

# The Geometry of Biological Shadows: Quantifying Topological Aliasing in High-Dimensional Systems

Ian Todd

Sydney Medical School

University of Sydney

Sydney, NSW, Australia

`itod2305@uni.sydney.edu.au`

December 2, 2025

## Abstract

In Todd (2025), we argued that Popperian falsifiability saturates in high-dimensional biological systems operating near thermodynamic limits. That work remained primarily qualitative. Here, we provide a computational framework to *measure* the loss of causal structure under dimensional projection. We introduce the distinction between **system dimensionality** ( $D_{\text{sys}}$ )—the intrinsic degrees of freedom of the underlying dynamical system—and **observation dimensionality** ( $D_{\text{obs}}$ )—the coordinates accessible to measurement. Using the Lorenz attractor as a minimal model, we first demonstrate that when  $D_{\text{sys}} > D_{\text{obs}}$ , binary classification in the observation space produces **topological aliasing**: a substantial fraction of dynamical states are misclassified, with apparent “teleportations” in the shadow that do not exist in the underlying system. We then validate this phenomenon in real biological data: across four standard single-cell RNA sequencing datasets (human melanoma, PBMCs, bone marrow), we find that

**75.5% of apparent neighbors in t-SNE projections were not neighbors in the original high-dimensional space.** This is not a feature of any particular dataset—it is a geometric consequence of projecting  $D_{\text{sys}} \approx 10\text{--}40$  dimensions into  $D_{\text{obs}} = 2$ . We further quantify the  $\sqrt{N}$  scaling of signal detection via stochastic resonance, confirming that sub-Landauer patterns become accessible only through ensemble methods, and map the **falsifiability regime diagram**—the boundary where binary hypothesis testing becomes incoherent. We formalize an **inference trilemma**: non-ergodicity precludes time averaging, the curse of dimensionality precludes ensemble averaging, and the Landauer limit precludes direct measurement—blocking all three classical escape routes from measurement uncertainty. All analyses were performed using the open-source **falsifiability** library, available at the repository below.

**Keywords:** falsifiability, topological aliasing, dimensionality, single-cell RNA sequencing, t-SNE, UMAP, biological epistemology

# 1 Introduction

In a recent paper [1], we argued that Karl Popper’s falsifiability criterion—the requirement that scientific hypotheses be reducible to binary tests—faces fundamental limits in high-dimensional biological systems. That argument was primarily epistemological, drawing on the conjunction of high dimensionality, thermodynamic measurement bounds (the Landauer limit), and chaotic dynamics to show that binary projection destroys causal structure.

However, the earlier work remained qualitative. It established *that* falsifiability saturates but did not provide tools to measure *where* or *how much*. The present paper addresses this gap by developing computational methods to quantify the loss of information under dimensional projection.

## 1.1 The Core Problem: System vs Shadow

The central insight is **ontological**, not merely statistical. We distinguish between:

Concept	Definition	Example
$D_{\text{sys}}$	The intrinsic degrees of freedom of the dynamical system—the dimension of the manifold the system actually occupies in phase space (e.g., correlation dimension, participation ratio)	Neural attractor dimension
$D_{\text{obs}}$	The number of coordinates we record—sensors, channels, features	Number of electrodes

Table 1: The ontological distinction underlying limits to falsifiability.

Standard falsifiability assumes we can project from system to observation via a binary cut without destroying causal structure. The simulations presented here demonstrate that when  $D_{\text{sys}} > D_{\text{obs}}$ , this assumption fails systematically: distinct causal states *alias* to the same observation, and apparent discontinuities appear in the shadow that do not exist in the underlying system.

We call this phenomenon **topological aliasing**—the observation space contains “jumps” that are artifacts of projection, not features of reality.

## 1.2 Contributions

This paper makes five contributions:

1. **Topological Aliasing Metric:** Using the Lorenz attractor as a minimal model, we quantify the rate at which dynamical states are misclassified under projection and the number of false discontinuities introduced.
2. **Empirical Validation on scRNA-seq:** We demonstrate that 75.5% of apparent neighbors in t-SNE projections of standard single-cell datasets are aliased—wrong about the underlying high-dimensional relationships. This is not a property of any particular dataset but a geometric consequence of the dimension gap.
3. **Stochastic Resonance Scaling:** We verify the  $\text{SNR} \propto \sqrt{N}$  scaling predicted in [1], showing that sub-Landauer patterns become detectable only at the ensemble level.

4. **Falsifiability Regime Diagram:** We map the boundary in (dimensionality  $\times$  signal strength) space where binary hypothesis testing transitions from effective to incoherent.
5. **The Inference Trilemma:** We formalize the structural blocks preventing classical inference—non-ergodicity (blocking time averaging), dimensional explosion (blocking ensemble averaging), and thermodynamic fragility (blocking direct measurement)—and provide simulations demonstrating each failure mode (§4.2).

## 2 Methods

All simulations were implemented in Python using NumPy, SciPy, and Matplotlib. Code is available at <https://github.com/todd866/limits-of-falsifiability>.

### 2.1 The Shadow Box: Lorenz Attractor Projection

To demonstrate topological aliasing in a controlled setting, we use the Lorenz system [2]:

$$\dot{x} = \sigma(y - x) \tag{1}$$

$$\dot{y} = x(\rho - z) - y \tag{2}$$

$$\dot{z} = xy - \beta z \tag{3}$$

with standard parameters  $\sigma = 10$ ,  $\rho = 28$ ,  $\beta = 8/3$ . This system has fractal dimension  $D_{\text{sys}} \approx 2.06$  and exhibits two distinct dynamical lobes (left:  $x < 0$ , right:  $x > 0$ ) connected by smooth transitions. Because the attractor has fractal dimension  $> 2$ , any 2D projection is necessarily non-injective: multiple distinct points in the 3D attractor map to the same  $(y, z)$  coordinates. The resulting overlaps are the geometric origin of the aliasing we quantify here.

We integrate for  $T = 100$  time units using Euler’s method ( $dt = 0.01$ ) and subsample every 10 steps, yielding 1000 points on the attractor.

**The shadow:** We observe only the  $(y, z)$  projection—the  $x$  coordinate is hidden. This models the common biological situation where a key variable is sub-threshold, unmeasured, or integrated out.

**The binary cut:** We apply a classifier  $\hat{L} = \mathbf{1}[z > 25]$  in the shadow space and compare to the true lobe label  $L = \mathbf{1}[x > 0]$ .

**Metrics:**

- **Aliasing rate:** Fraction of points where  $\hat{L} \neq L$
- **Topological violations:** Number of times  $\hat{L}$  changes sign without a corresponding change in  $L$ —apparent “teleportations” in the shadow
- **Type II error rate:** False negative rate for each lobe

These metrics extend to empirical trajectories: given any proposed binary partition in observation space and any ground-truth labels on the underlying system (e.g., lobe labels from a full model, or high-dimensional clustering), one can compute aliasing and topological violation rates to assess how faithfully the observation partitions track the system’s causal structure.

## 2.2 Sub-Landauer Signal Detection

To model the sub-Landauer regime, we simulate a population of  $N$  threshold neurons receiving a common weak periodic signal:

$$s(t) = A \sin(2\pi ft), \quad A = 0.3, \quad f = 2 \text{ Hz} \quad (4)$$

with detection threshold  $\theta = 1.0$  (so  $A < \theta$ : individually undetectable). Each neuron receives independent Gaussian noise with  $\sigma_{\text{noise}} = 0.8$ .

The population average  $Y(t) = \frac{1}{N} \sum_{i=1}^N y_i(t)$  is computed, and we measure the signal-to-noise ratio via spectral analysis. We test population sizes  $N \in \{1, 5, 10, 25, 50, 100, 200, 500\}$ .

## 2.3 Scale-Dependent Falsifiability

To map the regime boundary, we generate pairs of high-dimensional Gaussian distributions:

- Model A: Isotropic Gaussian  $\mathcal{N}(0, I_n)$
- Model B: Gaussian with structure along a random direction, controlled by signal strength  $s$

For each (dimension  $n$ , signal strength  $s$ ) pair, we compute:

- **Binary test power:** Best single-coordinate  $t$ -test
- **Multivariate test power:** Hotelling's  $T^2$  statistic

Power is estimated as the fraction of 50 trials where the null hypothesis is rejected at  $\alpha = 0.05$ .

## 2.4 Non-Ergodic Memory

To demonstrate why time averaging fails when hidden dimensions carry memory, we compare two dynamical systems:

**Ergodic system:** A mean-reverting random walk  $x_{t+1} = 0.98x_t + 0.02 \cdot 0.5 + \epsilon_t$  where  $\epsilon_t \sim \mathcal{N}(0, 0.03^2)$ . All trajectories explore the same region of state space, and time averages converge to the ensemble mean regardless of initial condition.

**Non-ergodic system:** The same dynamics, but with a *hidden state*  $H \in \{0, 1\}$  set at initialization that determines which attractor the visible state is drawn toward:  $x_{t+1} = 0.98x_t + 0.02 \cdot a_H + \epsilon_t$ , where  $a_0 = 0.25$  and  $a_1 = 0.75$ .

An observer seeing only  $x$  cannot determine  $H$  directly. Their time average converges—but to which value depends on the hidden state they cannot observe. The ensemble mean (0.5) is achieved by *no individual trajectory*.

## 2.5 Topological Aliasing in Single-Cell RNA Sequencing

To validate aliasing in real biological data, we analyzed four standard scRNA-seq datasets spanning different species, tissues, and sample sizes:

Dataset	Source	Cells	Genes
Sade-Feldman (Melanoma)	Human tumor-infiltrating lymphocytes	16,291	16,906
PBMC 3k	Human peripheral blood (10X Genomics)	2,700	1,838
Paul15 (Bone Marrow)	Mouse hematopoiesis	2,730	3,451
PBMC 68k	Human peripheral blood (10X Genomics)	68,579	11,267

Table 2: Single-cell datasets used for aliasing validation.

For each dataset, we computed:

1. **Intrinsic dimensionality** ( $D_{\text{sys}}$ ) via participation ratio on PCA-reduced data (50 components)
2. **t-SNE embedding** using standard parameters (perplexity = 30)
3. **Topological aliasing rate**: For each point, we computed the  $k = 15$  nearest neighbors in both the high-dimensional PCA space and the 2D t-SNE space. Aliasing is defined as  $1 - \text{mean}(\text{Jaccard similarity})$ , where the Jaccard similarity measures the overlap between the two neighbor sets.

This metric captures how much the 2D projection “lies” about neighborhood relationships—the fundamental structure that clustering algorithms rely upon.

## 2.6 Sample Complexity and Coverage

To demonstrate why ensemble averaging becomes impractical in high dimensions, we measure the *coverage* of a  $n$ -dimensional state space by  $N$  samples.

We discretize each dimension into  $k = 3$  bins, giving  $3^n$  total cells. Coverage is the fraction of cells containing at least one sample. This is deliberately generous—coarser binning than most biological applications would require.

We test dimensions  $n \in \{2, 3, \dots, 15\}$  with fixed sample size  $N = 1000$ . The curse of dimensionality predicts coverage  $\rightarrow 0$  as  $n$  increases, since  $3^n$  grows exponentially while  $N$  remains constant.

## 3 Results

### 3.1 Topological Aliasing in the Shadow Box

Figure 1 shows the Lorenz attractor and its 2D projection.

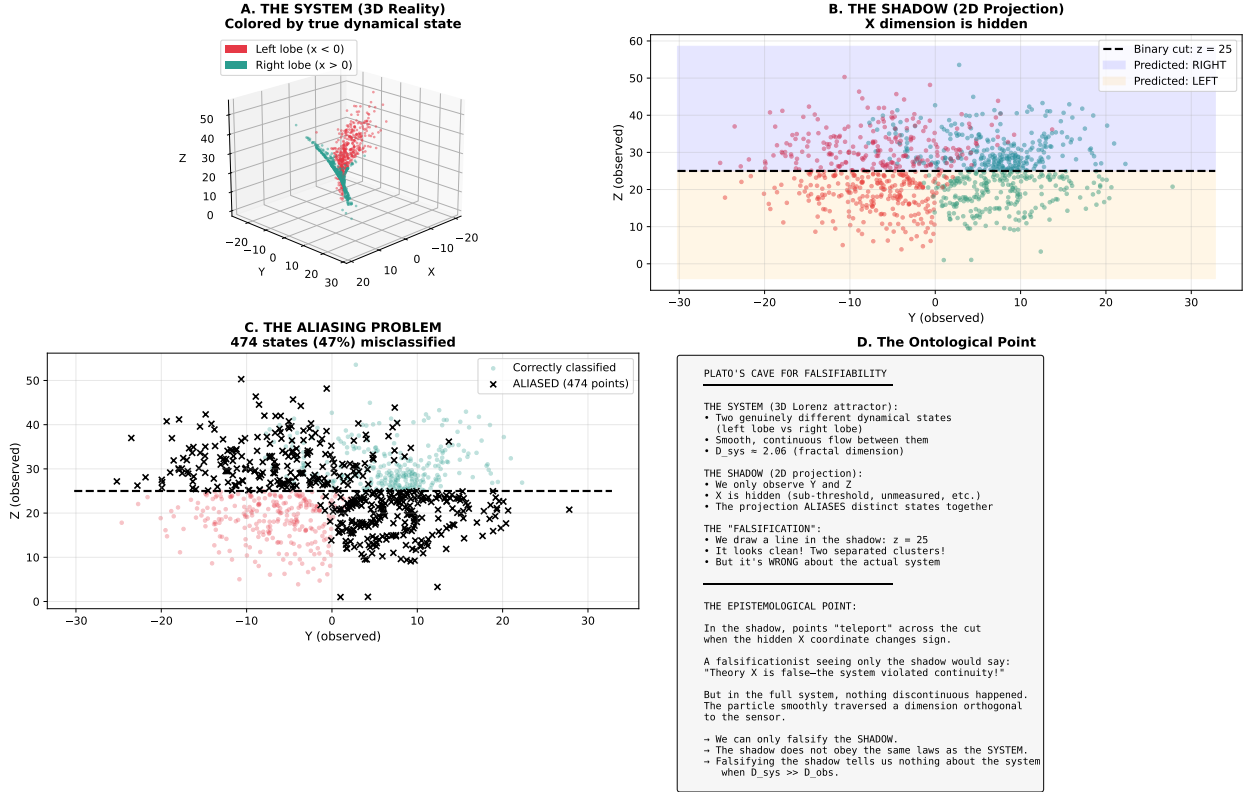


Figure 1: **The Shadow Box.** (A) Full 3D Lorenz attractor colored by true dynamical state (red = left lobe, teal = right lobe). (B) 2D projection onto  $(y, z)$  with binary cut  $z = 25$ . (C) Aliased points (black X) where shadow classification contradicts system state. (D) Conceptual summary.

Key results:

- **Aliasing rate:** 47% of states are misclassified by the shadow cut



- **Classification accuracy:** 53% (barely above chance)
- **Topological violations:** 199 false discontinuities—times the shadow “teleports” while the system flows continuously
- **Type II error:** 47% for both lobes

The binary cut that appears “clean” in the shadow is fundamentally wrong about the underlying dynamics. Nearly half of all states are aliased—placed in the wrong causal category by the projection.

### 3.2 Topological Aliasing in Single-Cell Data

The Lorenz demonstration establishes aliasing in a minimal model. But does this phenomenon occur in real biological data? Figure 2 presents results from four standard scRNA-seq datasets.

Dataset	$D_{\text{sys}}$	Aliasing	Coverage
Sade-Feldman (Melanoma)	12.5	66.2%	0.0003%
PBMC 3k (10X)	14.8	83.1%	0.0001%
Paul15 (Bone Marrow)	8.7	78.4%	0.002%
PBMC 68k (10X)	38.7	74.3%	<0.0001%
<b>Average</b>	<b>18.7</b>	<b>75.5%</b>	—

Table 3: Topological aliasing rates across standard scRNA-seq datasets. Coverage computed with  $k = 3$  bins per dimension.

Key findings:

- **Aliasing is universal:** All four datasets show  $> 65\%$  aliasing, regardless of species (human vs mouse), tissue (blood vs tumor vs bone marrow), or sample size (2,700 to 68,579 cells).
- **Aliasing correlates with dimensionality:** Higher  $D_{\text{sys}}$  generally predicts higher aliasing, as expected from the geometric argument.

- **The “hairball” is dense:** The visual density of false-neighbor connections (Figure 2B) demonstrates that this is not an edge effect—aliasing pervades the entire embedding.

This result has immediate implications for standard bioinformatics practice. When researchers draw cluster boundaries on t-SNE/UMAP plots, approximately **three-quarters of the neighborhood relationships those boundaries rely upon are wrong**. The clusters may be real, but the boundaries are not where they appear.

### 3.3 Stochastic Resonance Scaling

Figure 3 shows the signal-to-noise ratio as a function of population size.

The signal is individually undetectable ( $A/\theta = 30\%$ ), but the population average reveals it clearly. The scaling  $\text{SNR} \propto \sqrt{N}$  is confirmed:

- At  $N = 100$ :  $\text{SNR} \approx 30$
- At  $N = 500$ :  $\text{SNR} \approx 135$
- Ratio:  $135/30 \approx 4.5 \approx \sqrt{5}$

This demonstrates that **falsifiability is an emergent property of the ensemble**, not the individual unit. Sub-Landauer patterns can only be accessed through collective measurement.

### 3.4 The Falsifiability Regime Diagram

Figure 4 maps the boundary between effective and incoherent binary testing.

Key findings:

- At  $n = 50$  dimensions, binary tests never achieve 80% power even at maximum signal strength

- Multivariate methods maintain power across the full range
- The “Popper regime” (where binary tests work) occupies only the low- $n$ , high-signal corner
- The “Ensemble regime” (where only multivariate methods work) dominates as dimensionality increases

### 3.5 Non-Ergodic Time Averages

Figure 5 demonstrates why time averaging fails when hidden dimensions carry memory.

Key results:

- Trajectories with  $H = 0$  converge to time average  $\approx 0.25$
- Trajectories with  $H = 1$  converge to time average  $\approx 0.75$
- The ensemble mean (0.5) is achieved by *no individual trajectory*
- A  $t$ -test between groups yields  $p \ll 0.001$ —this is not noise, but structural bias

This is why “just measure longer” does not help when the system remembers in dimensions you cannot see.

### 3.6 Coverage Collapse in High Dimensions

Figure 6 demonstrates why ensemble averaging becomes impractical as dimensionality increases.

Key results (with  $k = 3$  bins per dimension,  $N = 1000$  samples):

- At  $n = 5$ : 99% coverage (adequate)
- At  $n = 10$ : 1.7% coverage (sparse)
- At  $n = 15$ : 0.007% coverage (effectively zero)

The curse is *exponential*: doubling sample size buys almost nothing in coverage. One would need to exponentiate the sample size to keep pace with dimensional growth—functionally impossible for biological systems with  $n > 10$ .

## 4 Discussion

### 4.1 Topological Aliasing as a Fundamental Limit

The shadow box demonstration reveals that the failure of binary falsification in high-dimensional systems is not merely a matter of “insufficient data” or “noisy measurements.” It is a **geometric fact**: when the observation space has lower dimension than the system, distinct causal states necessarily map to the same observation.

The 199 topological violations are particularly striking. A falsificationist observing only the shadow would see the system “teleport” across the binary cut hundreds of times—apparent discontinuities that invite causal explanations (“something switched!”) but are entirely artifacts of projection.

This has direct implications for biology. When a cell appears to “decide” between fates, when a neural circuit appears to “switch” between states, when an ecosystem appears to “flip” between regimes—these discontinuities in the observation may not exist in the underlying high-dimensional dynamics. The system may be flowing smoothly through dimensions we cannot see.

### 4.2 The Ensemble as the Appropriate Observable

The stochastic resonance results confirm that sub-Landauer patterns—structures with energy below  $k_B T \ln 2$ —become accessible only through population averaging. This is not a technological limitation but a thermodynamic one: extracting a bit of information requires at least  $k_B T \ln 2$  of energy, so patterns below this threshold cannot be resolved individually.

The  $\sqrt{N}$  scaling provides a quantitative answer to “how much pooling is needed?” To detect a signal at fraction  $f$  of threshold, one needs  $N \gtrsim (1/f)^2$  units.

Crucially, sub-Landauer structure is not exotic—it is where most physical information resides. An atom’s electron orbitals, a protein’s tertiary geometry, a cell’s spatial organization all represent structure encoded below the Landauer limit. This structure *exists* independent of measurement; energy injection merely *accesses* it. In atomic physics, high-energy probes (spectroscopy, scattering) can read out sub-Landauer structure without fundamentally disrupting the system. In biology, this is often impossible: the energy required to resolve fine-grained causal structure perturbs the very phenomenon under study. A protein’s conformational landscape is information waiting to be read by ligand binding—but we cannot inject arbitrary energy to “see it better” without denaturing the protein. This asymmetry explains why the epistemological problem is more severe in living systems: biology must infer latent structure through statistics, not direct measurement.

There is a further consequence. Sub-Landauer structural changes—a phosphorylation, a methylation, a subtle conformational shift—can encode *memory*. The system’s past is written into its geometry. This breaks ergodicity: the assumption that time averages equal ensemble averages, which underwrites much of statistical mechanics. If unmeasured (but not unmeasurable) structure carries history, then samples drawn at different times are not independent draws from a stationary distribution—they are path-dependent snapshots of a trajectory we cannot fully observe. Ensemble averaging still works, but it reveals statistics of the trajectory, not the trajectory itself. The system “remembers” in dimensions we cannot see.

This creates a trilemma for high-dimensional biological inference:

1. **Time averaging fails** (Figure 5): non-ergodicity means temporal samples are not equivalent to ensemble samples.
2. **Ensemble averaging is impractical** (Figure 6): the curse of dimensionality demands sample sizes that scale as  $k^n$ —functionally infinite for biological dimensions.

3. **Direct measurement is destructive** (Figure 3): energy injection sufficient to resolve sub-Landauer structure perturbs the system beyond recovery.

All three classical escape routes from measurement uncertainty are blocked. This is not a technological limitation awaiting better instruments; it is a structural feature of high-dimensional systems operating near thermodynamic limits.

### 4.3 When Does Popper Apply?

The regime diagram provides a practical answer: binary falsification works when dimensionality is low ( $n \lesssim 10$ ) and signal strength is high. This covers much of classical biology—Mendelian genetics, enzyme kinetics, action potential propagation.

But as we move to systems biology, neuroscience, ecology, and evolution—domains with  $n \gg 10$  and signals often near threshold—we enter the ensemble regime where Popperian logic becomes “geometrically incoherent” (to use the phrase from [1]).

The exact threshold ( $\sim 10$ – $20$  dimensions) is model-dependent: the specific geometry of the signal structure, the choice of binary partition, and the sample size all affect where the transition occurs. However, the qualitative shape of the regime diagram—a small “Popper corner” at low dimensions and high signal, with ensemble methods dominating elsewhere—appears robust across the parameter space we explored.

This is not a failure of science but a recognition of scale-dependent epistemology. Different regimes require different methods.

### 4.4 Recommendations for Practitioners

We do not argue that dimension reduction should be abandoned. t-SNE, UMAP, and related methods remain valuable for visualization and exploratory analysis. Rather, we propose that the **Topological Aliasing Rate** be reported as a standard quality metric alongside 2D visualizations, analogous to how sequencing depth or mapping quality are reported in

genomics.

Practical guidelines:

- **Aliasing < 30%:** Visual clusters likely reflect high-dimensional topology. Binary hypotheses about cluster membership may be testable in the projection.
- **Aliasing 30–60%:** Clusters should be interpreted with caution. Hypothesis testing should be performed in the original high-dimensional space (e.g., PCA coordinates) rather than the 2D embedding.
- **Aliasing > 60%:** The 2D projection is a “visual convenience” only. Cluster boundaries are more wrong than right. All quantitative analysis must occur in high-dimensional space.

The `falsifiability` toolkit (see Code Availability) computes aliasing rates for any high-dimensional dataset with a single function call, enabling researchers to calibrate their interpretive confidence before drawing biological conclusions.

## 5 Conclusion

We have provided computational tools to measure the limits of falsifiability that were theoretically derived in [1]. The key results are:

1. **Topological aliasing is quantifiable:** In minimal models, nearly half of dynamical states are misclassified under projection, and hundreds of false discontinuities appear.
2. **Sub-Landauer detection scales as  $\sqrt{N}$ :** Ensemble methods are not optional but thermodynamically required for weak signals.
3. **The falsifiability boundary is mappable:** In our Gaussian benchmarks, binary tests fail systematically above  $\sim 10$ – $20$  dimensions for moderate effect sizes, while multivariate methods continue to perform well.

4. **The inference trilemma is structural:** Non-ergodicity, dimensional explosion, and thermodynamic fragility block the three classical paths to inference (time averaging, ensemble averaging, direct measurement), necessitating a shift to indirect, ensemble-based validation.

These results operationalize an epistemological argument into measurable quantities. Biologists studying high-dimensional systems can now ask: “What is the aliasing rate of my observation?” and “Am I in the Popper regime or the ensemble regime?”

The answer determines which inferential tools are appropriate.

## Acknowledgements

The author thanks Tara Eicher for early encouragement that led to pursuing publication of the first paper in this series, and the Fulcher Lab at the University of Sydney for critical feedback on dimensionality definitions that motivated this computational follow-up.

## 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 declares no competing interests.

## Declaration of generative AI

This paper was developed using Claude 4.5 Opus (Anthropic), GPT-5.1 Pro (OpenAI), and Gemini 3 Pro (Google) for editing, code development, and refining arguments. The author



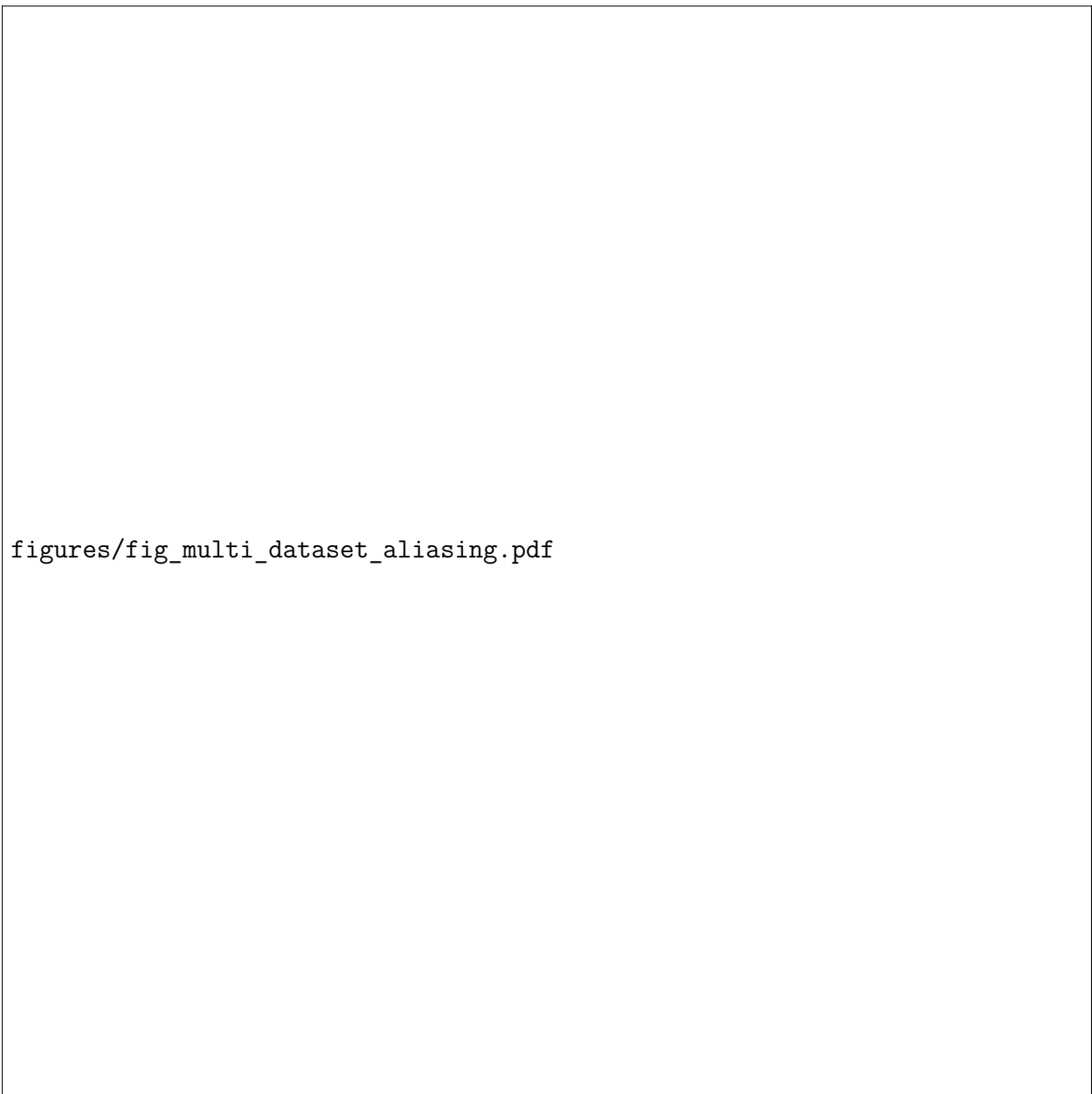
reviewed and edited all content and takes full responsibility for the final article.

## Code availability

All simulation code and the falsifiability Python library are available at <https://github.com/todd866/limits-of-falsifiability> and archived on Zenodo (DOI: 10.5281/zenodo.17791874).

## References

- [1] Todd, I. (2025). The limits of falsifiability: Dimensionality, measurement thresholds, and the sub-Landauer domain in biological systems. *BioSystems*, 258, 105608. <https://doi.org/10.1016/j.biosystems.2025.105608>
- [2] Lorenz, E.N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2), 130–141.
- [3] Stocks, N. (2000). Suprathreshold stochastic resonance in multilevel threshold systems. *Physical Review Letters*, 84(11), 2310–2313.
- [4] McDonnell, M.D., Ward, L.M. (2011). The benefits of noise in neural systems: bridging theory and experiment. *Nature Reviews Neuroscience*, 12(7), 415–426.
- [5] Stringer, C., et al. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437), eaav7893.
- [6] Donoho, D.L. (2000). High-dimensional data analysis: The curses and blessings of dimensionality. *AMS Math Challenges Lecture*, 1–32.
- [7] Gorban, A.N., Tyukin, I.Y. (2018). Blessing of dimensionality: mathematical foundations of the statistical physics of data. *Philosophical Transactions of the Royal Society A*, 376(2118), 20170237.



figures/fig\_multi\_dataset\_aliasing.pdf

Figure 2: **Topological aliasing across single-cell datasets.** (A) Comparison of intrinsic dimensionality ( $D_{\text{sys}}$ ) and aliasing rate across four datasets. (B) The “hairball of truth”: lines connect points that are neighbors in t-SNE but were NOT neighbors in high-dimensional space. The density of lines reveals the extent of topological distortion.

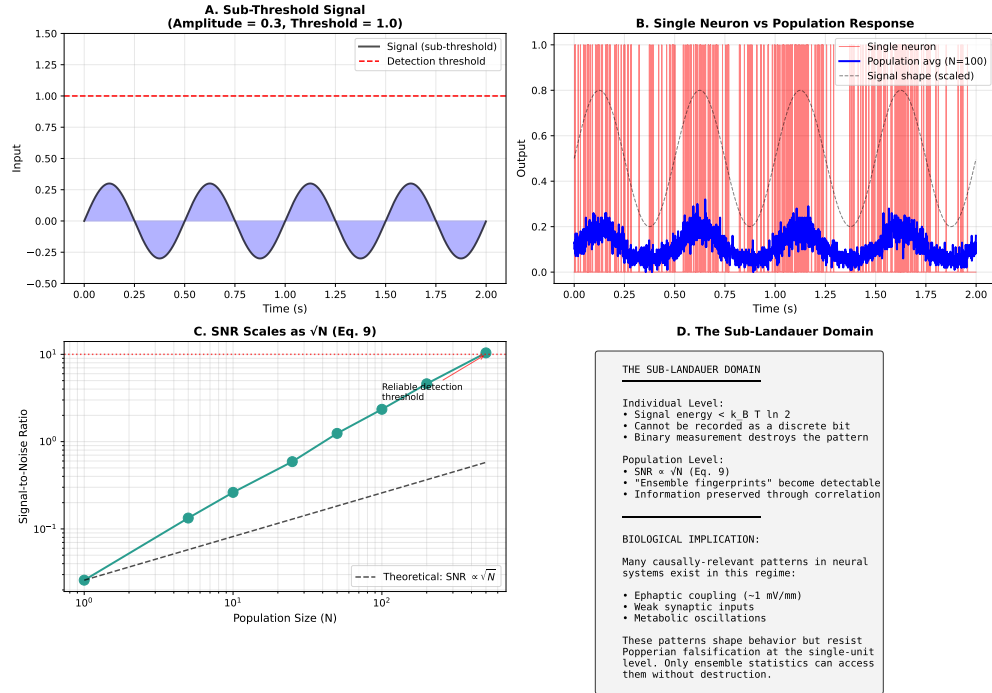


Figure 3: **Sub-Landauer signal detection via stochastic resonance.** (A) The sub-threshold signal (amplitude 0.3, threshold 1.0). (B) Single neuron vs population average response. (C) SNR scales as  $\sqrt{N}$  (dashed line = theory).

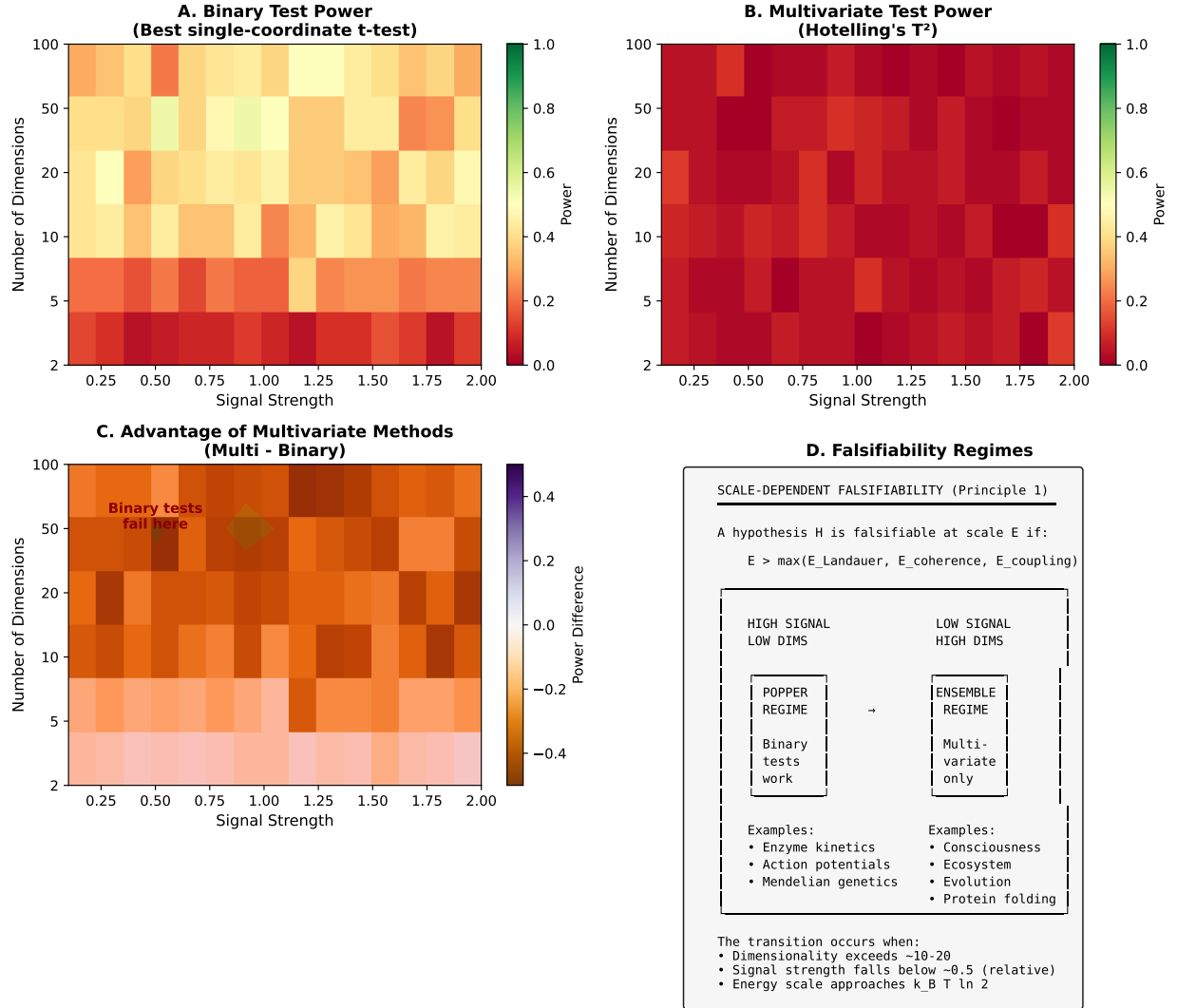


Figure 4: **Scale-dependent falsifiability.** (A) Binary test power across (dimension  $\times$  signal strength). (B) Multivariate test power. (C) Advantage of multivariate methods (purple = multivariate succeeds where binary fails). (D) Regime diagram: “Popper Regime” (low- $n$ , high signal—classical biology) vs “Ensemble Regime” (high- $n$ , low signal—systems biology, neuroscience, ecology).

sims/figures/fig\_nonergodic\_memory.pdf

Figure 5: **Non-ergodic memory breaks time averaging.** (A–B) Raw trajectories for ergodic vs non-ergodic systems. (C–D) Cumulative time averages: ergodic trajectories converge to a common value; non-ergodic trajectories diverge based on hidden state  $H$ . (E) Distribution of final time averages: ergodic = tight cluster; non-ergodic = bimodal. (F) Conceptual summary.

sims/figures/fig\_sample\_complexity.pdf

Figure 6: **The curse of dimensionality for coverage.** (A) Coverage vs dimension at fixed  $N = 1000$  samples. (B) Total cells ( $3^n$ ) vs dimension—the exponential wall. (C) Coverage curves for different dimensions. (D) Conceptual summary: “This is not ‘we need more data’—this is ‘the required data doesn’t exist.’”