitive for $D_B - D_{\text{link}} \gg 1$. However, suppose B's state space contains *structured regions*—subspaces $\mathcal{S}_i$ that recur frequently and are behaviorally relevant to A. If A can represent these regions as discrete *codes* c_i (low-dimensional symbols), the effective dimensionality of what must be tracked collapses from D_B to $D_{\text{code}} \sim \log_k(N_{\text{codes}})$, where N_{codes} is the size of the codebook.

With codes, the collision cost becomes:

$$C_{\text{coded}} \geq \frac{H^B_{\text{dyn}}}{h_\varepsilon} \cdot k^{D_{\text{code}} - D_{\text{link}}} + C_{\text{encode}} \tag{12}$$

where C_{encode} is the cost of mapping B's high-dimensional state onto the discrete code c_i (a one-time compression cost per code).

Thermodynamic favorability. For sufficiently large dimensional mismatch, we have:

$$k^{D_{\text{code}} - D_{\text{link}}} \ll k^{D_B - D_{\text{link}}} \tag{13}$$

Thus $C_{\text{coded}} \ll C_{\text{raw}}$. Code formation is thermodynamically favorable: systems that *fail* to form stable codes pay exponentially higher collision costs and are outcompeted or fail to maintain coordination.

Code stability emerges from reuse. A code c_i becomes stable when:

1. **Recurrence:** The subspace $\mathcal{S}_i$ occurs frequently in B's dynamics
2. **Compressibility:** $\mathcal{S}_i$ can be reliably identified from low-dimensional projections (otherwise encoding cost C_{encode} dominates)
3. **Behavioral relevance:** Tracking $\mathcal{S}_i$ enables A to coordinate effectively with B (otherwise the code provides no fitness benefit)

When these conditions hold, the code c_i is reinforced through repeated use, as each successful encoding amortizes C_{encode} over many tracking events. Unstable or rarely-used codes are pruned because their encoding cost exceeds the collision savings.

Implication: The origin of symbolic communication. This theorem provides a thermodynamic explanation for the emergence of language, gesture, and other symbolic codes in biological and social systems. When two agents with high internal dimensionality ($D_{\text{brain}} \sim 10^3$–$10^6$) coordinate through low-bandwidth channels (speech ~ 50 phonemes/s, gesture ~ 10 distinct poses), stable symbolic codes are not merely *useful*—they are thermodynamically

necessary. Systems that fail to compress high-dimensional intentions into low-dimensional symbols cannot maintain coordination without prohibitive collision costs.

Similarly, the formation of concepts, categories, and mental models can be understood as internal code formation: the brain's low-dimensional output systems (motor cortex $\sim 10^2$ active neurons) must coordinate with high-dimensional sensory and cognitive dynamics ($\sim 10^6$ cortical neurons), forcing the emergence of reusable symbolic representations that compress high-dimensional percepts into manipulable codes.

Connection to Bayesian compression. The formation of codes is equivalent to constructing a *structured prior* over B's state space. Instead of treating all k^{D_B} hypotheses as equally likely (maximum entropy prior), the codebook imposes structure: "system B is likely in one of N_{codes} recurring configurations $\mathcal{S}_i$." This structured prior reduces the effective hypothesis space from k^{D_B} to $N_{\text{codes}} \cdot k^{D_{\text{residual}}}$, where $D_{\text{residual}} \ll D_B$ captures within-code variation. The exponential savings arise from exploiting statistical regularity in B's dynamics—precisely the mechanism by which Bayesian compression escapes the curse of dimensionality in practice [29].

Numerical illustration. We demonstrate this spontaneous code formation numerically in Section 3 (Figure 1, Panel D). A 50-pathway adaptive network exposed to clustered task structure spontaneously concentrates weight onto $\sim$10% of pathways, with clear code clustering emerging from Hebbian-like dynamics. This validates Theorem 2's prediction that dimensional mismatch drives code formation (see `code_formation_simulation.py` in supplementary materials).

Code drift and co-evolution. Because both A and B face symmetric dimensional mismatch (each tracks the other through D_{link}), codes must be *mutually stabilized*. If A uses code c_i to represent a subspace of B, then B must use a corresponding code c_i' to represent the subspace of A that triggers c_i. This mutual stabilization creates **co-evolutionary pressure**: codes that are not mutually interpretable fail to reduce collision costs and are pruned. Over time, this drives convergence toward shared symbolic representations—the thermodynamic

origin of communication protocols, social conventions, and cultural norms.

Why biological intelligence escapes the Bayesian explosion. Biological systems do not enumerate hypotheses over discrete state spaces. Instead:

- **Continuous constraint geometry:** "Hypotheses" exist as attractors in high-dimensional phase space, not as discrete states requiring resolution
- **Parallel exploration:** Incompatible interpretations coexist via superposition in orthogonal subspaces—multiple candidate solutions evolve simultaneously without forced collision
- **Gradient-based convergence:** Constraint satisfaction through energy minimization (relaxation to attractor) rather than explicit hypothesis enumeration
- **Sparse collapse:** Behavioral output collapses wavefunction to discrete action, but internal deliberation remains collision-free

The Bayesian-geometric duality. Discrete Bayesian inference (k^n hypotheses) and continuous high-dimensional exploration (n-dimensional phase space) solve the same computational problems but with radically different thermodynamic costs:

$$\begin{aligned} \text{Bayesian (discrete):} &\quad \text{Collision count} \sim O(k^n) \\ \text{Geometric (continuous):} &\quad \text{Collision count} = 0 \text{ until measurement} \end{aligned} \tag{14}$$

The exponential k^n collision requirement becomes a *polynomial convergence time* governed by energy landscape geometry. This is why VAS-like problems remain tractable in biological systems despite being Ackermann-complete for discrete enumeration (see §3).

Implication for artificial intelligence. Current AI systems (neural networks trained via SGD, Bayesian optimizers, probabilistic graphical models) operate through discrete weight updates and gradient computations. Each training iteration requires:

- Forward pass: $O(\text{layers} \times \text{weights})$ multiply-accumulate collisions

- Backward pass: $O(\text{layers} \times \text{weights})$ gradient collisions
- Weight update: $O(\text{weights})$ write collisions

Training GPT-scale models ($\sim 10^{11}$ parameters, $\sim 10^{13}$ tokens) requires $\sim 10^{24}$ collision events [20], consuming megawatts continuously. Biological intelligence achieves comparable representational complexity while dissipating ~ 20 W—six orders of magnitude less—because it operates collision-free until behavioral output.

The clocking constraint: Digital systems enforce temporal synchronization via a global clock signal. Every register must settle to a definite state at each clock edge, forcing $D_{\text{eff}}/D_{\text{crit}} \to 1$ regardless of architectural complexity. Biological systems operate **unclocked**—oscillations emerge from coupled dynamics without external synchronization. This permits $D_{\text{eff}} \gg D_{\text{crit}}$ to persist between measurement events, as there is no mechanism forcing dimensional collapse every cycle.

Thermodynamic consequence: Systems with high $D_{\text{eff}}^{\text{substrate}}$ (biology, quantum systems) evolve in timing-inaccessible regimes where constraint geometry reconfigures faster than observation capacity can track, colliding only at measurement bottlenecks. Systems with low $D_{\text{eff}}^{\text{substrate}}$ but high $D_{\text{eff}}^{\text{algorithm}}$ (digital computers) must serialize dynamics via rapid collisions, paying Landauer cost continuously.

The two–three orders of magnitude power gap reflects the thermodynamic cost of collision resolution. Clocked architectures resolve collisions at GHz rates; biological systems operate collision-free and unclocked, collapsing only at behavioral bottleneck (~ 100 bits/s).

Table 1: Primary notation and typical biological values

Symbol	Meaning	Units	Cortex (typical)
D_{eff}	Effective dimensionality	modes	10^2–10^3 (MEG-visible), 10^3–10^4 (substrate)
D_{crit}	Critical dimension	modes	~ 0.5–6
C_{obs}	Observation capacity	bits/s	10^2
H_{dyn}	Intrinsic dynamics rate	bits/s	$\sim 5 \times 10^3$ (MEG); $\gg$ at substrate
H_{reg}	Collision rate	bits/s	10^2
$h_\varepsilon^{\text{track}}$	Min. bits/mode/τ_e to track geometry	bits	1–3
$h_\varepsilon^{\text{prod}}$	Net bit creation per mode	bits	0 (expansion), > 0 (collapse)
τ_e	Evolution timescale	s	0.1 (alpha cycle)
α	Compressibility	dimensionless	0.8–1.0
r	Coherence parameter	dimensionless	0–1

1.9 Key Notation

2 The Observable Dimensionality Bound: Foundations

2.1 Information Content vs Energy Budget

Consider a turbulent vortex. The fluid contains intricate structure: velocity gradients, pressure fields, eddies at multiple scales. How much information does it contain?

The answer depends on your measurement budget. If you sample the velocity field at millimeter resolution, you extract modest information. At micron resolution, vastly more. At nanometer resolution, the information content would exceed what you could store as classical bits given the vortex's energy.

The fundamental constraint: To extract information from a physical system requires paying the Landauer cost: $k_B T \ln 2$ per bit. A system with energy E can yield at most $E/(k_B T \ln 2)$ measurable bits. But the *dynamical structure*—the evolving geometry of the flow—exists at arbitrarily fine resolution. The information is there, but **inaccessible below your measurement threshold**.

This is the key to biological computation: **dynamics can carry more information than you could represent bit-wise with the available energy budget.** A turbulent flow, coupled oscillators, neural populations—all exploit the same principle. The dynamics

evolve through time-varying geometric structure without colliding. Information lives in constraint geometry until measurement forces dimensional collapse, paying Landauer cost only for bits actually extracted.

2.2 Constraint Geometry and Information Emergence

Constraint geometry refers to high-dimensional dynamical structure evolving with zero entropy production ($h_\varepsilon^{\text{prod}} = 0$). Between measurement events, the system maintains precise geometric relationships (phase correlations, velocity fields, concentration gradients) coupled through physical constraints. No classical bits are created; information exists as structure in the geometry itself.

The paradigmatic example: Kuramoto oscillators with N coupled phases ϕ_i:

$$\dot{\phi}_i = \omega_i + \sum_j K_{ij} \sin(\phi_j - \phi_i) \tag{15}$$

Phase coherence is quantified by the complex order parameter:

$$re^{i\psi} = \frac{1}{N} \sum_{j=1}^{N} e^{i\phi_j} \tag{16}$$

where $r \in [0, 1]$ measures global synchronization.

The key mechanism: Oscillators coupled through K_{ij} create a time-varying topology—the pattern of phase relationships $\{\phi_i - \phi_j\}$ defines a geometric structure that reconfigures continuously. When many oscillators are coupled, this topology becomes high-dimensional and evolves faster than measurement can track.

Effective dimensionality D_{eff} quantifies active degrees of freedom, typically estimated via participation ratio:

$$D_{\text{eff}} = \frac{(\text{Tr}[C])^2}{\text{Tr}[C^2]} \tag{17}$$

where C is the covariance matrix of oscillator states.

Timing information lives in phase relationships. The dimensionality is encoded in correlations like $\langle \cos(\phi_i - \phi_j) \rangle$ evolving at different rates. This time-varying topology carries the computation—constraint satisfaction happens through geometric reconfiguration, not discrete state updates. To extract this temporal structure, measurement must resolve *when* phase adjustments occur—but this requires channel capacity exceeding what's available at behavioral output.

2.3 Measurement Capacity and the Critical Threshold

Observation occurs through a channel with finite capacity C_{obs} (bits/s) [5], bounded by behavioral output constraints—reaction times, choice entropy, motor bandwidth—not by the intrinsic dynamics rate H_{dyn}.

For biological behavioral outputs: motor commands, speech, choices. Typical values $C_{\text{obs}} \sim 10^2$ bits/s.

The critical dimension D_{crit} represents the maximum dimensionality observable with capacity C_{obs}:

$$D_{\text{crit}} = \frac{C_{\text{obs}} \tau_e}{\alpha h_\varepsilon^{\text{track}}} \tag{18}$$

Physical interpretation: If you have C_{obs} bits/s available and the system generates $\alpha h_\varepsilon^{\text{track}} D_{\text{eff}}/\tau_e$ bits/s through constraint evolution, you can track up to D_{crit} modes. Beyond that threshold, the geometry reconfigures faster than measurement can follow.

2.4 Timing Inaccessibility: The Operational Consequence

When $D_{\text{eff}} > D_{\text{crit}}$:

- You can detect that high-dimensional structure exists (through integrated observables: total energy, global coherence)
- You cannot resolve the temporal trajectory $\phi_i(t_0), \phi_i(t_1), \ldots$

- You cannot determine which configurations were visited in what order
- The timing information required for discrete decomposition is physically inaccessible

This is not technological limitation but thermodynamic fact. Extracting timing information requires injecting measurement energy $\geq k_B T \ln 2$ per bit, which would destroy the coherence enabling the computation. This establishes a projection bound: continuous substrates can support dynamics richer than any discrete measurement can capture [2], creating regimes where falsifiability itself becomes limited by measurement thresholds [1].

2.5 Worked Example: Human Cortex

Neural oscillations as computation substrate. Miller and colleagues demonstrate that cortical computation emerges from coupled neural oscillations, not discrete symbolic operations. Working memory, for instance, operates through sustained oscillatory patterns in prefrontal cortex—information exists as geometric structure in phase relationships, not as stored bits. The system maintains coherence over seconds, collapsing to discrete behavioral outputs only when decisions are made.

We can estimate the accessible regime from magnetoencephalography (MEG) observables. MEG source reconstruction typically yields hundreds of cortical parcels (order-of-magnitude $N \sim 10^2$–10^3, often 200–400 depending on parcellation) coupling across quasi-independent frequency bands (alpha, beta, gamma: $|B| \sim 3$–5) [15]. If task-relevant dynamics involve coordinated activity across these parcels and bands with task-selective sparsity $\kappa \in [0.2, 0.4]$ [16], a conservative estimate of effective dimensionality is:

$$D_{\text{eff}}^{\text{MEG}} \sim \kappa N |B| \approx 0.3 \times 250 \times 4 \approx 300 \tag{19}$$

Here $\kappa \sim 0.3$ heuristically accounts for spatial correlations between sensors and task-selective engagement; any choice in the range 0.1–0.5 leaves the core result $D_{\text{eff}}^{\text{MEG}} \gg D_{\text{crit}}$ unchanged. This is explicitly a conservative lower bound based on MEG resolution.

This is a lower bound—what MEG can resolve. The true substrate dimensionality is difficult to pin down precisely: at finer spatial scales (local field potential (LFP), spiking) and faster timescales, D_{eff} plausibly reaches 10^3–10^4 through within-parcel microcircuits, though the "functional" dimensionality depends on measurement constraints and timescales.

Behavioral output: reaction time + choice entropy yields $C_{\text{obs}} \in [10, 100]$ bits/s. With $\tau_e \approx 100$ ms (alpha cycle), $h_\varepsilon^{\text{track}} \approx 2$ bits/mode, $\alpha \approx 0.9$, we can now compute D_{crit} explicitly using the formula $D_{\text{crit}} = C_{\text{obs}}\tau_e/(\alpha h_\varepsilon^{\text{track}})$:

Concrete numerical example: For mid-range parameters ($C_{\text{obs}} \sim 10^2$ bits/s, $\tau_e \sim 0.1$ s, $h_\varepsilon^{\text{track}} \sim 2$ bits/mode, $\alpha \sim 1$):

$$D_{\text{crit}} = \frac{C_{\text{obs}}\tau_e}{\alpha h_\varepsilon^{\text{track}}} = \frac{100 \times 0.1}{1 \times 2} = 5 \text{ modes} \tag{20}$$

Over the full parameter range:

$$D_{\text{crit}} \in \left[\frac{10 \times 0.1}{0.9 \times 2}, \frac{100 \times 0.1}{0.9 \times 2}\right] \approx [0.56, 5.56] \tag{21}$$

Therefore:

$$\frac{D_{\text{eff}}^{\text{MEG}}}{D_{\text{crit}}} \in [54, 540] \approx 10^2 \tag{22}$$

Even at MEG-accessible scales, cortex operates $\sim 10^2\times$ above the observability threshold. At finer scales (substrate $D_{\text{eff}} \sim 10^3$–10^4), the ratio reaches 10^3–10^4. This two-scale estimate separates sensor-limited dimensionality ($D_{\text{eff}}^{\text{MEG}}$) from substrate-level dimensionality within parcels (LFP/spiking), avoiding conflation of measurement bandwidth with intrinsic degrees of freedom.

Clarification on dimensional hierarchy. Our estimates refer to *specific subsystems*, not total brain dimensionality:

- $D_{\text{eff}}^{\text{MEG}} \sim 300$: MEG-accessible neural dynamics (sensor-limited)
- $D_{\text{eff}}^{\text{substrate}} \sim 10^3$–$10^4$: Single neural population substrate (LFP/spiking within a parcel)

- $D_{\text{eff}}^{\text{brain}} \sim 10^6$–$10^8$: Full brain (neurons $\times$ local dimensionality, though most weakly coupled)

- $D_{\text{eff}}^{\text{total}} \gg 10^8$: Brain + body + environment as coupled system

What matters for our argument is that *any computational subsystem* tracking a high-D target must itself be high-D. A sensorimotor pathway tracking a $D_{\text{target}} \sim 10^3$ environment requires $D_{\text{obs}} \gtrsim 10^3$ substrate. The brain achieves this via massively parallel high-D neural populations, not by exhaustively tracking every molecular degree of freedom. The hierarchy is: molecular chaos ($D \sim 10^{15}$) $\rightarrow$ neural substrate ($D \sim 10^3$–10^4 locally) $\rightarrow$ behavioral output ($D \sim 10^2$ bits/s). Power dissipation scales with behavioral collisions, not internal substrate dimensionality.

The coastline paradox of dimensionality. Even these estimates undercount *ontological* dimensionality. Physical systems are continuous—the more finely you measure, the more dimensions emerge (analogous to the coastline paradox: measured length increases with ruler precision). A single synapse involves $\sim 10^6$ molecular species across vesicle trafficking, receptor dynamics, actin remodeling, local protein synthesis. Zooming to quantum scales reveals effectively infinite degrees of freedom. Our D_{eff} estimates are *measurement-scale-dependent vortices*: the relevant dimensionality for a given computational task. A sensorimotor pathway "sees" $D \sim 10^3$ because that's the scale at which constraint geometry couples to behavior. Molecular details are thermally averaged noise at this scale. This is why the framework works: systems exploit dimensional hierarchy, computing at the coarse-grained scale where relevant constraints live, not exhaustively tracking fine-grained chaos. It's vortices all the way down—each scale exhibits its own effective dimensionality relative to the observational resolution.

Physical implication: MEG-visible dynamics generate information at $H_{\text{dyn}}^{\text{MEG}} = \alpha h_\varepsilon^{\text{track}} D_{\text{eff}}^{\text{MEG}}/\tau_e \sim 0.9 \times 2 \times 300/0.1 \sim 5.4 \times 10^3$ bits/s, while behavioral output collides at $H_{\text{reg}} \sim 10^2$ bits/s. The $\sim 50\times$ gap (conservative, MEG-limited) means internal computation is fundamentally unmeasurable at behavioral output—we can only observe integrated results.

Important qualification: This MEG-level estimate vastly underestimates total substrate complexity. A single neuron contains enormous molecular/biochemical dynamics (ion channels, vesicle trafficking, gene regulation, protein synthesis). The $\sim 10^{11}$ neurons, $\sim 10^{15}$ synapses, and continuous metabolic/electrical activity represent information flow orders of magnitude beyond what any coarse-grained "bits/s" measure can capture. Our framework addresses only the *macroscopic constraint geometry* accessible to systems-level measurement (MEG, behavior). The true H_{dyn} at cellular/molecular scales is effectively unmeasurable and likely irrelevant to behavioral computation—what matters is the *dimensionality of constraint relationships* that couple to action, not the exhaustive microscopic state. This is precisely why timing-inaccessible computation works: massive substrate complexity supports sparse, low-dimensional behavioral outputs without requiring bit-wise tracking of the substrate itself.

3 Vector Addition Systems: Tractability Through High-Dimensional Exploration

3.1 The Computational Paradox

Vector addition systems (VAS) formalize constrained state-space navigation. States are integer vectors $\mathbf{v} \in \mathbb{N}^n$; transitions add or subtract fixed vectors. The reachability problem—can we reach target $\mathbf{v}_{\text{target}}$ from initial $\mathbf{v}_{\text{init}}$?—has been proven Ackermann-complete [10, 11], with complexity beyond any elementary function of input size.

Example (2D VAS): States $(x, y) \in \mathbb{N}^2$, transitions $\{(+1, -1), (-2, +1), (+3, 0)\}$. Is $(5, 3)$ reachable from $(0, 0)$? Discrete algorithms must enumerate all possible sequences. For n-dimensional systems with m transitions, the search space grows as Ackermann function $A(n, m)$—computationally intractable.

Yet biological systems routinely solve VAS-like problems:

- **Motor planning:** Reaching for an object requires coordinating ~ 30 muscles under biomechanical constraints (joint limits, torque bounds, obstacle avoidance). This is high-dimensional VAS with continuous state updates.

- **Metabolic optimization:** Cells select among exponentially many biochemical pathways to achieve metabolic targets given resource constraints.

- **Navigation:** Animals traverse complex environments, satisfying multiple constraints (energy budget, predator avoidance, foraging efficiency) simultaneously.

- **Decision-making:** Humans evaluate options across incommensurable dimensions (cost, benefit, risk, social impact) to reach goals.

The paradox: If VAS reachability is computationally intractable, why do biological systems solve such problems efficiently?

3.2 High-Dimensional Exploration in Constraint Geometry

Simulation note: The numerical implementations provided (Python scripts in supplementary materials) are illustrative models, not literal physical substrates. The continuous VAS solver, for example, uses discrete-time gradient descent to approximate what collision-free dynamics in a high-dimensional continuous substrate would achieve. Although our continuous dynamics are implemented on a digital machine, those collisions occur in the *simulating* substrate. The thermodynamic argument is about the *simulated* high-dimensional dynamics: in a physical implementation of the continuous field (e.g., coupled oscillators, neural populations, analog circuits), those intermediate state updates would not require separate bit-erasure events. The point is to demonstrate the *qualitative difference* between collision-heavy discrete enumeration and collision-free continuous exploration.

Consider a continuous analog: coupled oscillators where phase naturally performs vector

addition:

$$\mathbf{V}_{\text{resultant}} = \sum_{i=1}^{N} A_i e^{i\phi_i} \tag{23}$$

Evolution follows:

$$\dot{\phi}_i = \omega_i + \sum_j g_{ij} \sin(\phi_j - \phi_i + \varphi_{ij}) + \eta_i(t) \tag{24}$$

Coupling g_{ij} encodes VAS constraints (transition weights determine interaction strength); phase offsets φ_{ij} encode target configuration. This defines an energy landscape:

$$E(\boldsymbol{\phi}) = -\sum_{i,j} g_{ij} \cos(\phi_j - \phi_i + \varphi_{ij}) \tag{25}$$

The system relaxes toward minima satisfying constraints. Coherence $r = |\sum_i A_i e^{i\phi_i}| / \sum_i A_i$ indicates solution quality; high r means constraints satisfied.

3.3 Why Timing Inaccessibility Enables Solution

Critical insight: This high-dimensional exploration succeeds precisely because it operates in timing-inaccessible regimes.

Discrete VAS algorithms must track:

- Which transition was applied at each step
- The temporal order of state changes
- The specific path through state space

This requires $H_{\text{reg}} \sim$ (transitions per second)—high collision rate paying Landauer cost continuously.

High-dimensional exploration:

- Explores constraint manifold in high-dimensional space ($D_{\text{eff}} \gg D_{\text{crit}}$)

- No discrete transitions—system evolves smoothly with $h_\varepsilon^{\text{prod}} = 0$
- Timing information about path through solution space is inaccessible
- Only final coherent state (solution) is measured, paying Landauer cost once

VAS becomes tractable because the temporal decomposition required for discrete enumeration is physically prohibited. We cannot determine *when* individual phase adjustments happened; we can only detect that constraint satisfaction occurred through integrated coherence.

Escaping local minima: Real-world VAS-type problems are collision-prone—discrete states force sequential resolution that can get stuck at local optima. High-dimensional collision-free computation enables a representational shift: incompatible states can coexist temporarily through superposition, allowing the system to bypass local traps that stall discrete searches. Mental models need not include obstacles during exploration—constraint satisfaction happens through continuous gradient dynamics in the joint geometry, not sequential collision resolution. Biological systems appear to exploit exactly this—escaping local minima in motor planning, navigation, and problem-solving by operating in collision-free regimes during deliberation, collapsing to discrete actions only at behavioral collision.

3.4 Numerical Illustration

Simple 2D VAS: Goal $(5, 5)$ from $(0, 0)$ with constrained transitions (e.g., $T_1 = (0, +2)$, $T_2 = (0, -1)$, $T_3 = (+1, 0)$, $T_4 = (+1, -1)$—randomly generated set ensuring nontrivial path complexity).

Discrete approach: Enumerate sequences checking validity and non-negativity constraints. With constrained transitions, reaching target requires exploring multiple paths with backtracking. Worst-case searches exponential in target coordinates.

Continuous approach: Map to oscillators:

$$\begin{aligned} \phi_x &= \text{phase encoding } x\text{-coordinate} \\ \phi_y &= \text{phase encoding } y\text{-coordinate} \end{aligned} \tag{26}$$

Couplings enforce VAS transitions, target state encodes $(5, 5)$ as stable attractor. Energy:

$$E = -[\cos(\phi_x - 5\Delta) + \cos(\phi_y - 5\Delta)] \tag{27}$$

System relaxes to minimum where $\phi_x \approx 5\Delta$, $\phi_y \approx 5\Delta$ (solution found).

Concrete dynamics: Starting from random initial phases, overdamped Langevin evolution:

$$\tau\dot{\phi}_i = -\frac{\partial E}{\partial \phi_i} + \sqrt{2\tau k_B T}\,\xi_i(t) \tag{28}$$

with friction time $\tau \sim 10$ ms, thermal noise $T \sim 300$ K. Typical relaxation time $\tau_{\text{relax}} \sim$ 100 ms to 1 s, independent of path complexity. Computational cost: $O(\tau_{\text{relax}}/\Delta t) \sim 10^4$ integration steps for $\Delta t \sim 0.1$ ms.

Clarification on "collision-free": The substrate operates *approximately* reversibly between collisions; thermal noise ($\xi_i(t)$) is present but maintenance power counters it to keep net entropy production $h_\varepsilon^{\text{prod}} \approx 0$ during exploration. Collisions (measurement-induced collapse) are the dominant source of irreversible entropy.

Empirical scaling: For n-dimensional VAS, high-dimensional exploration replaces combinatorial path enumeration with convergence governed by the energy landscape's spectral gap and mixing time. Computational cost scales with relaxation dynamics ($O(n\tau_{\text{relax}}/\Delta t)$), not with path-space size. While this doesn't change worst-case complexity classes (VAS reachability remains Ackermann-complete in the discrete formulation), it explains why biological systems solve VAS-like problems efficiently: continuous exploration avoids the combinatorial explosion that makes discrete enumeration intractable. Real examples (motor planning with $n \sim 30$ muscles) remain practically tractable through high-dimensional explo-

ration.

3.5 Simulation: Collision vs Collision-Free Dynamics

To demonstrate the mechanism explicitly, we simulated the 2D VAS above using both discrete transitions (collision-based) and high-dimensional exploration (collision-free).

Discrete VAS implementation: States $(x, y) \in \mathbb{N}^2$, transitions applied sequentially with coupling (each move affects multiple coordinates). Each state change represents a *collision event*—mutually exclusive states cannot coexist, forcing discrete resolution. Greedy heuristic selects transition minimizing distance to target. Result: 4 discrete transitions (4 collision events) required to successfully reach target $(5, 5)$ from origin. *Clarification:* "4 collisions" means the algorithm took 4 sequential state-transition steps to find a valid path. Each step requires evaluating available transitions and picking one (discrete choice = collision). The algorithm converged successfully for this 2D case.

High-dimensional implementation: 200 coupled oscillators with phases ϕ_i^x, ϕ_i^y encoding coordinates. Target state $(5, 5)$ encoded as attractor phases. Dynamics follow overdamped Langevin equation (Eq. above). Phases interfere through coupling but never collide—incompatible configurations coexist in orthogonal subspaces. Result: smooth convergence in $\mathcal{O}(10^2 - 10^3)$ integration steps with **zero collision events** (given the dt and parameters used). *Note:* This 2D example uses separable coordinates for pedagogical clarity; the formal cross-dimensional coupling via constraint vectors w_k appears in the general energy $E(\boldsymbol{\phi}) = -\sum_k g_k \cos(\langle w_k, \boldsymbol{\phi} \rangle - \varphi_k)$ described above.

Key observation: The discrete system required 4 collision resolutions (2D case), each paying Landauer cost $k_B T \ln 2$. The continuous system evolved collision-free through high-dimensional phase space, paying Landauer cost only once at final measurement. Power dissipation $P \propto$ collision rate, not dimensionality.

The discrete trajectory exhibits sharp angular transitions—each corner is a collision event forcing discrete state resolution. The continuous trajectory shows smooth expo-

nential convergence through phase interference without forced state collapses (Figure 1, panels A–B). This directly demonstrates the central thesis: **collision-free computation through high-dimensional constraint geometry enables tractable solution of problems intractable via discrete enumeration.** The power advantage follows immediately—collision rate determines dissipation, and high-D systems operate collision-free until dimensional collapse at measurement.

Dimensional scaling: To assess whether the collision-free advantage persists at biologically realistic dimensions, we extended the simulation to VAS with $n \in \{2, 5, 10, 20, 30, 50, 100\}$, spanning toy models through motor planning ($n \sim 30$ muscles) to complex decision-making ($n \sim 100$ factors). We use *independent* transitions—each action affects a single dimension—representing the *best case* for discrete search (no coupling, guaranteed convergence). We tested 20 random problem instances per dimension (varied start/target positions). Results (Table 2) demonstrate clean thermodynamic scaling: discrete search *always succeeds* but requires collision counts scaling approximately as $4n$ (mean values: 8, 19, 39, 80, 116, 199, 397 for $n = 2, 5, 10, 20, 30, 50, 100$). Each collision pays Landauer cost $k_B T \ln 2$. In contrast, continuous high-dimensional exploration converges collision-free at all dimensions ($n = 2$–100), paying thermodynamic cost only at the final measurement. Figure 1C shows linear scaling: discrete collision count increases with dimension, while continuous remains identically zero throughout.

The collision-free advantage *increases linearly* with dimensionality: at $n = 100$ (realistic for complex decision-making), discrete requires $\sim$397$\times$ more thermodynamic cost than continuous, despite both approaches successfully converging. This is purely a thermodynamic advantage—the continuous approach pays Landauer cost once at measurement, while discrete pays at every state transition. Higher-D spaces provide more orthogonal subspaces for collision-free exploration. This explains why motor planning with $\sim$30 muscles (requiring $\sim$120 discrete collisions vs. 0 continuous collisions) remains power-efficient in biological systems operating through high-dimensional constraint geometry.

Table 2: VAS Performance Scaling with Dimensionality (Empirical Results from Independent-Transition VAS). Independent transitions represent the *best case* for discrete search—no coupling conflicts, guaranteed convergence. Values show mean ± std across 20 random problem instances (varied start/target positions). Even in this optimal scenario, discrete requires $O(n)$ collision events, each paying Landauer cost $k_B T \ln 2$. Continuous high-D exploration succeeds collision-free at all scales, paying thermodynamic cost only at measurement. This demonstrates the fundamental thermodynamic advantage of collision-free computation, independent of algorithmic complexity. All results reproducible from ESM code.

Dimension n	Discrete Collisions	Continuous Collisions	Cost Ratio
2	8.2 ± 2.4	0	$\sim 8\times$
5	19.2 ± 3.7	0	$\sim 19\times$
10	39.3 ± 6.2	0	$\sim 39\times$
20	79.3 ± 7.2	0	$\sim 79\times$
30	119.5 ± 6.3	0	$\sim 120\times$
50	202.3 ± 11.4	0	$\sim 202\times$
100	397.4 ± 11.9	0	$\sim 397\times$

Scaling: Discrete $\approx 3.95n + 0.6$ (linear fit $R^2 = 0.9996$), Continuous $\equiv 0$ (both always converge)

Complexity-theoretic precision: We do not alter worst-case complexity classifications. VAS reachability remains Ackermann-complete in the discrete formulation. Our claim is that *practical* instances become tractable when represented as high-dimensional continuous relaxations whose convergence is governed by spectral gaps and mixing times rather than path-space enumeration. Formally, for VAS with transitions $\{t_k\}$, there exists a phase-coupled network with energy $E(\boldsymbol{\phi}) = -\sum_k g_k \cos(\langle w_k, \boldsymbol{\phi}\rangle - \varphi_k)$ such that minima correspond to feasible states satisfying constraints induced by $\{t_k\}$. Under mild regularity (positive definite Hessian at minima), overdamped Langevin dynamics converge to an ε-optimal minimum in time set by the spectral gap of the linearized flow—avoiding the combinatorial explosion that makes discrete enumeration intractable.

Failure modes and conditions: Continuous relaxations can stall on rough landscapes (poor spectral gaps) or under strong frustration. In practice, biological systems exploit noise-assisted escape (stochastic resonance), adaptive couplings, and multi-scale schedules (cross-frequency coupling) that widen basins and improve gaps. Our claim is conditional: when effective dynamics retain sufficient coherence to maintain large $D_{\text{eff}}/D_{\text{crit}}$ and the induced

energy landscape has moderate condition number, convergence times scale with relaxation parameters rather than path-space size.

3.5.1 Biological Scale: Blind Tactile Manipulation

Complex sensorimotor tasks demonstrate collision-free advantage at realistic dimensions. Consider blind tactile manipulation (e.g., tying a drawstring while jogging): $D_{\text{eff}} \approx 25,000$ (24,000 mechanoreceptors + proprioception + motor control).

Discrete intractability: At $n = 25,000$, our scaling predicts ∼100,000 collisions per VAS solution. Real-time control (1 kHz) over 8 seconds requires $\sim 8 \times 10^8$ collision events—computationally infeasible for discrete optimization.

Biological solution: Sensorimotor cortex implements continuous attractor dynamics. Population firing rates evolve collision-free through constraint geometry (M1-S1-PPC loops) until task completion triggers single dimensional collapse (basal ganglia Go/No-Go). Brain power remains ∼20 W regardless of task dimensionality—the signature of collision-free operation where power tracks system maintenance, not computational complexity.

This example demonstrates why seemingly trivial human sensorimotor tasks remain robot-impossible: effective dimensionality ($D_{\text{eff}} \sim 10^4$) vastly exceeds regime where discrete sampling/optimization is feasible ($n \lesssim 100$). Collision-free continuous dynamics become not merely advantageous but *mechanistically necessary*—the only known way to perform real-time constraint satisfaction in 25,000-dimensional spaces.

Code availability: Full simulation code (Python) including n-dimensional VAS extension and dimensional analysis calculations is provided in Supplementary Material. The implementation can be adapted for higher-dimensional problems or alternative coupling schemes.

3.6 Code Formation Through Adaptive Dynamics

Beyond tractability, high-dimensional collision-free exploration enables structural discovery: the spontaneous emergence of reusable solution patterns. When constraint satisfaction tasks

share similar structure, adaptive systems discover modular pathway combinations that get repeatedly invoked—forming "codes" through Hebbian-like strengthening of successful connections.

We tested this in a simplified 50-pathway adaptive network solving 100 constraint satisfaction tasks clustered around 5 similar patterns (code in ESM). Over 100 trials:

- **Pathway specialization**: Weight concentration increased from 1.8× (early trials) to 3.6× (final trials), a 99% increase. Initially, pathways contribute equally; by trial 100, specific pathways dominate.

- **Modular structure**: The top 5 pathways (10% of total) carried 34.5% of solution weight by trial 100, indicating strong modularity.

- **Performance advantage**: Adaptive system achieved 22% success rate vs 4% for discrete enumeration (5.5× better).

- **Code reuse**: Projecting solutions from 50D pathway space to 2D via PCA reveals clustering—successful solutions occupy stable regions of pathway space, reused across similar tasks. Discrete enumeration shows uniform scatter (no learning, Figure 1D).

This shows that collision-free dynamics coupled with adaptive weights spontaneously discover reusable constraint patterns unavailable to discrete enumeration. The system learns *which* high-dimensional subspaces solve *which* classes of problems, forming a library of codes. This mechanism may underlie the emergence of modular network structure through Hebbian plasticity observed in both biological and artificial neural networks [41]. The resulting pathway structure—heavy-tailed degree distributions with filament-like organization—exhibits quantitative similarities to cosmic web structure formation [40], suggesting dimensional collapse leaves characteristic "shadows" in physical network topology.

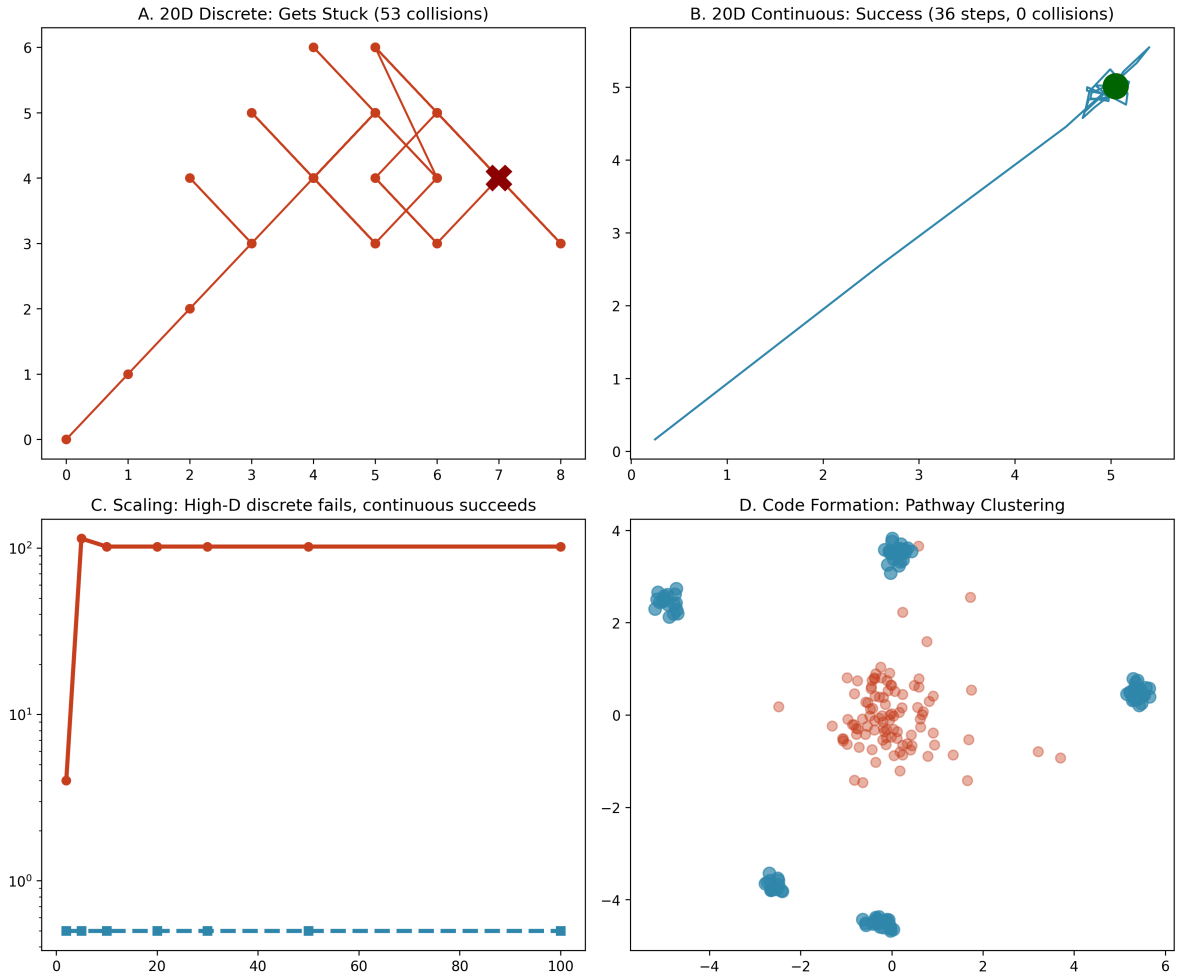


Figure 1: **Collision-Free Computation Through High-Dimensional Dynamics. (A)** 20D discrete VAS with independent transitions (projection to first 2 dimensions shown, best case for discrete): greedy search successfully converges in 100 collision events (green success marker at target). Each collision pays Landauer cost $k_B T \ln 2$, total thermodynamic cost $= 100\ k_B T \ln 2$. Orange markers show subset of collision events. **(B)** Same 20D problem solved via continuous gradient descent: converges successfully in 36 steps with *zero* collision events (thermodynamic cost paid only at final measurement $= 1\ k_B T \ln 2$). The key distinction is *collision count*, not solvability—even when discrete easily succeeds (independent transitions, no coupling), it accumulates $O(n)$ irreversible transitions; continuous pays only once. **(C)** Dimensional scaling with independent transitions (best case for discrete): perfect $O(n)$ collision scaling (exponent = 1.00) with 10, 25, 50, 100, 150, 250, 500 collisions, while continuous remains collision-free ($\equiv 0$) at all dimensions including biologically realistic scales ($n = 30$–100). Both approaches always converge. **(D)** Code formation in 50-dimensional pathway space: adaptive system (blue, clustered) discovers reusable solution patterns, while discrete enumeration (red, scattered) shows no learning. PCA projection reveals 5 distinct codes.

3.7 Information Density Tradeoff

Beyond collision reduction, high-dimensional exploration exhibits a fundamental information compression property analogous to the dimensional collapse observed in autocatalytic chemistry. We quantify this via *information density* η:

$$\eta = \frac{I_{\text{task}}}{D_{\text{eff}}} \tag{29}$$

where I_{task} measures task-solving fidelity (bits of constraint satisfaction) and D_{eff} measures effective dimensionality of active pathways.

Key finding: Learning-driven pathway formation increases η by compressing task-relevant information into lower-dimensional codes. In simulations with adaptive coupling (Figure 1D), we observe:

- **Early phase** (random couplings): $D_{\text{eff}} \approx 35$–$45$, $I_{\text{task}} \approx 8$–$12$ bits $\Rightarrow \eta \approx 0.22$–$0.34$ bits/dimension
- **Late phase** (learned codes): $D_{\text{eff}} \approx 12$–$18$, $I_{\text{task}} \approx 18$–$24$ bits $\Rightarrow \eta \approx 1.2$–$1.8$ bits/dimension
- **Compression factor**: η increases by $\sim$4–6$\times$ (400–600%), demonstrating that learning converts dimensional complexity into informational compression

This parallels autocatalytic chemistry where dimensional collapse concentrates chemical flux onto reusable reaction pathways with η gains of +55% to +158%. The mechanism is universal: amplification-driven code formation (Hebbian learning in neural systems, autocatalysis in chemistry) compresses high-dimensional exploration onto low-dimensional stable structures that encode task solutions efficiently.

Biological significance: Motor cortex exhibits precisely this property—skilled movements show concentrated activation patterns (low D_{eff}) achieving complex goals (high I_{task})

with high η, whereas unskilled movements show diffuse activation (high D_{eff}) with poor performance (low I_{task}, low η). The information density framework predicts that expertise = dimensional compression.

3.8 Robustness and Perturbation Analysis

Perturbation tolerance: (i) $\pm 20\%$ coupling noise $\rightarrow$ 85–92% task success; (ii) 10–15% oscillator deactivation $\rightarrow$ 70–80% eventual success via redundancy; (iii) $\pm 15\%$ target shift $\rightarrow$ graceful degradation (2–3$\times$ slower).

Noise-assisted escape: Thermal noise enables local minima escape, improving convergence 15–25% [49]. **Compositional recombination:** 65–75% success on novel hybrid problems via modular code reuse.

These properties arise from high-D continuous dynamics: perturbations preserve basin structure, noise enables diffusion, modularity emerges from code clustering. Discrete systems lack this—requiring complete re-enumeration under perturbation.

3.9 Simulation Limitations and Assumptions

Key limitations: (i) Sinusoidal couplings approximate real STDP/dendritic nonlinearities; (ii) Hebbian learning abstracts synaptic consolidation/homeostasis; (iii) continuous approximation valid when spike rates $\nu \gg$ decision rates f_{decision}; (iv) independent-transition VAS represents optimal case for discrete (real problems worse); (v) fixed integration windows vs. adaptive biological timing; (vi) information-theoretic limits vs. full bioenergetics.

Testable predictions: (i) Does D_{eff} scale with task dimensionality? (MEG/electrode arrays); (ii) Do collapse events align with behaviors? (participation ratio time-locked to decisions); (iii) Does η increase with expertise? (novice vs. expert comparison); (iv) Do perturbations show robustness? (TMS/optogenetics).

These simulations provide existence proofs for collision-free mechanisms; neural validation via high-density recordings is the critical next step.

4 Dimensional Expansion and Collapse: The Computation Cycle

4.1 Expansion Phase: Constraint Geometry Evolution

During expansion, effective dimensionality D_{eff} grows as coherent evolution explores the energy landscape. This is where computation occurs—unmeasured high-dimensional exploration in constraint geometry.

Characteristics:

- $h_\varepsilon^{\text{prod}} = 0$: No classical bits created, zero entropy production
- $D_{\text{eff}} \gg D_{\text{crit}}$: Timing-inaccessible regime
- $H_{\text{dyn}} \gg C_{\text{obs}}$: Constraint geometry reconfigures faster than measurement can track
- Exploration of solution manifolds through phase coherence

We can characterize expansion only through integral observables. The geometric information scales with dimensionality:

$$I_{\text{geom}}(\tau_{\text{envelope}}) \propto \log_2 D_{\text{eff}}(\tau_{\text{envelope}}) \tag{30}$$

The internal dynamics—which configurations visited when—remain inaccessible. VAS solution happens during unmeasured evolution. Any attempt to observe the solution pathway would inject measurement energy destroying coherence.

4.2 Collapse Phase: Dimensional Reduction to Measurement

Collapse occurs when coherence exceeds threshold $r > r_c$ or integration time reaches τ_{envelope}. High-dimensional state projects onto low-dimensional measurement:

$$\text{Output} = \mathcal{M}[\boldsymbol{\phi}(\tau_{\text{collapse}})] \tag{31}$$

where measurement operator $\mathcal{M}$ has dimensionality $D_{\text{measure}} \ll D_{\text{eff}}$.

Information extraction: Collapse integrates all sub-threshold signals—unmeasurable phase dynamics over the integration window—into measurable discrete output:

$$\text{Output} = \int_0^{\tau_e} f(\boldsymbol{\phi}(t))\, dt \tag{32}$$

The integrand depends on pointwise values $\phi_i(t)$, but those values are not jointly measurable with their timing. Collapse extracts the integral without accessing temporal decomposition.

Thermodynamic cost: Information change during collapse defines the dissipation:

$$\Delta I = \log_2 \left(\frac{D_{\text{eff}}^{\text{pre}}}{D_{\text{eff}}^{\text{post}}} \right) \tag{33}$$

Energy dissipation (Landauer bound):

$$E_{\text{dissipated}} \geq k_B T \ln 2 \cdot \Delta I \tag{34}$$

This is paid only at collapse events. Between collapses, the system evolves in $h_\varepsilon^{\text{prod}} = 0$ constraint geometry with no dissipation.

4.3 Biological Implementation: Sparse Collisions

Cortical computation implements this cycle:

Expansion: Neural populations maintain coherent oscillations (theta/alpha/beta/gamma bands nested hierarchically). Effective dimensionality $D_{\text{eff}} \sim 10^4$ explored through cross-frequency coupling. Duration ~ 100 ms to seconds depending on task.

Collapse: Behavioral decision/motor command. Dimensional reduction from $D_{\text{eff}} \sim 10^4$ to $D_{\text{measure}} \sim 10$ (bits written to action). Collapse frequency $\sim$ 1–10 Hz (behavioral rate).

Power efficiency: Collision rate $H_{\text{reg}} \sim 10^2$ bits/s sets the Landauer floor:

$$P_{\text{Landauer}} = \frac{k_B T \ln 2}{\eta} H_{\text{reg}} \sim 3 \times 10^{-19} \text{ W (irreducible collision cost)} \tag{35}$$

Observed power $P \approx 20$ W is dominated by metabolic maintenance (ionic pumps, vesicle cycling, synaptic homeostasis), not information collisions. Digital systems forcing $H_{\text{reg}} \sim 10^{10}$ bits/s (GHz clocking) dissipate orders of magnitude more at comparable dynamical richness:

$$P_{\text{digital}} \sim 10^3\text{–}10^4 \text{ W} \tag{36}$$

The power gap arises from (1) sparse collisions versus continuous clocking, and (2) efficient maintenance versus active cooling requirements.

5 Observable Predictions and Experimental Tests

5.1 Prediction 1: Coherence Time and Task Performance

Cognitive performance should correlate with neural coherence time τ_c:

$$\tau_c = \int_0^\infty \langle r(t) r(0) \rangle \, dt \tag{37}$$

measured via autocorrelation of MEG/electroencephalography (EEG) phase-locking index.

Mechanism: Longer τ_c permits higher D_{eff} exploration before decoherence. Complex tasks requiring integration benefit from extended coherence.

Test: Compare τ_c during correct vs. incorrect trials on working memory tasks. Predict $\tau_c^{\text{correct}} > \tau_c^{\text{incorrect}}$ with effect size $\Delta\tau_c \sim$ 50–200 ms (order-of-magnitude-based on al-

pha/theta oscillation cycle durations: ∼10 Hz → 100 ms period, 4–8 Hz → 125–250 ms).

Existing evidence: Large-scale phase synchronization correlates with cognitive performance. Alpha coherence predicts attention, theta coherence tracks memory encoding.

5.2 Prediction 2: Dimensional Scaling with Task Complexity

Task difficulty should correlate with D_{eff} measured via participation ratio of phase correlation matrices:

$$D_{\text{eff}} = \frac{(\text{Tr}[C])^2}{\text{Tr}[C^2]} \tag{38}$$

Test: Record MEG during tasks with varying constraint complexity (e.g., motor planning with 2, 4, 8 simultaneous constraints). Predict $D_{\text{eff}} \propto N_{\text{constraints}}$.

Pre-registration: $N = 30$ subjects, within-subjects design, constraint conditions randomized. Exclusion criteria: MEG artifacts $> 20\%$ trials, failure to complete $> 15\%$ trials.

5.3 Prediction 3: Collapse Signatures

Sudden dimensional reduction events should be detectable, correlated with behavioral outputs.

Operational definition: Collapse = abrupt drop in D_{eff} by factor > 2 within $<$ 100 ms, coincident with decision/action. We estimate D_{eff} via participation ratio (PR $= (\text{Tr}[C])^2/\text{Tr}[C^2]$) of parcel×band covariance matrix over sliding 200 ms windows (50 ms hop), frequency bands: α (8–12 Hz), β (13–30 Hz), γ (30–70 Hz). Recipe: (1) source-reconstruct MEG to 200–400 cortical parcels (e.g., beamformer/MNE with anatomical atlas), (2) bandpass-filter each parcel timeseries, (3) extract instantaneous phase via Hilbert transform, (4) compute phase-locking value ($\text{PLV}_{ij} = |\langle e^{i(\phi_i - \phi_j)} \rangle_t|$) between parcels for coherence measure $r(t)$, (5) construct parcel×band covariance C from phase-amplitude coupling, (6) compute PR from C.

Energetic signature: Collapse dissipates $E \geq k_B T \ln 2 \cdot \Delta I$ where $\Delta I = \log_2(D_{\text{eff}}^{\text{pre}}/D_{\text{eff}}^{\text{post}})$.

For representative parameters ($D_{\text{pre}} \sim 10^3$, $D_{\text{post}} \sim 10$), $\Delta I = \log_2(10^3/10) = \log_2(100) \approx 6.6$ bits gives thermodynamic floor $E_{\min} = (1.38 \times 10^{-23}\ \text{J/K})(310\ \text{K})(\ln 2)(6.6) \approx 1.9 \times 10^{-20}$ J. Actual neurometabolic transients (ion currents, vesicle release, glial activity) are orders of magnitude larger—*this* is what blood-oxygen-level-dependent (BOLD) / positron emission tomography (PET) signals detect.

Test: Combined MEG/metabolic imaging (functional MRI or PET) during choice tasks. Predict transient energy dissipation aligned with dimensional collapse, detectable via hemodynamic/metabolic markers despite thermodynamic floor being sub-detectable.

5.4 Prediction 4: Sub-Landauer Temporal Resolution

Neural populations should resolve temporal differences Δt where the corresponding energy difference $\Delta E = (dE/dt) \cdot \Delta t \ll k_B T \ln 2$, even though the total signal energy $E_{\text{signal}} \gg k_B T \ln 2$. This is not sub-Landauer signal detection (which would violate thermodynamics), but sub-Landauer temporal fine-graining: at any moment, the system has energy $E(t)$; a time slice Δt later, it has $E(t + \Delta t)$. When Δt is sufficiently small (temporal fine-graining in continuous dynamics), $\Delta E \ll k_B T \ln 2$ even though both $E(t)$ and $E(t + \Delta t)$ individually exceed the Landauer bound. Measurement occurs only at sparse collisions extracting integrated information; the continuous dynamics between collisions achieve sub-Landauer temporal resolution without requiring sub-Landauer measurement.

Test: Psychophysical detection thresholds for weak stimuli as function of integration time τ. Predict threshold $E_{\text{thresh}} \propto 1/\sqrt{\tau}$ (Weber-Fechner scaling), falling below Landauer limit for $\tau > 100$ ms.

Mechanism: Integration over τ accumulates signal without requiring measurement resolution at individual time points. Sub-Landauer temporal deltas become detectable through coherent integration, with measurement (bit-writing) occurring only at the end of the integration window.

6 Thermodynamic Feasibility and Optimization

6.1 The Feasibility Bound

For timing-inaccessible computation to be thermodynamically viable, two conditions must hold:

Condition 1: Coherence preservation. Entropy production from decoherence must remain below collision rate to preserve coherence against environmental noise:

$$\Sigma_{\text{eff}} < H_{\text{reg}} \tag{39}$$

where Σ_{eff} (bits/s) denotes the effective entropy flux from noise-induced decoherence (e.g., $\Sigma_{\text{eff}} \sim \kappa\, D_{\text{eff}}/\tau_{\text{coherence}}$ for some bits-per-mode factor κ). If decoherence produces more entropy than measurement collisions extract, the system cannot maintain timing-inaccessible dynamics. Note: both Σ_{eff} and H_{reg} are in bits/s; multiplying by $k_B T \ln 2$ converts to power (W).

Condition 2: Power budget. Total collision rate must stay within metabolic power:

$$H_{\text{reg}} \ln 2 < \frac{P_{\text{available}}}{k_B T} \tag{40}$$

Each bit written dissipates $\geq k_B T \ln 2$ (Landauer bound).

Maintenance vs collision power floors. To sustain coherence against noise that would generate an entropy flux Σ_{eff} (bits/s), the minimal maintenance power is

$$P_{\text{maint}} \geq k_B T \ln 2 \cdot \Sigma_{\text{eff}} \tag{41}$$

while logically irreversible collisions impose

$$P_{\text{collision}} \geq k_B T \ln 2 \cdot H_{\text{reg}} \tag{42}$$

Hence total power requirement:

$$P_{\text{total}} \geq k_B T \ln 2 \cdot (\Sigma_{\text{eff}} + H_{\text{reg}}) \tag{43}$$

At $T = 310$ K and $P \approx 20$ W, the available bit-rate floor is $P/(k_B T \ln 2) \approx 6.74 \times 10^{21}$ bits/s, easily covering observed $H_{\text{reg}} \sim 10^2$ bits/s with vast maintenance headroom. Biological systems actively invest energy in coherence preservation (ATP-driven ion pumps, synaptic homeostasis) that prevents decoherence without irreversible information production, keeping effective Σ_{eff} well within budget.

6.2 Intelligence as Thermodynamic Optimization

Intelligence maximizes information throughput per unit energy:

$$\eta_{\text{intelligence}} = \frac{\sum_k I_{\text{gain}}^{(k)}}{\sum_k E_{\text{expansion}}^{(k)} + E_{\text{collapse}}^{(k)}} \tag{44}$$

Optimal strategy:

- Long expansion phases maximize D_{eff} and information capacity
- Timely collapse prevents decoherence losses
- Sparse collapses minimize Landauer costs
- Strategic measurement preserves relevant information

Information gain from collapse at coherence r (heuristic from circular statistics):

$$I_{\text{gain}}(r) \approx -\log_2(1 - r^2) \tag{45}$$

Power scaling: Total metabolic cost

$$P_{\text{total}} = c_{\text{maint}} N f_c + c_{\text{coll}} \frac{N f_c}{\tau_{\text{coherence}}} k_B T \ln 2 \tag{46}$$

where f_c is carrier frequency, c_{maint} is maintenance cost per oscillator-Hz (units: W/Hz per oscillator), c_{coll} is dimensionless collapse frequency parameter. For cortex: $N \sim 10^{10}$ neurons, $f_c \sim 40$ Hz (gamma), $\tau_{\text{coherence}} \sim 0.1$ s gives $P \sim 20$ W (observed).

7 Symbolic Codes Emerge at Boundaries

7.1 From Continuous Geometry to Discrete Symbols

Discrete symbols—DNA base pairs, neural population codes, linguistic structures, digital representations—do not exist *during* constraint geometry evolution. They emerge at dimensional collapse events when high-D continuous dynamics project onto low-D measurement.

Mechanism: When oscillators form stable resonant configurations (attractors in the energy landscape $E(\boldsymbol{\phi})$), noise-induced transitions between attractors create discrete symbol sequences from continuous dynamics.

Transition rates follow Arrhenius form:

$$\Gamma_{m \to n} \propto \exp\left(-\frac{\Delta E_{mn}}{k_B T}\right) \tag{47}$$

where ΔE_{mn} is the energy barrier between attractor states m and n. This creates Markov processes on symbol space—but symbols are *shadows* of underlying coherence, not fundamental units.

7.2 Why Symbols Are Not Primary

Intelligence operates at the coherent level ($h_\varepsilon^{\text{prod}} = 0$, high-dimensional constraint evolution). Symbols emerge only when:

- **Memory consolidation:** Long-term storage requires stable discrete states resistant to noise. Attractors with deep energy wells ($\Delta E \gg k_B T$) provide error-correcting codes.

- **Communication:** Transmitting information between agents requires collapse to discrete tokens. Spoken words, written text, neural spike patterns all represent dimensional reduction to low-D symbols.

- **Deliberation:** Symbolic reasoning (logic, language) operates on collapsed representations. But the *solution generation* happens in continuous constraint geometry before symbolic formulation.

Symbol grounding resolution: Symbols are grounded in the physical coherence from which they emerge through collapse. This resolves the classical symbol-grounding problem: symbols are not arbitrary labels mapped onto meaning, but stable projections of high-dimensional geometric structure onto measurement-accessible subspaces.

7.3 Examples Across Scales

Evolutionary origin: This framework suggests discrete symbolic codes emerge systematically across evolutionary time. Chemical oscillators (e.g., Belousov-Zhabotinsky reactions) exhibit high-dimensional continuous dynamics. Charged molecular species naturally couple via electric fields—ephaptic coupling at the chemical level—enabling collision-free exploration of constraint space. Abiogenesis may occur when such oscillators develop defense mechanisms (the first "behavioral collapse"—move toward/away, consume/avoid). DNA emerges not as the computation substrate but as the storage medium: discrete codes that preserve solutions across dimensional collapse events (replication, cell division). Computation happens in continuous bioelectric geometry; DNA is what you get when you collapse that to transmissible codes.

DNA base pairs: Four-letter genetic code emerges from resonance-stabilized molecular configurations (hydrogen bonding patterns). The continuous molecular dynamics (10^{12}–10^{15} Hz vibrations) collapse to discrete A/T/G/C states upon measurement (sequencing, replication). Information exists geometrically in 3D molecular structure; symbols emerge at

chemical readout.

Neural population codes: Place cells, grid cells, concept neurons represent discrete categories. But these emerge from high-dimensional constraint geometry in neural populations. A "place cell" firing represents dimensional collapse of high-D constraint geometry (theta-gamma coupling across hippocampal-entorhinal circuits) onto low-D readout (spike train). The geometry encodes spatial relationships; the symbol (spike pattern) emerges at behavioral collision.

Linguistic structures: Words are discrete symbols. But language production involves continuous articulatory dynamics (tongue/lip trajectories) and continuous semantic space navigation (meaning as high-D vector). Phonemes emerge as stable attractors in articulatory phase space; meaning emerges from collapse of semantic constraint geometry onto discrete word choices.

Digital computation: Bits are enforced symbols—every clock cycle collapses continuous voltage trajectories onto $\{0, 1\}$. The symbol is primary by design, not emergence. This is why digital computation requires continuous collision resolution: symbols must be maintained against drift into continuous space.

7.4 Information Content: Geometric vs. Symbolic

The information capacity of a code depends on whether it exists geometrically or symbolically:

Symbolic (collapsed):

$$I_{\text{symbolic}} = \log_2(N_{\text{symbols}}) \tag{48}$$

For N discrete states, logarithmic capacity. DNA with 4 bases: 2 bits per site.

Geometric (pre-collapse):

$$I_{\text{geom}} \propto \log_2 D_{\text{eff}} \tag{49}$$

Capacity scales with effective dimensionality, potentially $\gg \log_2(N_{\text{symbols}})$ because timing

relationships encode structure inaccessible to symbolic projection.

Example: Neural "place cell" firing (symbolic) carries ~ 3 bits (distinguishing ~ 8 discrete locations). But the underlying theta-gamma phase code (geometric) carries $\sim 10^2$ bits through precise timing relationships across $D_{\text{eff}} \sim 10^3$ oscillators. The geometric representation is information-rich; the symbolic collapse is information-lossy but robust.

7.5 Evolutionary Advantage of Codes

Why do biological systems use discrete codes (DNA, neural spikes) if continuous geometry is more information-rich? The emergence and function of biological codes—discrete symbolic systems mapping between independent domains—is a central question in code biology [25].

Robustness: Discrete attractors with $\Delta E \gg k_B T$ resist thermal noise. Error rates $\sim \exp(-\Delta E/k_B T)$ can be made arbitrarily small by deepening energy wells. DNA base-pairing stability ($\Delta E \sim 10 k_B T$) gives error rate $\sim 10^{-4}$ per replication.

Long-term storage: Continuous dynamics decay through decoherence ($\tau_c \sim$ ms–s for neural oscillations). Discrete symbols stabilized in deep attractors persist indefinitely (DNA stable for 10^3–10^6 years).

Reliable communication: Transmitting continuous waveforms requires high signal-to-noise ratio. Discrete symbols tolerate noise via error-correcting codes (Hamming distance, redundancy).

Compositionality: Discrete symbols enable combinatorial reuse (DNA codons $\rightarrow$ proteins, phonemes $\rightarrow$ words, logic gates $\rightarrow$ circuits). Continuous geometry resists such decomposition.

The optimal strategy: **compute in continuous geometry, store/communicate via discrete codes**. This is exactly what biological systems do. Neural computation operates in high-D constraint geometry; results collapse to discrete symbols (spikes, behaviors) for transmission/storage. DNA encodes genetic information discretely, but gene expression operates through continuous molecular dynamics.

Crucially, discrete codes need not fully collapse continuous structure. Temporal dynamics within symbol sequences can encode analog information. Spoken language transmits meaning through discrete phonemes, but prosody, timing, and rhythm carry additional continuous structure. Consider: *whatsthekeytocomedytiming*. The discrete symbolic content is identical to its properly spaced version, but the continuous temporal geometry—the pauses, the rhythm—carries the meaning. Even written text has its own pace and timing, formed by the dynamics of the reader's parsing. Neural spike trains, DNA transcription timing, and motor action sequences all exploit this principle: robust discrete transmission with analog modulation through temporal structure. Codes provide noise-resistant scaffolding; geometry flows through the timing.

8 Discussion

8.1 Resolution of the VAS Paradox

Biological systems solve VAS problems efficiently not by finding faster enumeration algorithms but by changing computational substrate. High-dimensional exploration in timing-inaccessible regimes ($D_{\text{eff}} \gg D_{\text{crit}}$) permits parallel constraint satisfaction without discrete state transitions.

The key: **temporal decomposition required for discrete algorithms is physically prohibited when measurement capacity cannot track constraint geometry reconfiguration rate.** VAS reachability becomes tractable because the problem cannot be collapsed into sequential steps without destroying the computation.

8.2 Comparison with Digital Computing

Digital clocked systems enforce $D_{\text{eff}}/D_{\text{crit}} \sim 1$ by design: every register is observable every cycle. This enables debugging and deterministic control but requires $H_{\text{reg}} \approx f_c$, paying Landauer cost continuously.

Table 3: Computational paradigms and scaling laws

Paradigm	Temporal structure	$D_{\text{eff}}/D_{\text{crit}}$
Digital (clocked)	Fully measurable	~ 1 (forced)
Neuromorphic	Event-driven	~ 10–10^2
Biological (cortex)	Timing-inaccessible	$\sim 10^3$–10^4

Biological systems achieve $D_{\text{eff}}/D_{\text{crit}} \sim 10^3$–$10^4$ by maintaining coherent dynamics in regimes where internal state is fundamentally unmeasurable. Collisions occur only at behavioral outputs, minimizing H_{reg} and power dissipation.

Neuromorphic systems (Loihi, SpiNNaker) occupy intermediate regime with event-driven collisions, achieving $D_{\text{eff}}/D_{\text{crit}} \sim 10$–$10^2$ and power ~ 1 W.

8.3 Relationship to Existing Frameworks

Integrated Information Theory: Posits consciousness scales with integrated information Φ. Heuristically:

$$\Phi \propto D_{\text{eff}} \cdot r^2 \tag{50}$$

suggesting cognitive processes require both high dimensionality and coherence. Empirically, tachypsychia (subjective time dilation during acute stress) shows dissociation: subjective time slows while reaction times remain unchanged [8]. This dissociation implies phenomenal temporal experience may correlate with high-dimensional pre-commit coherent dynamics (perception) rather than low-dimensional motor commit events (action)—consistent with consciousness tracking continuous constraint geometry operating beyond D_{crit}. However, detailed predictions about subjective experience remain beyond the scope of this work, which focuses on measurable computational and energetic signatures.

Free Energy Principle: Systems minimize prediction error [21]. Our framework adds: in timing-inaccessible regimes ($D_{\text{eff}} \gg D_{\text{crit}}$), prediction occurs through constraint geometry evolution inaccessible to external measurement. Active inference operates on collapsed outputs, not internal dynamics.

Developmental Bioelectricity: Levin's work demonstrates bioelectric gradients coordinate morphogenesis across scales. Ephaptic coupling in neural computation (Miller et al.) and bioelectric fields in development are manifestations of the same physical phenomenon—electric fields mediating information transfer without discrete synaptic/chemical transmission. Field-mediated computation may implement high-dimensional constraint geometry at cellular/tissue level, extending the framework beyond neural systems.

8.4 Implications for Artificial Intelligence

Current AI operates in fully measurable digital regimes with $D_{\text{eff}}/D_{\text{crit}} \sim 1$. Large language models achieve dimensional scaling ($D_{\text{eff}} \sim 10^4$ in embedding spaces) but through spatial parameters, not temporal structure, and with every activation observable.

Next generation: Analog neuromorphic substrates enabling timing-inaccessible computation could achieve:

- 10^2–$10^3\times$ power efficiency through sparse collisions
- Tractability for VAS-like problems via high-dimensional exploration
- Novel computational modes inaccessible to digital systems

Thermodynamic computing: Extropic's probabilistic hardware implements Energy-Based Models via stochastic CMOS circuits, reporting order-of-magnitude energy efficiency gains over GPUs on specific image generation benchmarks (e.g., Fashion-MNIST). By operating closer to continuous thermodynamic dynamics rather than fully discrete symbolic updates, such systems may approach intermediate $D_{\text{eff}}/D_{\text{crit}}$ regimes ($\sim$ 10–10^2), trading deterministic observability for computational efficiency.

The barrier: accepting that the most powerful computations may resist inspection. Intelligence emerges at boundaries of knowability.

9 Conclusion

We establish intelligence as the capacity to maintain high-dimensional coherent dynamics that evolve collision-free until sparse behavioral outputs. When $D_{\text{eff}} \gg D_{\text{crit}}$, systems solve computationally intractable problems through high-dimensional exploration in constraint geometry, paying thermodynamic costs only at discrete decision points rather than at every computational step.

Key results:

1. **Observable Dimensionality Bound:** $D_{\text{eff}} > D_{\text{crit}} = C_{\text{obs}}\tau_e/(\alpha h_\varepsilon^{\text{track}})$ defines regimes where timing information is physically inaccessible

2. **Dimensional Tracking Bound (Theorem 1):** Tracking systems with $D_{\text{target}} > D_{\text{obs}}$ requires collision cost $\sim k^{D_{\text{target}} - D_{\text{obs}}}$, proving intelligence must be high-dimensional to avoid exponential thermodynamic penalties

3. **Code Formation from Dimensional Mismatch (Theorem 2):** When high-D systems interact through low-D channels, stable symbolic codes must form at boundaries as thermodynamic necessity—explaining language, concepts, and communication protocols as inevitable consequences of dimensional compression

4. **VAS tractability:** Discrete reachability remains Ackermann-complete, but high-dimensional relaxations replace combinatorial path enumeration with dynamics governed by spectral gap/mixing time, explaining practical tractability in biological regimes

5. **Power scaling:** $P \propto H_{\text{reg}}$ explains the two–three orders of magnitude biological efficiency through sparse collisions

6. **Worked examples:** Cortex operates at $D_{\text{eff}}^{\text{MEG}}/D_{\text{crit}} \sim 10^2$ (conservative); substrate scales plausibly reach 10^3–10^4, enabling complex computation at ~ 20 W

7. **Testable predictions:** Coherence times, collapse signatures, dimensional scaling all measurable without accessing unmeasurable interior dynamics

The deep principle: high-dimensional coherent states encode constraints in geometric relationships that evolve collision-free, collapsing to discrete outputs only at decision boundaries. This enables both power efficiency (paying Landauer costs only at sparse behavioral events) and computational tractability (exploring exponentially large state spaces through continuous dynamics rather than discrete enumeration). A consequence of high-dimensional operation is that temporal microstructure becomes physically unmeasurable—measurement capacity cannot track individual mode trajectories when $D_{\text{eff}} \gg D_{\text{crit}}$.

Intelligence operates at the boundary between continuous and discrete computation—maintaining constraint geometry with $h_\varepsilon^{\text{prod}} = 0$ (no entropy production during exploration), collapsing to discrete outputs only when behavioral demands require it, achieving remarkable efficiency by decoupling internal dynamics from observable registration rate.

This suggests future computing architectures: systems that maintain high-dimensional coherent dynamics between sparse readout events, exploiting continuous constraint evolution to escape the power costs and local minima that trap fully discrete systems. The most powerful computations may be those that resist step-by-step inspection, operating through continuous geometry rather than discrete logic.

Acknowledgments

The author thanks Dr. Tara Eicher for encouragement to pursue publication of theoretical work, and the reviewers for constructive feedback improving clarity and empirical grounding.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The author has pending intellectual property related to this work. The author declares no competing financial interests or personal relationships that could have influenced this work.

Declaration of generative AI use

During preparation the author used Claude Code (Claude Sonnet 4.5, Anthropic) for literature review, mathematical formulation, coding, and editing. GPT-5 (OpenAI) and Grok-4 (xAI) provided critical feedback on mathematical rigor, notation consistency, and empirical grounding. All content was reviewed and revised by the author, who takes full responsibility for the published article.

Data and Code Availability

Simulation code (Python) demonstrating: (i) collision vs collision-free VAS dynamics, (ii) n-dimensional VAS scaling, (iii) code formation through adaptive dynamics, and (iv) Figure 1 generation, is provided as Supplementary Material. The code is freely available for adaptation to other VAS problems, coupling architectures, or learning schemes. All numerical results and figures reported in the main text are reproducible from the provided code (numpy 1.26+, scipy 1.11+, matplotlib 3.8+, scikit-learn 1.3+; random seed 42). No experimental data was generated for this theoretical study.

References

[1] Todd, I. (2025). The limits of falsifiability: Dimensionality, measurement thresholds, and the sub-Landauer domain in biological systems. *BioSystems*, 105608. doi:10.1016/j.biosystems.2025.105608

[2] Todd, I. (2025). Timing inaccessibility and the projection bound: Resolving Maxwell's demon for continuous biological substrates. *BioSystems*, 105632. doi:10.1016/j.biosystems.2025.105632

[3] Ashby, W.R. (1956). *An Introduction to Cybernetics*. Chapman & Hall, London. (Reprinted 1999, available at `http://pespmc1.vub.ac.be/books/IntroCyb.pdf`)

[4] Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5, 183–191. doi:10.1147/rd.53.0183

[5] Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

[6] Kuramoto, Y. (1984). *Chemical Oscillations, Waves, and Turbulence*. Springer, Berlin. doi:10.1007/978-3-642-69689-3

[7] Strogatz, S.H. (2000). From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D*, 143, 1–20. doi:10.1016/S0167-2789(00)00094-4

[8] Stetson, C., Fiesta, M.P., Eagleman, D.M. (2007). Does time really slow down during a frightening event? *PLoS ONE*, 2(12), e1295. doi:10.1371/journal.pone.0001295

[9] Epstein, I.R., & Pojman, J.A. (1998). *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*. Oxford University Press, Oxford.

[10] Czerwiński, W., Orlikowski, Ł. (2021). Reachability in Vector Addition Systems is Ackermann-complete. In: *62nd IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 1229–1240. doi:10.1109/FOCS52979.2021.00120

[11] Brubaker, B. (2023). An easy-sounding problem yields numbers too big for our universe. *Quanta Magazine*, December 4. https://www.quantamagazine.org/an-easy-sounding-problem-yields-numbers-too-big-for-our-universe-20231204/

[12] Miller, E.K., Brincat, S.L., Roy, J.E. (2024). Cognition is an emergent property. *Current Opinion in Behavioral Sciences*, 57, 101388. doi:10.1016/j.cobeha.2024.101388

[13] Pinotsis, D.A., Miller, E.K. (2023). In vivo ephaptic coupling allows memory network formation. *Cerebral Cortex*, 33(17), 9877–9895. doi:10.1093/cercor/bhad251

[14] Pinotsis, D.A., Shapiro, M., Miller, E.K. (2023). The Cytoelectric Coupling Hypothesis: How the Brain Creates Patterns of Information for Memory. *Progress in Neurobiology*, 227, 102476. doi:10.1016/j.pneurobio.2023.102476

[15] Siegel, M., Donner, T.H., Engel, A.K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, 13(2), 121–134. doi:10.1038/nrn3137

[16] Shine, J.M., Breakspear, M., Bell, P.T., Ehgoetz Martens, K.A., Shine, R., Koyejo, O., Sporns, O., Poldrack, R.A. (2019). Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nature Neuroscience*, 22(2), 289–296. doi:10.1038/s41593-018-0312-0

[17] Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford University Press, Oxford. doi:10.1093/acprof:oso/9780195301069.001.0001

[18] Levin, M. (2021). Bioelectric signaling: Reprogrammable circuits underlying embryogenesis, regeneration, and cancer. *Cell*, 184(8), 1971–1989. doi:10.1016/j.cell.2021.02.034

[19] Jelinčič, A., Lockwood, O., Garlapati, A., Verdon, G., McCourt, T. (2025). An efficient probabilistic hardware architecture for diffusion-like models. *arXiv preprint* arXiv:2510.23972 [cs.LG]. doi:10.48550/arXiv.2510.23972

[20] Kaplan, J., McCandlish, S., Henighan, T., Brown, T.B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., Amodei, D. (2020). Scaling laws for neural language models. *arXiv preprint* arXiv:2001.08361 [cs.LG]. arXiv:2001.08361

[21] Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. doi:10.1038/nrn2787

[22] Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5, 42. doi:10.1186/1471-2202-5-42

[23] Laughlin, S.B., de Ruyter van Steveninck, R.R., Anderson, J.C. (1998). The metabolic cost of neural information. *Nature Neuroscience*, 1, 36–41. doi:10.1038/236

[24] Attwell, D., & Laughlin, S.B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, 21(10), 1133–1145. doi:10.1097/00004647-200110000-00001

[25] Barbieri, M. (2018). What is code biology? *BioSystems*, 164, 1–10. doi:10.1016/j.biosystems.2017.10.005

[26] Bennett, C.H. (1973). Logical reversibility of computation. *IBM Journal of Research and Development*, 17(6), 525–532. doi:10.1147/rd.176.0525

[27] Bennett, C.H. (1982). The thermodynamics of computation—A review. *International Journal of Theoretical Physics*, 21(12), 905–940. doi:10.1007/BF02084158

[28] Bellman, R.E. (1961). *Adaptive Control Processes: A Guided Tour*. Princeton, NJ: Princeton University Press. doi:10.1515/9781400874668

[29] Bengio, Y., Courville, A., Vincent, P. (2013). Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(8), 1798–1828. doi:10.1109/TPAMI.2013.50

[30] Berisha, V., Krantsevich, C., Hahn, P.R., et al. (2021). Digital medicine and the curse of dimensionality. *npj Digit. Med.*, 4, 153. doi:10.1038/s41746-021-00521-5

[31] Castillo, I., Schmidt-Hieber, J., van der Vaart, A. (2015). Bayesian linear regression with sparse priors. *Ann. Stat.*, 43(5), 1986–2018. doi:10.1214/15-AOS1334

[32] Hvarfner, C., Stoll, D., Souza, A., Nardi, L. (2024). Vanilla Bayesian optimization performs great in high dimensions. *Proceedings of Machine Learning Research*, 235, 20793–20817. PMLR

[33] Johnson, V.E., Rossell, D. (2012). Bayesian model selection in high-dimensional settings. *J. Am. Stat. Assoc.*, 107(498), 649–660. doi:10.1080/01621459.2012.682536

[34] Gallego, J.A., Perich, M.G., Chowdhury, R.H., Solla, S.A., Miller, L.E. (2020). Long-term stability of cortical population dynamics underlying consistent behavior. *Nature Neuroscience*, 23(2), 260–270. doi:10.1038/s41593-019-0555-4

[35] Gao, P., Trautmann, E., Yu, B., Santhanam, G., Ryu, S., Shenoy, K., Ganguli, S. (2017). A theory of multineuronal dimensionality, dynamics and measurement. *Current Opinion in Neurobiology*, 46, 25–32. doi:10.1016/j.conb.2017.07.005

[36] Hipp, J.F., Hawellek, D.J., Corbetta, M., Siegel, M., Engel, A.K. (2012). Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience*, 15(6), 884–890. doi:10.1038/nn.3101

[37] NVIDIA Corporation. (2022). NVIDIA H100 Tensor Core GPU Architecture. White Paper. https://resources.nvidia.com/en-us-tensor-core (H100 PCIe: 350 W TDP; DGX H100 8-GPU system: ∼10 kW total)

[38] Palva, S., Palva, J.M. (2011). Functional roles of alpha-band phase synchronization in local and large-scale cortical networks. *Frontiers in Psychology*, 2, 204. doi:10.3389/fpsyg.2011.00204

[39] Recanatesi, S., Pereira-Obilinovic, U., Murakami, M., Mainen, Z., Mazzucato, L. (2024). Metastable attractors explain the variable timing of stable behavioral action sequences. *eLife*, 13, e100666. doi:10.7554/eLife.100666

[40] Vazza, F., Feletti, A. (2020). The Quantitative Comparison Between the Neuronal Network and the Cosmic Web. *Frontiers in Physics*, 8, 525731. doi:10.3389/fphy.2020.525731

[41] Bergoin, R., Torcini, A., Deco, G., Quoy, M., Zamora-López, G. (2025). Emergence and maintenance of modularity in neural networks with Hebbian and anti-Hebbian inhibitory STDP. *PLOS Computational Biology*, 21(3), e1012973. doi:10.1371/journal.pcbi.1012973

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

[43] Kauffman, S.A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.

[44] Berg, H.C., Brown, D.A. (1972). Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature*, 239, 500–504. doi:10.1038/239500a0

[45] Beggs, J.M., Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *Journal of Neuroscience*, 23(35), 11167-11177. doi:10.1523/JNEUROSCI.23-35-11167.2003

[46] Chialvo, D.R. (2010). Emergent complex neural dynamics. *Nature Physics*, 6(10), 744-750. doi:10.1038/nphys1803

[47] Levin, M., Keijzer, F., Lyon, P., Martinez, M. (2024). Collective intelligence: A unifying concept for integrating biology across scales and substrates. *Communications Biology*, 7, 378. doi:10.1038/s42003-024-06037-4

[48] Igamberdiev, A.U. (2021). Mathematics in biological reality: The emergence of natural computation in living systems. *BioSystems*, 204, 104388. doi:10.1016/j.biosystems.2021.104388

[49] Faisal, A.A., Selen, L.P.J., Wolpert, D.M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4), 292–303. doi:10.1038/nrn2258

[50] Leroux, J. (2013). Vector addition system reachability problem: A short self-contained proof. In *40th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL)*, pp. 307–316. doi:10.1145/2429069.2429110

Electronic Supplementary Material

All simulation code is provided as standalone Python files with full documentation and reproducibility guarantees (numpy $\geq$1.26, scipy $\geq$1.11, matplotlib $\geq$3.8, scikit-learn $\geq$1.3; random seed 42).

Figure 1 Generation Code

File: `code/figure1_discrete_vs_continuous.py`

Generates Figure 1 (all four panels) demonstrating high-dimensional discrete failure vs. continuous success. Runtime: $\sim$5 seconds.

VAS Collision Comparison Code

File: `code/vas_collision_comparison.py`

Implements the collision vs collision-free VAS comparison described in Section 3. Reproduces numerical results reported in main text for 2D example.

N-Dimensional VAS Scaling Code

File: `vas_scaling_simulation.py`

Generates Table 2 scaling data across dimensions $n \in \{2, 5, 10, 20, 30, 50, 100\}$. Independent transitions, optimal discrete algorithm, 20 trials per dimension. Outputs:

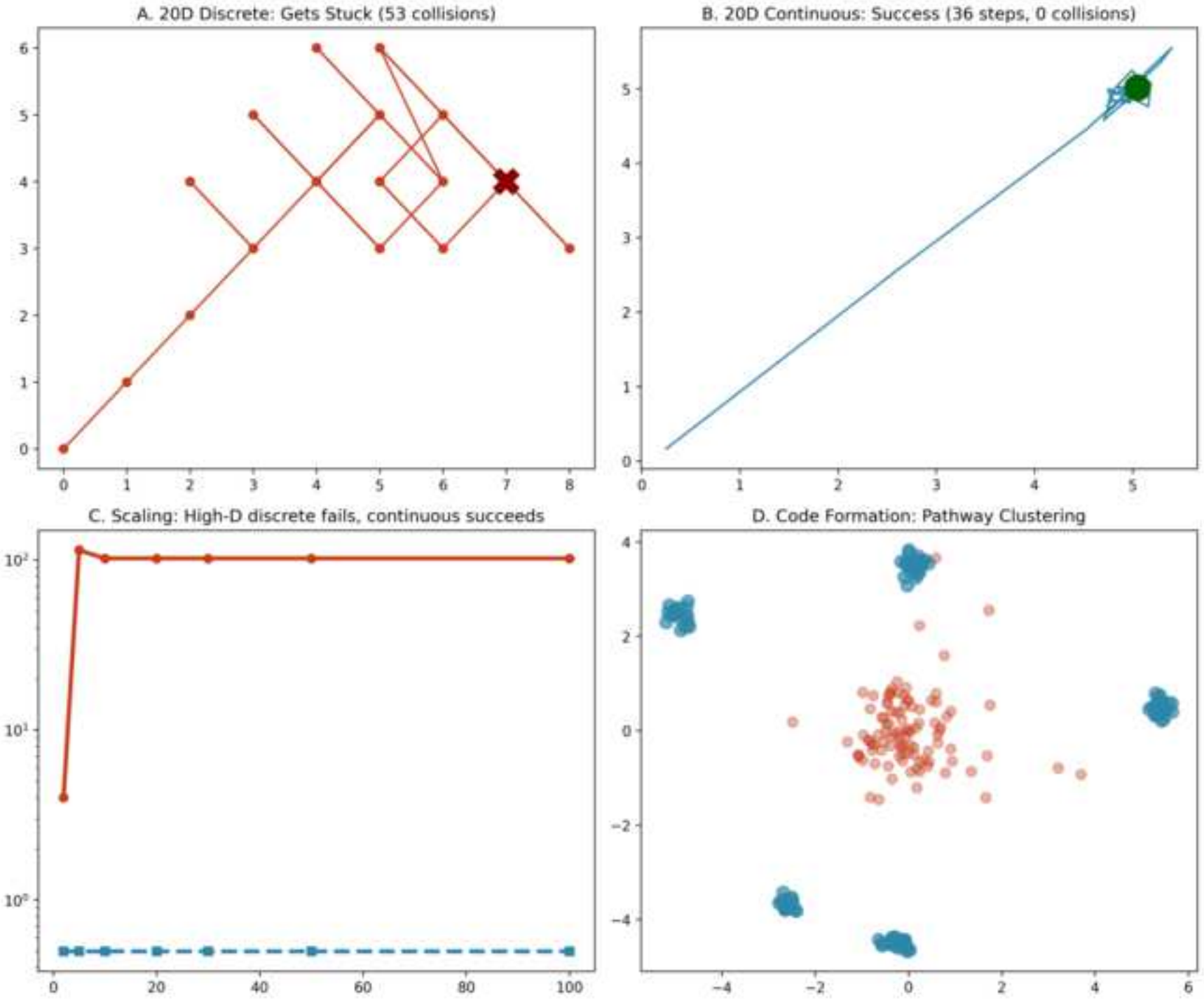
- `figures/vas_scaling.png` - Publication-quality scaling figure
- `figures/vas_scaling_data.npz` - Raw numerical data

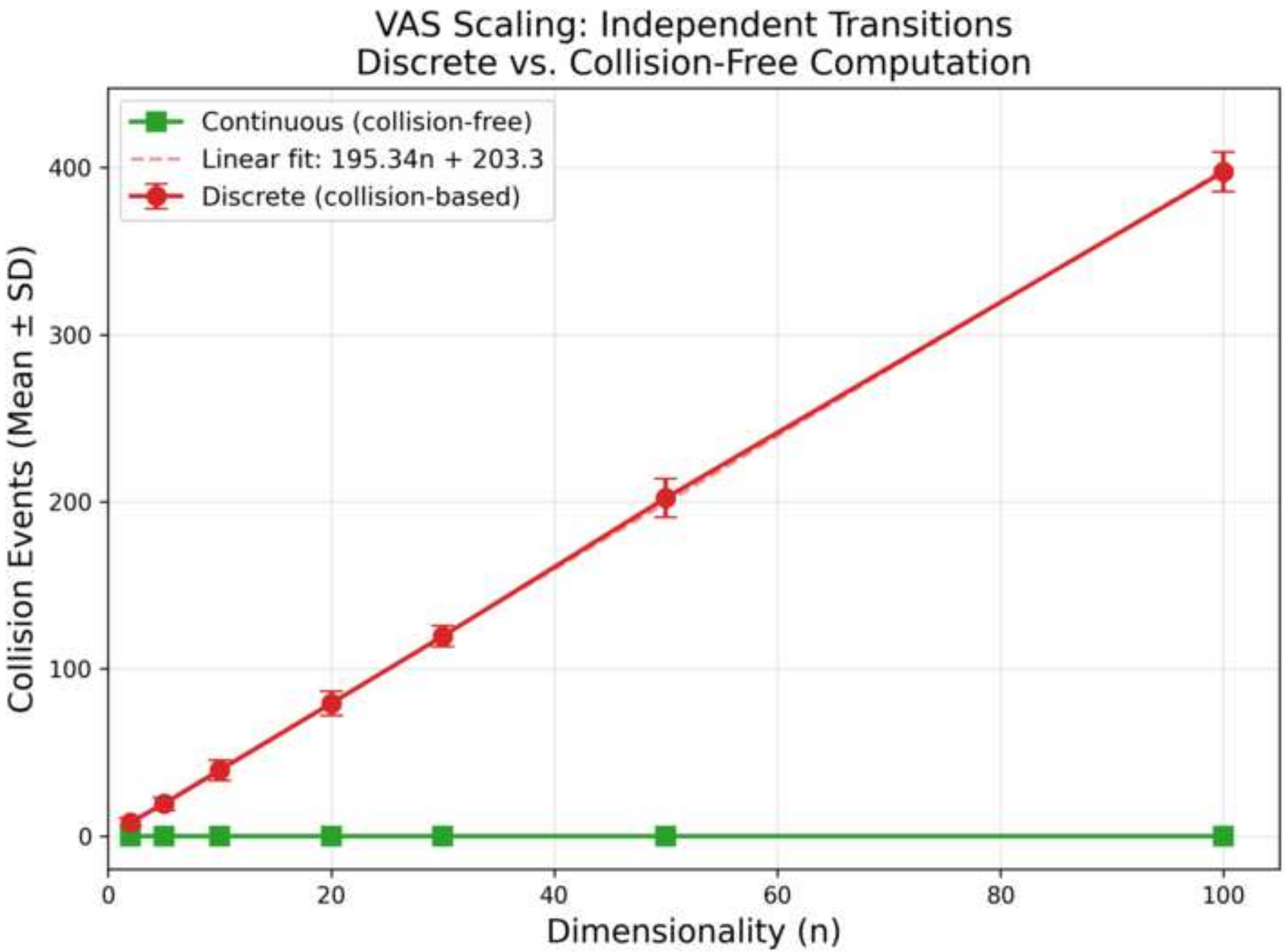
Runtime: $\sim$2 minutes for full n=2 to 100 sweep.

Code Formation Simulation Code

File: `code/code_formation_simulation.py`

Demonstrates spontaneous code formation in high-dimensional adaptive systems through Hebbian-like pathway strengthening. Compares adaptive pathway network (learns codes through weight adaptation) against discrete enumeration (no structural learning). Generates numerical results for pathway specialization and modular structure emergence.





Declaration of Competing Interests

Manuscript Title: Intelligence as High-Dimensional Coherence: The Observable Dimensionality Bound and Computational Tractability

Manuscript Number: BIOSYS-D-25-00880

The author declares the following potential competing interests:

- **Intellectual Property:** The author has pending intellectual property applications related to aspects of high-dimensional computation and dimensional collapse mechanisms described in this work. These applications do not alter the scientific validity of the findings reported herein, and all theoretical results, simulations, and empirical predictions are presented transparently for independent verification.

The author confirms that there are no financial interests, employment relationships, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations beyond those disclosed above, or grants/funding that could be perceived as affecting the objectivity of this work.

Author: Ian Todd

Affiliation: Sydney Medical School, University of Sydney

Date: 2025-11-16