The Code-Constraint Problem in Biological Systems: How Low-Dimensional Interfaces Shape High-Dimensional Dynamics

Ian Todd

Sydney Medical School

University of Sydney

Sydney, NSW, Australia

itod2305@uni.sydney.edu.au

December 20, 2025

Abstract

A recurring problem spans biophysics, systems biology, and neuroscience: high-dimensional systems (protein ensembles, gene networks, neural populations) are characterized through low-dimensional descriptions (order parameters, principal components, expression markers), yet the relationship between interface and underlying dynamics remains poorly understood. When is dimensional reduction faithful compression versus systematic distortion? Why do some low-dimensional codes work while others fail?

This Perspective proposes a unifying interpretation: low-dimensional interfaces between coupled systems function not only as information channels, but as *stabilizing constraints* that shape downstream dynamics. We synthesize evidence from dynamical systems theory, statistical mechanics, and information theory to argue that the recurring interpretive problems arise when this constraint role is overlooked.

Using coupled oscillator simulations as a minimal physical exemplar, we demonstrate a characteristic signature of constraint: bandwidth-limited coupling induces systematic changes in responding-system complexity—often a collapse—while maintaining bounded tracking. Operationally: if reducing interface bandwidth causes complexity collapse in the responding system while tracking error remains bounded, the interface functions as a constraint rather than merely losing information. Critically, this requires structured projections that capture coherent collective variables; random projections of the same dimensionality produce the opposite effect (whitening, not collapse). A systematic parameter sweep reveals a sharp boundary between constrained and unconstrained regimes, with finite-size scaling showing the effect strengthens with system size. The same signature appears in gene regulatory network dynamics, confirming generality beyond oscillatory systems.

We connect this mechanism to protein folding, single-cell manifolds, morphogenetic fields, and neural coding, and offer practical criteria for researchers: how to choose reaction coordinates, interpret manifold structure, design multi-scale interfaces, and distinguish constraint from representation. The framework suggests that biological codes often function as dimensional valves enabling persistent organization, a role complementary to—but distinct from—information transmission.

Keywords: Dimensional reduction, coarse-graining, biological codes, constraint dynamics, order parameters, model identifiability

1 A Common Problem Across Fields

Across biophysics, systems biology, and neuroscience, a recurring tension appears: the systems we study operate in high-dimensional state spaces, yet we characterize them through low-dimensional descriptions. This dimensional mismatch is not merely a practical limitation—it generates systematic interpretive problems that recur across domains.

In protein biophysics, conformational ensembles explore thousands of degrees of free-

dom, yet function is characterized by a handful of order parameters (Bryngelson et al., 1995; Dill and Chan, 1997). These are typically treated as faithful summaries of folding dynamics. However, reaction coordinates that appear to govern kinetics may hide multiple parallel pathways (Pande et al., 2010), and parameter estimation is generically "sloppy"—many parameter combinations produce indistinguishable low-dimensional outputs (Gutenkunst et al., 2007; Transtrum et al., 2015; Machta et al., 2013). When order-parameter descriptions are reliable remains actively debated (Best and Hummer, 2005; Noé and Clementi, 2017).

In single-cell genomics, RNA sequencing reveals that cell states occupy low-dimensional manifolds embedded in ∼20,000-dimensional gene expression space (Trapnell, 2015; Weinreb et al., 2018). These manifolds are often interpreted as biologically meaningful structure (McInnes et al., 2018; Kobak and Berens, 2019). Yet dimensionality reduction can create spurious clusters, collapse distinct states, and distort neighborhood relationships (Chari and Pachter, 2023; Cooley et al., 2019; Lähnemann et al., 2020). Pseudotime analyses rely on manifold structure that may reflect technical artifacts rather than biological reality.

In developmental biology, tissues coordinate through field-like gradients—morphogen concentrations, bioelectric potentials, mechanical stresses (Levin, 2021; Turing, 1952; Kicheva and Briscoe, 2015). Classically, these gradients are understood as positional information that cells "read" to determine fate (Wolpert, 1969). Yet these fields may possess rich dynamics at molecular and ionic scales, while individual cells respond through bandwidth-limited interfaces (ion channels, receptors). The mismatch between field complexity and cellular readout is structural, not merely practical.

In neuroscience, neural populations fire in high-dimensional pattern spaces, yet must coordinate across distant brain regions through apparent low-dimensional codes (Gallego et al., 2017; Sadtler et al., 2014; Cunningham and Yu, 2014). These codes are often interpreted as representations optimized for downstream processing. Yet order parameters that appear to govern dynamics may be shadows of higher-dimensional processes (Flack, 2017; Balduzzi and Tononi, 2008), and learning requires operating within existing neural

manifolds—a constraint whose mechanistic origin remains unclear.

These problems share a common structure: in each case, a high-dimensional driving system couples to a high-dimensional responder through a low-dimensional interface whose structure and bandwidth determine downstream dynamics. Yet the relationship between interface properties and system behavior is poorly understood. When can we trust a low-dimensional model? When is dimensional reduction a faithful compression versus a systematic distortion?

1.1 This Review: Codes as Constraints

This review proposes a unifying interpretation of these phenomena, rather than a new biological mechanism. We synthesize insights from dynamical systems theory, statistical mechanics, and information theory to argue that the recurring problems above arise from a common source: treating dimensional reduction purely as compression, overlooking its role in shaping dynamics.

The core reframing is that low-dimensional interfaces between coupled systems function not only as information channels, but as *stabilizing constraints* that regulate downstream dynamics.

Consider any biological interface: a cell membrane transducing environmental signals, a morphogenetic gradient coordinating tissue development, a neural code coupling sensory input to motor output. In each case, a high-dimensional "driving" system couples to a high-dimensional "responding" system through a low-dimensional bottleneck. The question is: what does this bottleneck *do*?

The standard information-theoretic framing asks how much information about the driving system's state is preserved or lost. We propose a different question: how does the bottleneck *shape the dynamics* of the responding system?

Using a minimal model of coupled oscillator lattices, we demonstrate that bandwidth-limited coupling produces a distinctive signature:

1. **Complexity collapse**: The responding system's effective dimensionality decreases systematically with code bandwidth.

2. **Bounded tracking**: Alignment between systems remains stable despite information loss.

3. **Structure dependence**: The effect requires projections onto coherent collective variables; random projections fail.

This pattern—complexity collapse with bounded tracking—is qualitatively different from information loss (which would produce tracking failure) or simple filtering (which would not systematically reduce effective dimensionality). We argue it represents a general mechanism by which biological coding structures enable persistent organization.

Positioning relative to existing theory: The framework relates to Haken's synergetics, where fast variables become "slaved" to slow order parameters—but extends this to coupled systems where the order parameter constrains a *separate* responding system. It differs from the information bottleneck principle, which optimizes compression for predictive relevance; we instead measure how compression constrains downstream dynamics, regardless of predictive content. It complements Markov blanket formalism by adding bandwidth as a quantitative knob: not just whether states are separated, but how many degrees of freedom can cross the boundary.

To be explicit about the novel claim: we are not proposing a better compression or a better predictor. We propose a *distinct dynamical regime diagnostic*: complexity collapse in the responding system while tracking error remains bounded, where "bounded" means alignment with the *coarse-grained reconstruction*, not full state reconstruction. This regime appears only for structured projections (coherent collective variables) rather than arbitrary low-dimensional reductions.

Table 1 summarizes the dimensional mismatch problem across four biological domains, with predicted constraint signatures and example measurements.

Table 1: Dimensional mismatch across biological domains and predicted constraint signatures.

Domain	**High-D system**	**Low-D interface**	**Known failure**	**Constraint test**	**Key refs**
Protein folding	Conformational ensemble	Reaction coordinate	Sloppiness; pathway degeneracy	B-factor collapse upon binding	Gutenkunst et al. (2007); Best and Hummer (2005)
Single-cell	Gene expression (∼20k)	Marker programs	Spurious clusters; trajectory artifacts	Manifold D_{eff} vs. coupling	Chari and Pachter (2023); Trapnell (2015)
Morphogenesis	Tissue microstates	Voltage gradient	Positional ambiguity	Bioelectric bandwidth manipulation	Levin (2021); Pietak and Levin (2017)
Neural coding	Local circuits	Population manifold	Learning constrained to manifold	Phase coherence vs. coupling	Sadtler et al. (2014); Gallego et al. (2017)

2 Biophysical Contexts

Before presenting a formal model, we review the dimensional mismatch problem as it appears across biological scales. Each domain has developed its own approaches and encountered characteristic difficulties. Our aim is to show that these difficulties share a common structure amenable to a unified interpretation.

2.1 Protein Conformational Ensembles

The protein folding problem offers perhaps the clearest example of dimensional mismatch in molecular biophysics. A typical protein explores a conformational space with thousands of degrees of freedom, yet researchers characterize function using a handful of order parameters: radius of gyration, fraction of native contacts, or principal components of the backbone motion (Frauenfelder et al., 1991; Henzler-Wildman and Kern, 2007).

The "energy landscape" perspective, developed by Wolynes, Onuchic, and colleagues (Bryngelson et al., 1995; Onuchic et al., 1997), frames folding as descent through a funnel-shaped landscape. This picture has been enormously productive, yet it is explicitly a low-dimensional projection of an astronomically high-dimensional reality. The funnel is not the landscape itself but a coarse-grained representation that may obscure parallel pathways, kinetic traps, and conformational heterogeneity (Dill and Chan, 1997).

Advances in molecular dynamics have made it possible to simulate folding at atomic resolution over millisecond timescales (Shaw et al., 2010; Klepeis et al., 2009). Yet even with such data, the choice of reaction coordinates remains contentious. Best and Hummer (2005) showed that apparent folding pathways depend sensitively on which collective variables are used to project the trajectory. Markov State Models attempt to address this by discretizing conformational space into metastable states (Pande et al., 2010; Husic and Pande, 2018), but the results depend on how states are defined—a choice that is itself a form of dimensional reduction.

A persistent finding is "sloppiness": many different parameter combinations in molecular models produce indistinguishable low-dimensional outputs (Gutenkunst et al., 2007; Transtrum et al., 2015). This is not merely a fitting problem but reflects a fundamental asymmetry between the high-dimensional microscopic reality and the low-dimensional observables we can measure. Machta et al. (2013) argue that sloppiness is generic in complex systems and reflects parameter space compression—the microscopic details that don't matter for macroscopic predictions.

The debate here is whether low-dimensional descriptions are *summaries* of high-dimensional dynamics (faithful compression) or *constraints* that shape what dynamics are possible (active reduction). The information-theoretic view treats reaction coordinates as optimal compressions for prediction. An alternative view—the one we develop below—treats them as interfaces that constrain downstream dynamics regardless of their predictive power.

Translating to the constraint framework: In protein-ligand binding, the ligand's confor-

mational dynamics in solution (high-dimensional) couple to the binding pocket (also high-dimensional) through a low-dimensional interface: the subset of conformations that can physically dock. If binding functions as constraint rather than information transfer, we predict: (1) reduced conformational entropy of the bound pocket relative to apo form, (2) stable binding geometry despite ligand fluctuations, and (3) this effect failing for decoy ligands that match steric volume but not the structured collective variables of true binding. Crystallographic B-factors could test this: cognate binding should induce greater complexity collapse than non-cognate binding of similar affinity.

2.2 Single-Cell State Manifolds

Single-cell RNA sequencing has transformed our ability to characterize cellular heterogeneity, revealing that cell states occupy low-dimensional manifolds embedded in ∼20,000-dimensional gene expression space (Trapnell, 2015). This observation—that high-dimensional data lies on low-dimensional structure—has become foundational for trajectory inference, pseudotime analysis, and RNA velocity methods (Weinreb et al., 2018; Schiebinger et al., 2019).

The standard dimensionality reduction tools—t-SNE (Van der Maaten and Hinton, 2008) and UMAP (McInnes et al., 2018)—have become ubiquitous in single-cell analysis. Yet their reliability has been sharply questioned. Chari and Pachter (2023) demonstrated that these methods can create spurious clusters, collapse distinct states, and systematically distort neighborhood relationships. Cooley et al. (2019) showed that different reduction methods applied to the same data can yield contradictory biological conclusions. The compression ratio (20,000 dimensions to 2–3) is so extreme that information loss is inevitable; the question is whether the loss is systematic or random.

A deeper problem identified by Lähnemann et al. (2020) is that the field lacks principled criteria for choosing among dimensionality reduction methods. PHATE (Moon et al., 2019) preserves different structural features than UMAP, which preserves different features than

diffusion maps. Each method makes implicit assumptions about which aspects of high-dimensional structure matter, but these assumptions are rarely made explicit or justified biologically.

The debate in the field centers on interpretation: do low-dimensional manifolds reflect intrinsic regulatory simplicity, or are they artifacts of the measurement and analysis pipeline? Some argue that developmental trajectories are genuinely low-dimensional because gene regulatory networks channel cells through stereotyped paths (Moris et al., 2016; Huang et al., 2005). Others argue that apparent low-dimensionality may reflect technical factors: dropout, batch effects, or the particular genes included in analysis (Lähnemann et al., 2020).

Our framework suggests a different question entirely: rather than asking whether manifolds *represent* cell states accurately, ask whether they *constrain* cell fate transitions. A gene regulatory motif that restricts accessible trajectories (like a developmental "valve") functions differently from one that merely correlates with cell state. The former shapes dynamics; the latter reflects them. This reframing shifts attention from visualization fidelity to functional consequence.

2.3 Morphogenetic Fields and Bioelectric Gradients

Developmental biology has long grappled with how spatial patterns emerge from locally-interacting cells. Two paradigms have dominated: Wolpert's "positional information" model, where cells read concentration gradients to determine their position (Wolpert, 1969), and Turing's reaction-diffusion model, where patterns emerge from local chemical dynamics (Turing, 1952). Recent work has sought to integrate these perspectives (Green and Sharpe, 2015; Kicheva and Briscoe, 2015), but both share a common feature: high-dimensional molecular dynamics are projected onto low-dimensional "morphogenetic fields" that guide tissue organization.

A growing body of work emphasizes bioelectric gradients as a parallel signaling system. Levin (2021) has demonstrated that manipulating membrane voltage patterns can reprogram

tissue-scale outcomes—inducing eye formation in non-eye locations, or converting tumor cells to normal phenotypes. The key insight is not that bioelectric fields are simple, but that cellular responses are constrained to low-dimensional projections of field dynamics. Cells "listen" through ion channels that act as bandpass filters, extracting only certain features of the voltage landscape (Levin, 2014; Adams and Levin, 2013).

Computational models of bioelectric signaling reveal a striking asymmetry: the ionic dynamics underlying membrane potential involve dozens of channel types, second messengers, and feedback loops (Pietak and Levin, 2017), yet the signal that propagates between cells is a single voltage value. This is dimensional reduction by physical necessity—the membrane itself acts as an interface that compresses high-dimensional intracellular dynamics into a low-dimensional intercellular signal.

The debate in developmental biology mirrors debates in protein folding and single-cell analysis: are morphogenetic fields faithful representations of tissue state, or are they constraints that shape what developmental trajectories are accessible? The positional information view treats fields as signals carrying instructions. The constraint view—developed below—treats them as interfaces that restrict the dynamical complexity of responding cells.

2.4 Neural Population Codes

Neuroscience faces the dimensional mismatch problem acutely. A cortical column contains ∼100,000 neurons, each firing in patterns that could encode vast amounts of information. Yet behavior is low-dimensional: we reach to a target, recognize a face, decide left or right. Somewhere between neural firing and behavior, dimensionality collapses.

The discovery that neural population activity is confined to low-dimensional manifolds (Churchland et al., 2012; Cunningham and Yu, 2014) has reshaped the field. Gallego et al. (2017) showed that motor cortex activity during reaching lies on manifolds of dimension 10–20, far below the number of recorded neurons. Sadtler et al. (2014) demonstrated that learning new behaviors requires either operating within existing manifolds or slowly expand-

ing them—a finding that suggests manifolds are not just descriptions but constraints on what the brain can do.

Two competing interpretations have emerged. The *representational* view holds that low-dimensional manifolds reflect the low-dimensional structure of task variables: reaching has few degrees of freedom, so neural activity encoding it should too (Georgopoulos et al., 1986; Pouget et al., 2000). The *mechanistic* view holds that manifolds reflect constraints imposed by network architecture and connectivity, independent of task structure (Sadtler et al., 2014).

A separate literature emphasizes oscillatory coordination. Buzsáki and Draguhn (2004) argued that brain rhythms provide a "syntactic" structure for neural communication. Fries (2015) proposed that gamma-band synchronization gates information flow between regions—the "communication through coherence" hypothesis. Varela et al. (2001) and Engel et al. (2001) developed the idea that large-scale integration requires phase synchronization across distributed populations.

These oscillatory frameworks implicitly invoke dimensional reduction: the phase of a rhythm is a low-dimensional variable that summarizes the collective state of millions of neurons. Miller et al. (2018) proposed that low-frequency oscillations (alpha, beta) coordinate activity across regions while high-frequency activity (gamma) carries local information. This frequency hierarchy is effectively a bandwidth hierarchy—slow oscillations couple over long distances because they average over local fluctuations.

The constraint interpretation suggests a synthesis: neural manifolds may reflect not the dimensionality of task variables, but the bandwidth of inter-regional coupling. Local circuits can support high-dimensional dynamics, but only low-dimensional projections can propagate across the coupling bottleneck to influence distant regions. Learning a new behavior isn't just finding a new representation—it's finding a representation that can propagate through bandwidth-limited interfaces.

3 A Minimal Physical Exemplar

To make the constraint interpretation concrete, we demonstrate its signature in a minimal model: two coupled lattices of phase oscillators where interaction is restricted to a bandwidth-limited projection. This is not intended as a model of any specific biological system, but as a *didactic exemplar*—a physically realizable system that exhibits the constraint mechanism in a form amenable to precise measurement. The same qualitative signature emerges in gene regulatory network dynamics (Section 5.2), confirming that the mechanism is not specific to oscillatory dynamics.

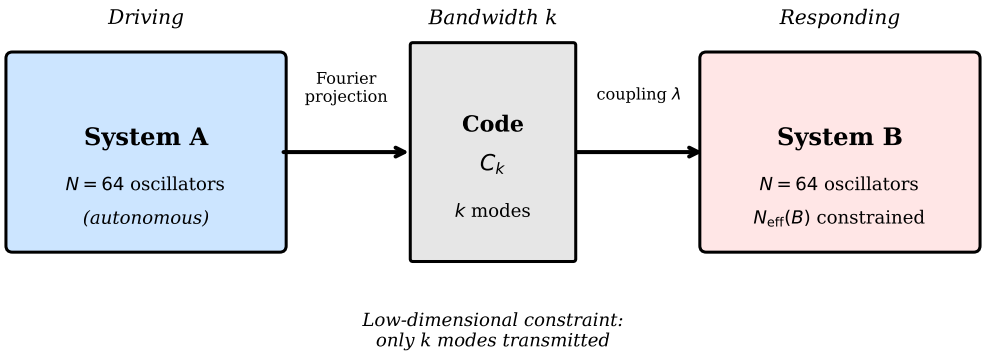

Figure 1: Model architecture. System A (driving) couples to System B (responding) only through a bandwidth-limited code C_k consisting of k Fourier modes. The code constrains B's dynamics without transmitting full state information.

3.1 Model Architecture

Consider two systems, each a one-dimensional lattice of N locally coupled phase oscillators. System A (the "driving" system) evolves autonomously:

$$\dot{\theta}_i^A = \omega_i + K \sum_{j \in \mathcal{N}(i)} \sin(\theta_j^A - \theta_i^A) + \eta_i^A(t), \tag{1}$$

where ω_i are intrinsic frequencies, K is coupling strength, $\mathcal{N}(i)$ denotes nearest neighbors, and $\eta_i^A(t)$ is Gaussian noise.

System B (the "responding" system) couples to A only through a low-dimensional code:

$$\dot{\theta}_i^B = \omega_i + K \sum_{j \in \mathcal{N}(i)} \sin(\theta_j^B - \theta_i^B) + \lambda \sin(\hat{\theta}_i^A - \theta_i^B) + \eta_i^B(t), \tag{2}$$

where $\hat{\theta}^A$ is a bandwidth-limited reconstruction of A's state.

The code C_k consists of the first k Fourier modes of A's phase field:

$$C_m = \frac{1}{N} \sum_{i=1}^{N} e^{i\theta_i^A} e^{-i2\pi mi/N}, \quad m = 0, \ldots, k. \tag{3}$$

The bandwidth-limited reconstruction is obtained by inverting with only the retained modes:

$$\hat{\theta}_i^A = \arg \left(\sum_{m=0}^{k} C_m e^{i2\pi mi/N} \right). \tag{4}$$

We define bandwidth k as the maximum retained mode index (including DC), so the code contains $k + 1$ complex coefficients.

This architecture (Figure 1) captures the essential structure: two high-dimensional systems coupled through a low-dimensional interface. The bandwidth k is a tunable parameter controlling code dimensionality.

3.2 Quantifying Complexity and Tracking

We measure *effective dimensionality* via spectral entropy of the Fourier amplitude spectrum. Specifically, the amplitude $|C_m|$ of each Fourier mode is computed, normalized to form a probability distribution $p_m = |C_m| / \sum_m |C_m|$, and the Shannon entropy $H = -\sum_m p_m \log p_m$ is calculated. The effective dimensionality is then $N_{\text{eff}} = e^H$, which gives the *effective support size*—the number of modes that contribute substantially to the dynamics, weighted by their relative amplitudes. This is not a topological dimension but a complexity measure analogous to the participation ratio in localization theory. The same qualitative results obtain using alternative metrics—PCA participation ratio (covariance-based), spatial gradient energy, and

Kuramoto order parameter—none of which depend on the Fourier basis (Supplementary Figure S2). This confirms that complexity collapse is not an artifact of measuring complexity in the same basis as the code.

Tracking quality is measured as mean circular distance between systems: $\Delta = N^{-1} \sum_i |\sin((\theta_i^A - \theta_i^B)/2)|$.

4 Demonstrating the Constraint Signature

The preceding sections reviewed the dimensional mismatch problem across fields. We now demonstrate concretely what a "constraint signature" looks like, using the minimal model described above. The goal is not to model any specific biological system but to establish measurable criteria that distinguish constraint from information loss.

4.1 Complexity Collapse Under Bandwidth Reduction

Figure 2 shows the central phenomenon. As code bandwidth k decreases:

- System B's effective dimensionality decreases systematically ($N_{\text{eff}} = 16.7 \to 10.9$)
- System A's complexity remains constant ($N_{\text{eff}} \approx 17.0$)
- The relationship is approximately monotonic

The bottleneck constrains only the responding system; the driving system's dynamics are unaffected by how it is observed. This asymmetry is the hallmark of constraint rather than mutual information loss.

4.2 Bounded Tracking Despite Complexity Collapse

Mismatch between systems increases only modestly as bandwidth decreases (Figure 3): full-state mismatch $\Delta(\theta^A, \theta^B) = 0.38 \to 0.50$, a $\sim$32% change, compared to $\sim$35% reduction in $N_{\text{eff}}(B)$.

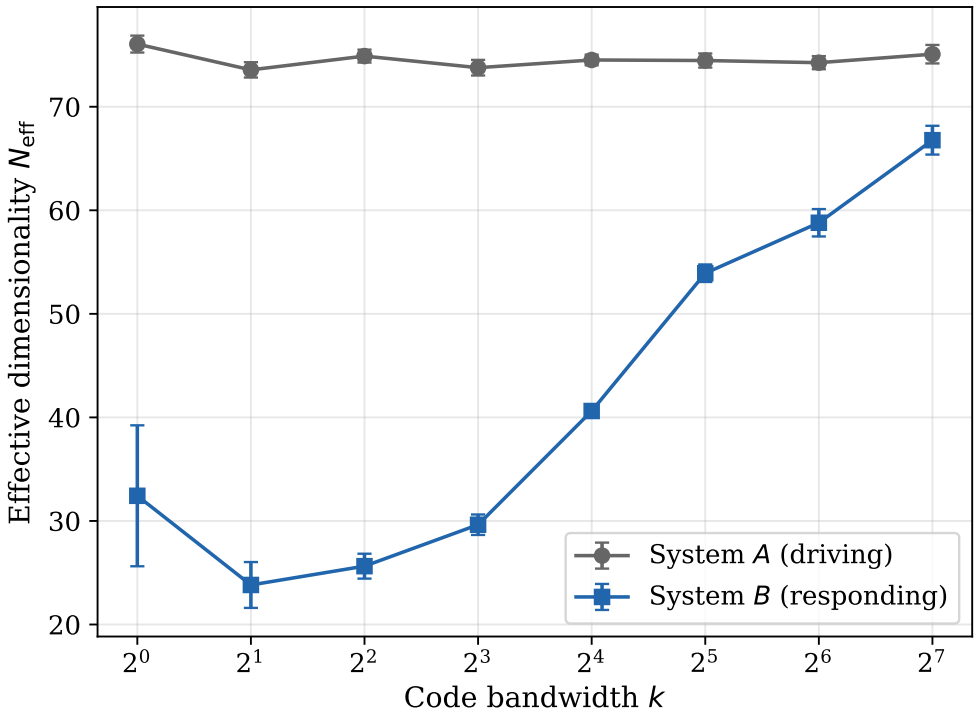

Figure 2: Effective dimensionality as a function of code bandwidth. System A (gray) maintains constant complexity; system B (blue) exhibits systematic collapse. Error bars: standard error over 15 trials.

This asymmetry is crucial. Figure 3 shows that mismatch to the reconstructed code, $\Delta(\hat{\theta}^A, \theta^B)$, remains substantially lower than full-state mismatch across all bandwidths. This is the operational sense in which tracking remains "bounded": B tracks the low-dimensional constraint even when full state information is missing. If the code merely lost information, we would expect commensurate tracking failure. Instead, B tracks *the code* faithfully—it aligns with A's coarse-grained representation. The code constrains B's behavioral complexity without compromising alignment.

4.3 Structure Matters: Random Mode Selection Fails

A critical control: we repeated experiments using random Fourier mode subsets instead of the lowest k modes. To ensure a fair comparison, random projections retain *exactly the same number of modes* as structured projections: both conditions retain $k+1$ Fourier coefficients,

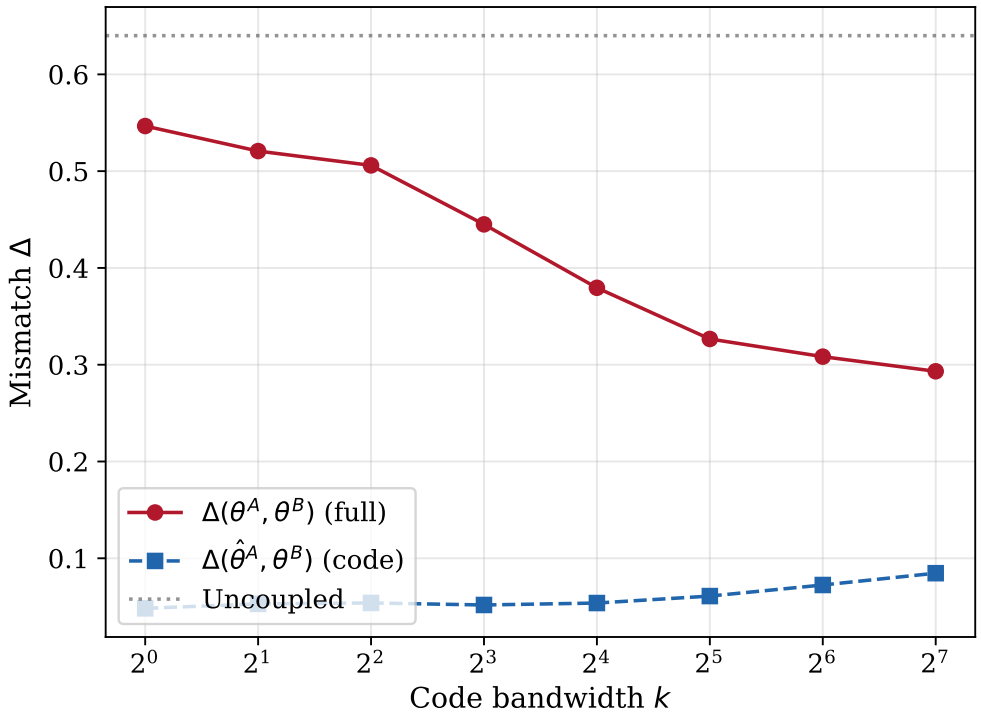

Figure 3: Tracking quality as a function of code bandwidth. Full-state mismatch $\Delta(\theta^A, \theta^B)$ (red solid) shows how well B tracks A's true state; code mismatch $\Delta(\hat{\theta}^A, \theta^B)$ (blue dashed) shows how well B tracks the coarse-grained reconstruction. The uncoupled control (gray dotted) provides a baseline. Code mismatch remains substantially lower than full-state mismatch across all bandwidths, confirming that B tracks the constraint, not the full state.

always including DC ($m = 0$). The structured projection uses modes $\{0, 1, \ldots, k\}$; the random projection uses mode 0 plus k modes selected uniformly at random from positive frequencies (with conjugate symmetry enforced for real-valued reconstruction). The only difference is *which* non-DC modes are retained—contiguous low frequencies versus scattered random frequencies.

Results differ strikingly (Figure 4):

- Random projections produce *no* complexity collapse ($N_{\text{eff}} \approx 15$–$17$)
- Mismatch is substantially higher
- At low k, random codes increase B's complexity above A's

The constraining effect is not about dimensionality per se, but about *structure*. Low-

frequency modes capture spatially coherent patterns; random modes mix signal with noise. Biological codes, we argue, must similarly capture coherent collective variables rather than arbitrary dimensional reductions.

Notably, at low k the random projection actually *increases* B's complexity above baseline ($N_{\text{eff}} \approx 18$ vs. 17 for the uncoupled system). This occurs because sparse random projections destroy local spatial correlations in the driving signal, effectively acting as a whitening filter that forces the responding lattice to desynchronize locally while still responding to the global coupling. The random code injects high-frequency structure rather than removing it.

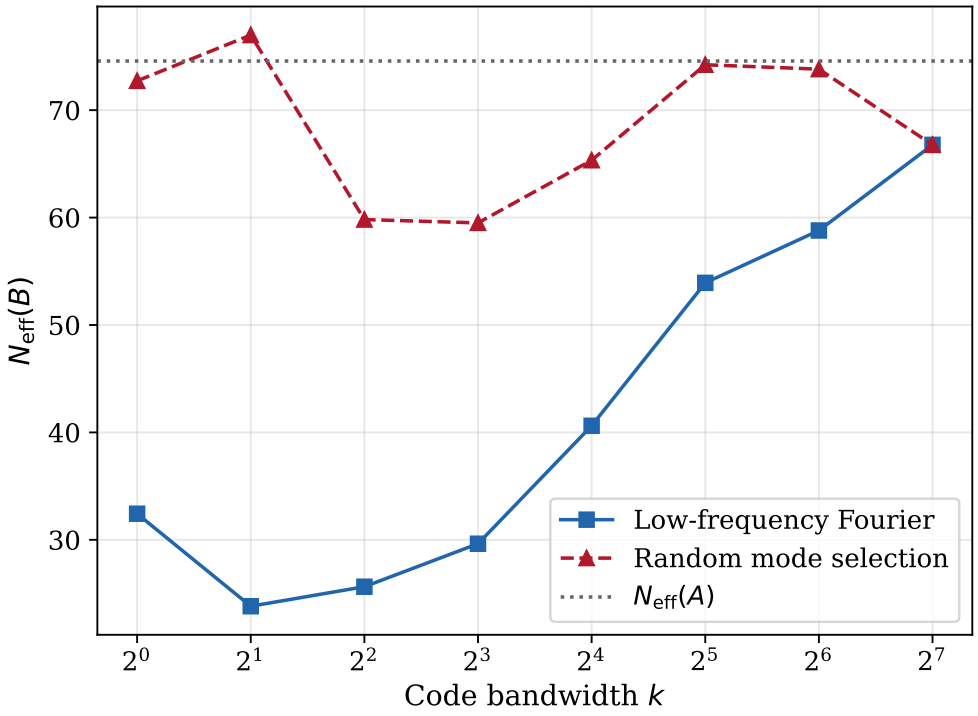

Figure 4: Low-frequency Fourier coupling (blue) vs. random mode selection (red). Random projections of the same dimensionality fail to induce complexity collapse.

Robustness: The qualitative pattern—complexity collapse for structured projections, not for random ones—persists across the parameter ranges tested: noise amplitude $\sigma \in [0.1, 0.5]$, coupling strength $\lambda \in [0.25, 4.0]$, and lattice sizes $N \in [32, 128]$. The effect is not an artifact of specific parameter choices. We expect other structured bases (Laplacian eigenmodes, wavelets, principal components of the driving system) to produce similar constraint signa-

tures, though the specific relationship between k and N_{eff} will depend on the basis.

5 Generality Across Scales and Dynamics

A key question for any proposed unifying framework is whether it applies beyond its original demonstration system. If the constraint signature is specific to phase oscillators, its relevance to protein folding or gene regulation would be limited. We therefore tested whether the same signature appears in systems with qualitatively different dynamics.

5.1 Large-Scale Kuramoto Systems

We repeated the main analysis at $N = 512$ oscillators (Supplementary Figure S1), confirming that the complexity collapse signature persists at larger system sizes:

- At $N = 512$: $N_{\text{eff}}(B)$ ranges from ≈ 102 at $k = 4$ to ≈ 143 at $k = 128$
- Random projections continue to show no collapse, with $N_{\text{eff}}(B)$ remaining near the uncoupled baseline

Further scaling to $N \in \{1024, 2048\}$ is presented in Section 5.6, where we characterize the full phase structure of the transition. The constraint signature is not a finite-size artifact; it persists and indeed *sharpens* as system dimensionality increases.

5.2 Gene Regulatory Network Dynamics

The biological contexts reviewed in Section 2 include gene regulatory networks, where transcription factor concentrations serve as low-dimensional interfaces between cellular subsystems. To test whether the constraint signature appears in this qualitatively different dynamics, we implemented a GRN model with Hill-function activation:

$$\dot{x}_i = -\frac{x_i}{\tau} + \sigma\left(\sum_j W_{ij} x_j\right) + \eta_i(t), \tag{5}$$

where $\sigma(\cdot)$ is a sigmoid activation function, W is a sparse random connectivity matrix, and τ is the decay timescale.

The "code" in this context represents transcription factor readout: a low-dimensional projection (via DCT) of one network's expression state constrains the other network's dynamics.

Results (Figure 5, $N = 256$ genes) show a complementary signature to the Kuramoto system: $N_{\text{eff}}(A)$ remains constant at ≈ 193, while $N_{\text{eff}}(B)$ collapses from 192 at $k = 1$ to 121 at $k = 32$ (37% reduction). Mismatch increases modestly (0.22 to 0.37) as bandwidth increases.

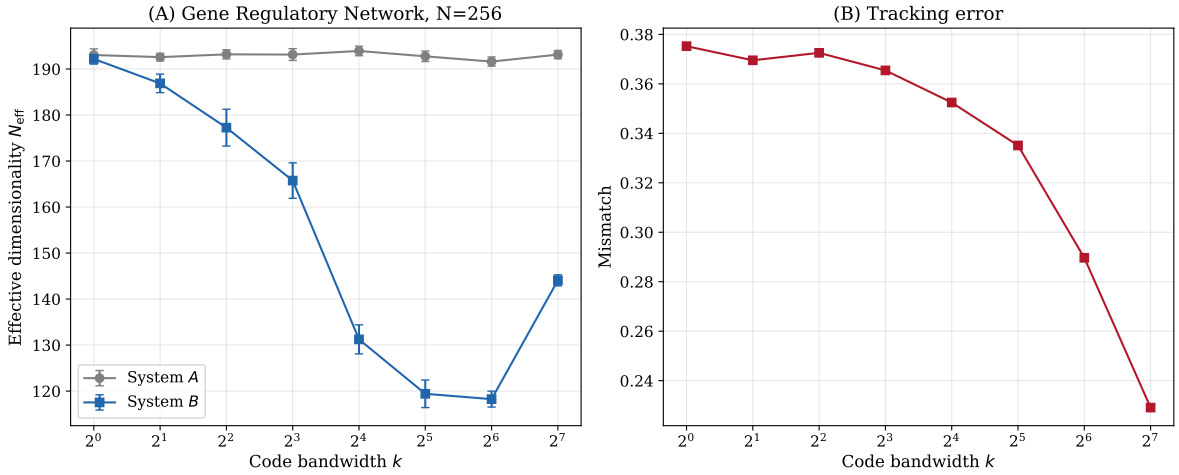

Figure 5: Complexity collapse in gene regulatory network dynamics ($N = 256$). (A) Effective dimensionality as a function of code bandwidth. System A (gray) maintains constant complexity while System B (blue) shows 37% reduction as k increases—a trend *opposite* to oscillators (Fig. 2) due to tracking saturation rather than mode suppression. (B) Tracking error (expression difference) increases at high bandwidth. Both systems demonstrate that low-dimensional interfaces constrain responding systems, but the mechanism depends on architecture: oscillators collapse at low k when codes suppress coherent modes; GRNs collapse at high k when high-variance inputs drive saturating activation functions.

The GRN results demonstrate that complexity collapse under bandwidth-limited coupling is not specific to phase oscillators, but the mechanism differs. In oscillator networks, narrow codes suppress fast modes, causing collapse at low bandwidth. In switch-like GRNs with sigmoid activation, high-bandwidth inputs drive nodes into the saturation regions of the

activation function, effectively freezing degrees of freedom and causing collapse when the code exceeds the system's tracking capacity.

Unifying the two regimes: Rather than viewing these as contradictory results, we propose they represent two canonical *constraint regimes*:

- **Filtering-limited regime** (Kuramoto-like): the code removes fast/incoherent modes, forcing the responder onto a low-dimensional attractor. Collapse occurs at low k.
- **Tracking-limited regime** (GRN-like): the responder cannot follow high-bandwidth drive due to intrinsic saturation or timescale separation. Collapse occurs at high k.

The common signature is that *responder complexity is systematically regulated by interface structure*, with the direction of collapse depending on where the limiting nonlinearity resides—in the code (filtering) or in the responder (tracking). This architectural dependence is itself a prediction: systems with strong intrinsic oscillations should show filtering-limited collapse, while systems with saturating response functions should show tracking-limited collapse.

5.3 Alternative Complexity Metrics

A potential concern is that measuring complexity via spectral entropy of Fourier amplitudes is tautological, since the code is also Fourier-based. We therefore computed three additional metrics that do not depend on the Fourier basis (Supplementary Figure S2):

- **PCA participation ratio**: effective dimensionality from covariance eigenvalues, $D_{\text{eff}} = (\sum \lambda)^2 / \sum \lambda^2$
- **Spatial gradient energy**: mean squared phase gradient $\langle |\nabla \theta|^2 \rangle$, measuring spatial smoothness
- **Kuramoto order parameter**: global synchrony $R = |\langle e^{i\theta} \rangle|$

All three metrics show the same qualitative pattern: System B's organization changes systematically with code bandwidth (lower gradient energy and higher order at low k), while System A remains constant. This confirms that complexity collapse is a genuine dynamical phenomenon, not an artifact of the measurement basis.

5.4 Coupling Strength Robustness

We verified that the qualitative pattern persists across a range of coupling strengths $\lambda \in [0.25, 4.0]$ (Supplementary Figure S3). Stronger coupling produces greater complexity collapse and tighter tracking, but the signature—complexity collapse with bounded mismatch—appears across the tested range.

5.5 Fractal Noise Robustness

A potential limitation of the main simulations is the use of white noise forcing. Real biological systems exhibit long-range temporal correlations characteristic of $1/f^{\beta}$ dynamics—from neural oscillations to gene expression fluctuations. We tested whether the constraint signature persists under fractal driving by replacing white noise ($\beta = 0$) with pink ($\beta = 1$) and brown ($\beta = 2$) noise (Supplementary Figure S4).

The constraint mechanism proves robust to noise structure. While fractal noise increases the baseline complexity of the driving system (more correlated modes contribute to the dynamics), the responding system still shows systematic constraint at low bandwidth. Complexity collapse at $k = 1$ is 65% under white noise, 46% under pink noise, and 43% under brown noise. The constraint is modulated but not eliminated by realistic temporal correlations, supporting applicability to biological systems with $1/f$-like dynamics.

5.6 Parameter Structure and Finite-Size Scaling

Having established robustness across coupling strengths and alternative metrics, we now examine the full (k, λ) parameter space to characterize the structure of the constraint regime. This analysis addresses a fundamental question: is complexity collapse a gradual effect or does it exhibit sharp transition-like behavior?

We performed a systematic sweep over code bandwidth $k \in [1, 256]$ and coupling strength $\lambda \in [0, 8]$ at $N = 1024$. The resulting phase diagram (Figure 6, panel A) reveals clear structure: a *constrained regime* at low k and high λ where $N_{\text{eff}}(B)$ drops to ~ 50 (compared to baseline ~ 270), and an *unconstrained regime* at high k or low λ where complexity remains near baseline. The transition between regimes is surprisingly sharp—not a gradual continuum but a distinct boundary in parameter space.

To test whether this sharpness reflects phase-transition-like behavior, we performed finite-size scaling analysis across $N \in \{256, 512, 1024, 2048\}$ at fixed $\lambda = 2.0$ (Figure 6, panel B). The results are striking:

- At each system size, the normalized complexity $N_{\text{eff}}(B)/N_{\text{eff}}^{\text{baseline}}$ follows a similar curve when plotted against k/N
- The collapse deepens with increasing N: at $k/N \approx 0.005$, normalized complexity drops from ~ 0.28 at $N = 256$ to ~ 0.23 at $N = 2048$
- This sharpening with system size is consistent with a transition that becomes more pronounced in the thermodynamic limit, rather than washing out

We do not claim universal critical exponents here; rather, we use finite-size sharpening as evidence of a well-defined regime boundary between constrained and unconstrained dynamics.

This parameter structure has implications for biological coding systems. The sharp boundary suggests that biological interfaces operating near it may be sensitive to small changes in bandwidth or coupling strength. Conversely, systems deep in the constrained

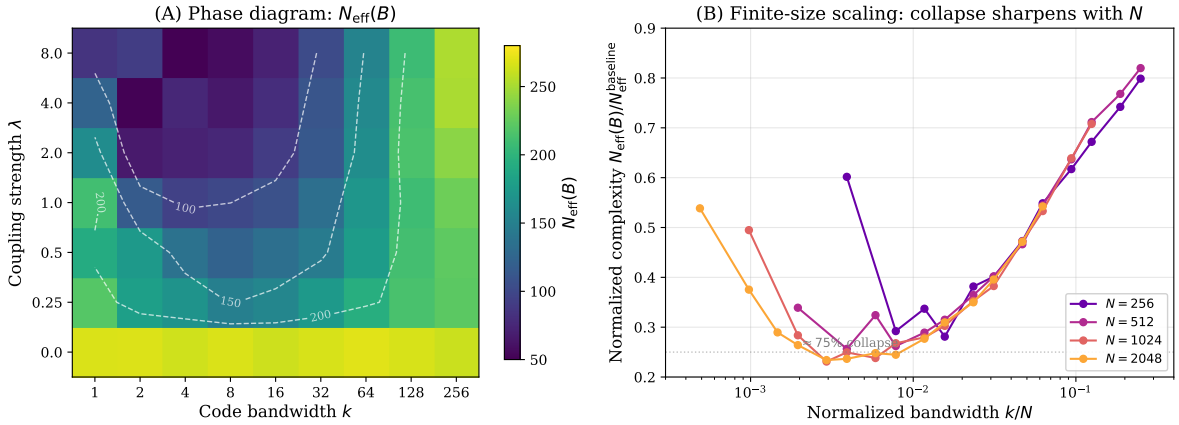

Figure 6: Parameter structure of the constraint regime. (A) Parameter diagram showing $N_{\text{eff}}(B)$ as a function of code bandwidth k and coupling strength λ ($N = 1024$). Contour lines mark $N_{\text{eff}} = 100, 150, 200$. A sharp boundary separates constrained (low k, high λ) from unconstrained regimes. (B) Finite-size scaling at $\lambda = 2.0$ for $N \in \{256, 512, 1024, 2048\}$. Plotting normalized complexity against k/N reveals systematic sharpening with system size—the constraint effect becomes more pronounced as dimensionality increases. At $k/N \approx 0.005$, complexity collapses to $\sim 23\%$ of baseline.

regime should be robust to parameter variation. The scaling behavior—constraint effects strengthening with system size—suggests that dimensional bottlenecks become *more* effective as systems grow more complex, rather than washing out in the thermodynamic limit.

6 Conceptual Framework: Codes as Dimensional Valves

6.1 The Valve Interpretation

The simulation results suggest a general principle: biological interfaces that couple high-dimensional systems through low-dimensional projections function as *dimensional valves*—they constrain the behavioral complexity of responding systems without requiring full state reconstruction.

This reframes the function of biological coding structures:

- **Gene regulatory motifs** don't merely "represent" cellular state; they constrain accessible developmental trajectories.

- **Morphogenetic gradients** don't merely "signal" position; they restrict the dynamical complexity of responding tissues.
- **Neural codes** don't merely "encode" stimuli; they constrain downstream processing to behaviorally relevant dimensions.

The key insight is that constraint and representation are distinct functions. A code can effectively constrain dynamics while having poor predictive power—indeed, discarding destabilizing microstate detail may be precisely why the code works.

6.2 When Coarse-Grained Models Are Reliable

The framework offers practical criteria for evaluating low-dimensional models:

1. **Does the projection capture coherent collective variables?** Random or arbitrary projections will not produce stable constraint. Effective codes must select dynamically relevant degrees of freedom.
2. **Is there a responding system being constrained?** If the "code" is merely an observer's summary with no causal role, it may misrepresent the underlying dynamics. Constraint requires causal coupling.
3. **Does complexity collapse predict stability?** If reducing the code's bandwidth produces chaotic rather than simplified dynamics in the responding system, the projection is not capturing constraining structure.

These criteria distinguish cases where order-parameter descriptions are reliable (because they capture genuine constraints) from cases where they are misleading (because they are arbitrary projections onto a richer dynamics).

6.3 The Membrane as Paradigm Case

The cell membrane exemplifies this framework (Kholodenko, 2006). It is a physical dimensional bottleneck: the external environment and internal cytoplasm are each enormously high-dimensional, yet coupling occurs through a restricted set of interface variables—receptor occupancy, ion channel conductance, transmembrane voltage.

The membrane does not "represent" the environment; it *constrains* which environmental features can influence cellular dynamics. Evolution has tuned both the bandwidth (interface dimensionality) and structure (which degrees of freedom couple) to maximize actionable constraint per unit metabolic cost (Niven and Laughlin, 2008).

This perspective suggests that "information processing" metaphors capture only part of membrane function. The membrane is not merely a channel that transmits environmental information; it is also a valve that restricts cellular dynamics to viable trajectories given environmental context.

7 Relation to Existing Theory

Our framework connects to several established theoretical traditions:

Synergetics and slaving: Haken's synergetics describes how fast variables become "slaved" to slow order parameters (Haken, 1983). We extend this to coupled systems: the code is a bandwidth-limited order parameter that constrains a separate responding system.

Information bottleneck: The information bottleneck principle (Tishby et al., 1999) optimizes compression for predictive relevance. Our framework differs: we measure how compression constrains downstream dynamics, not how much predictive information is preserved.

Markov blankets: The Markov blanket formalism (Friston, 2013) identifies boundaries separating internal from external states. We add a quantitative dimension: bandwidth k specifies how many degrees of freedom can cross the blanket.

Computational mechanics: ε-machines characterize minimal sufficient statistics for prediction (Crutchfield and Young, 1989). We propose a sibling concept: *minimal stabilizing interfaces*—projections that constrain dynamics while potentially being poor predictors.

8 Methodological Guidelines

The constraint interpretation suggests practical changes to how researchers across fields approach low-dimensional descriptions. We offer these not as prescriptions but as reframings that may prove useful.

Choosing reaction coordinates in molecular dynamics: The standard criterion for a good reaction coordinate is that it captures the slowest relaxation timescale (i.e., separates folded from unfolded). The constraint perspective suggests an additional criterion: does the coordinate *restrict accessible conformational dynamics* in the bound or functional state? A coordinate that merely summarizes the ensemble (thermometer) differs from one that actively constrains it (thermostat). The signature would be complexity collapse in conformational fluctuations when the coordinate is held fixed.

Interpreting single-cell manifolds: When UMAP or diffusion maps reveal low-dimensional structure, the standard interpretation is that this reflects underlying regulatory simplicity. The constraint perspective suggests an alternative: low-dimensional structure may reflect bandwidth limitations of inter-cellular coupling (e.g., signaling through a limited number of receptor types). Testing this would involve comparing manifold dimensionality in tissues with different coupling architectures.

Designing interfaces between models: Multi-scale modeling often couples coarse and fine scales through ad hoc "handshaking" variables. The constraint framework suggests a principled criterion: the interface should select coherent collective variables that induce complexity collapse in the fine-scale system without unbounded tracking error. Random or arbitrary coupling will fail this test.

Understanding why some low-D codes work: Neural population codes, morphogenetic gradients, and genetic regulatory motifs all achieve robust function through low-dimensional interfaces. The constraint interpretation explains *why* they work: they capture coherent collective variables that constrain downstream dynamics onto viable trajectories. Codes fail when they project onto incoherent modes that cannot stabilize the responding system.

Distinguishing constraint from information: Mutual information measures how much an observer can learn about one system from another. But a thermostat and a thermometer have the same mutual information with temperature—only one constrains it. When evaluating biological codes, ask: does reducing the code's bandwidth produce complexity collapse (constraint) or tracking failure (information loss)? The two functions are distinct and should not be conflated.

9 Limitations and Future Directions

We emphasize what the simulation demonstrates directly: complexity collapse and bounded tracking under bandwidth-limited coupling. We hypothesize, but do not directly test, that dimensionality reduction is a prerequisite for attractor formation and persistent organization. Perturbation studies would be needed to verify resilience of the constrained dynamics.

The Fourier basis is convenient analytically but not uniquely privileged. The key property is capturing spatially coherent structure; other structured bases (wavelets, Laplacian eigenfunctions) might serve equally well. The critical distinction is between coherent projections that enable constraint and incoherent projections that do not.

Future work should test these predictions in experimental systems: does manipulating interface bandwidth (e.g., via optogenetic control of receptor expression) produce the predicted complexity collapse in responding systems?

10 Conclusion

This review has argued that the recurring interpretive problems surrounding dimensional mismatch in biology—sloppiness in protein modeling, artifacts in single-cell manifolds, shadows in neural codes—arise from a common source: the assumption that dimensional reduction preserves information rather than imposing constraints.

We proposed a unifying interpretation: low-dimensional interfaces between coupled high-dimensional systems function as stabilizing constraints rather than information channels. The signature of effective constraint is complexity collapse with bounded tracking: the responding system's effective dimensionality decreases systematically with code bandwidth while maintaining alignment with the coarse-grained representation. Critically, this requires structured projections onto coherent collective variables; random projections fail.

The framework offers practical guidance across fields: criteria for choosing reaction coordinates in molecular dynamics, for interpreting manifold structure in single-cell data, for designing interfaces in multi-scale models, and for distinguishing constraint from information in any coding system. The key question is not "how much information does the code preserve?" but "does the code constrain downstream dynamics onto viable trajectories?"

Beyond biology, this architecture suggests a blueprint for *neuromorphic constraint computing*, where system states are controlled not by precise error-correction of individual bits, but by shaping the topological manifolds of high-dimensional substrates.

Speculative extension: social codes. The framework may extend to social and linguistic codes, where signals often appear inefficient by information-theoretic standards yet effectively synchronize group dynamics (Grueter et al., 2020). We note this as a direction for future work rather than a claim of the present review.

The deeper insight may be that biological codes serve dual roles: they transmit information *and* constrain dynamics. These functions are distinct, and focusing exclusively on representation may obscure the equally important constraint role. Recognizing this duality

offers new criteria for evaluating when low-dimensional models are reliable—and why some codes work while others fail.

Acknowledgments

No external funding was received for this work.

Declaration of generative AI use in the writing process

During the preparation of this work the author used Claude (Anthropic) for manuscript drafting, editing, and code development. The author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

CRediT author statement

Ian Todd: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization.

Declaration of competing interest

The author declares no competing interests.

Data availability

All simulation code and data are publicly available at `https://github.com/todd866/code-constraint`.

References

D. S. Adams and M. Levin. Endogenous voltage gradients as mediators of cell-cell communication: strategies for investigating bioelectrical signals during pattern formation. *Cell and Tissue Research*, 352(1):95–122, 2013. doi: 10.1007/s00441-012-1329-4.

D. Balduzzi and G. Tononi. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Computational Biology*, 4(6):e1000091, 2008. doi: 10.1371/journal.pcbi.1000091.

R. B. Best and G. Hummer. Reaction coordinates and rates from transition paths. *Proceedings of the National Academy of Sciences*, 102(19):6732–6737, 2005. doi: 10.1073/pnas.0408098102.

J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Structure, Function, and Bioinformatics*, 21(3):167–195, 1995. doi: 10.1002/prot.340210302.

G. Buzsáki and A. Draguhn. Neuronal oscillations in cortical networks. *Science*, 304(5679):1926–1929, 2004. doi: 10.1126/science.1099745.

T. Chari and L. Pachter. The specious art of single-cell genomics. *PLoS Computational Biology*, 19(8):e1011288, 2023. doi: 10.1371/journal.pcbi.1011288.

M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, and K. V. Shenoy. Neural population dynamics during reaching. *Nature*, 487(7405):51–56, 2012. doi: 10.1038/nature11129.

S. M. Cooley, T. Hamilton, E. J. Deeds, and J. C. J. Ray. A novel metric reveals previously unrecognized distortion in dimensionality reduction of scRNA-seq data. *bioRxiv*, page 689851, 2019. doi: 10.1101/689851.

J. P. Crutchfield and K. Young. Inferring statistical complexity. *Physical Review Letters*, 63 (2):105–108, 1989. doi: 10.1103/PhysRevLett.63.105.

J. P. Cunningham and B. M. Yu. Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, 17(11):1500–1509, 2014. doi: 10.1038/nn.3776.

K. A. Dill and H. S. Chan. From Levinthal to pathways to funnels. *Nature Structural Biology*, 4(1):10–19, 1997. doi: 10.1038/nsb0197-10.

A. K. Engel, P. Fries, and W. Singer. Dynamic predictions: oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, 2(10):704–716, 2001. doi: 10.1038/35094565.

J. C. Flack. Coarse-graining as a downward causation mechanism. *Philosophical Transactions of the Royal Society A*, 375(2109):20160338, 2017. doi: 10.1098/rsta.2016.0338.

H. Frauenfelder, S. G. Sligar, and P. G. Wolynes. The energy landscapes and motions of proteins. *Science*, 254(5038):1598–1603, 1991. doi: 10.1126/science.1749933.

P. Fries. Rhythms for cognition: communication through coherence. *Neuron*, 88(1):220–235, 2015. doi: 10.1016/j.neuron.2015.09.034.

K. Friston. Life as we know it. *Journal of the Royal Society Interface*, 10(86):20130475, 2013. doi: 10.1098/rsif.2013.0475.

J. A. Gallego, M. G. Perich, L. E. Miller, and S. A. Solla. Neural manifolds for the control of movement. *Neuron*, 94(5):978–984, 2017. doi: 10.1016/j.neuron.2017.05.025.

A. P. Georgopoulos, A. B. Schwartz, and R. E. Kettner. Neuronal population coding of movement direction. *Science*, 233(4771):1416–1419, 1986. doi: 10.1126/science.3749885.

J. B. A. Green and J. Sharpe. Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development*, 142(7):1203–1211, 2015. doi: 10.1242/dev.114991.

C. C. Grueter, X. Qi, D. Zinner, T. Bergman, M. Li, Z. Xiang, P. Zhu, A. B. Migliano, A. Miller, M. Krützen, J. Fischer, D. I. Rubenstein, T. N. C. Vidya, B. Li, M. Cantor, and L. Swedell. Multilevel organisation of animal sociality. *Trends in Ecology & Evolution*, 35 (9):834–847, 2020. doi: 10.1016/j.tree.2020.05.003.

R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS Computational Biology*, 3(10):e189, 2007. doi: 10.1371/journal.pcbi.0030189.

H. Haken. *Synergetics: An Introduction*. Springer, 1983. doi: 10.1007/978-3-642-88338-5.

K. Henzler-Wildman and D. Kern. Dynamic personalities of proteins. *Nature*, 450(7172): 964–972, 2007. doi: 10.1038/nature06522.

S. Huang, G. Eichler, Y. Bar-Yam, and D. E. Ingber. Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Physical Review Letters*, 94(12):128701, 2005. doi: 10.1103/PhysRevLett.94.128701.

B. E. Husic and V. S. Pande. Markov state models: From an art to a science. *Journal of the American Chemical Society*, 140(7):2386–2396, 2018. doi: 10.1021/jacs.7b12191.

B. N. Kholodenko. Cell-signalling dynamics in time and space. *Nature Reviews Molecular Cell Biology*, 7(3):165–176, 2006. doi: 10.1038/nrm1838.

A. Kicheva and J. Briscoe. Developmental pattern formation in phases. *Trends in Cell Biology*, 25(10):579–591, 2015. doi: 10.1016/j.tcb.2015.07.006.

J. L. Klepeis, K. Lindorff-Larsen, R. O. Dror, and D. E. Shaw. Long-timescale molecular dynamics simulations of protein structure and function. *Current Opinion in Structural Biology*, 19(2):120–127, 2009. doi: 10.1016/j.sbi.2009.03.004.

D. Kobak and P. Berens. The art of using t-SNE for single-cell transcriptomics. *Nature Communications*, 10:5416, 2019. doi: 10.1038/s41467-019-13056-x.

D. Lähnemann, J. Köster, E. Szczurek, D. J. McCarthy, S. C. Hicks, M. D. Robinson, C. A. Vallejos, K. R. Campbell, N. Beerenwinkel, A. Mahfouz, et al. Eleven grand challenges in single-cell data science. *Genome Biology*, 21:31, 2020. doi: 10.1186/s13059-020-1926-6.

M. Levin. Molecular bioelectricity: how endogenous voltage potentials control cell behavior and instruct pattern regulation in vivo. *Molecular Biology of the Cell*, 25(24):3835–3850, 2014. doi: 10.1091/mbc.e13-12-0708.

M. Levin. Bioelectric signaling: Reprogrammable circuits underlying embryogenesis, regeneration, and cancer. *Cell*, 184(6):1971–1989, 2021. doi: 10.1016/j.cell.2021.02.034.

B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna. Parameter space compression underlies emergent theories and predictive models. *Science*, 342(6158):604–607, 2013. doi: 10.1126/science.1238723.

L. McInnes, J. Healy, and J. Melville. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018. doi: 10.48550/arXiv.1802.03426.

E. K. Miller, M. Lundqvist, and A. M. Bastos. Working memory 2.0. *Neuron*, 100(2):463–475, 2018. doi: 10.1016/j.neuron.2018.09.023.

K. R. Moon, D. van Dijk, Z. Wang, S. Gigante, D. B. Burkhardt, W. S. Chen, K. Yim, A. van den Elzen, M. J. Hirn, R. R. Coifman, et al. PHATE: Visualizing transitions and structure for biological data exploration. *Nature Biotechnology*, 37(12):1482–1492, 2019. doi: 10.1038/s41587-019-0336-3.

N. Moris, C. Pina, and A. M. Arias. Transition states and cell fate decisions in epigenetic landscapes. *Nature Reviews Genetics*, 17(11):693–703, 2016. doi: 10.1038/nrg.2016.98.

J. E. Niven and S. B. Laughlin. Energy limitation as a selective pressure on the evolution

of sensory systems. *Journal of Experimental Biology*, 211(11):1792–1804, 2008. doi: 10.1242/jeb.017574.

F. Noé and C. Clementi. Collective variables for the study of long-time kinetics from molecular trajectories: theory and methods. *Current Opinion in Structural Biology*, 43:141–147, 2017. doi: 10.1016/j.sbi.2017.02.006.

J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes. Theory of protein folding: the energy landscape perspective. *Annual Review of Physical Chemistry*, 48:545–600, 1997. doi: 10.1146/annurev.physchem.48.1.545.

V. S. Pande, K. Beauchamp, and G. R. Bowman. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods*, 52(1):99–105, 2010. doi: 10.1016/j.ymeth.2010.06.002.

A. Pietak and M. Levin. Bioelectric gene and reaction networks: computational modelling of genetic, biochemical and bioelectrical dynamics in pattern regulation. *Journal of The Royal Society Interface*, 14(134):20170425, 2017. doi: 10.1098/rsif.2017.0425.

A. Pouget, P. Dayan, and R. Zemel. Information processing with population codes. *Nature Reviews Neuroscience*, 1(2):125–132, 2000. doi: 10.1038/35039062.

P. T. Sadtler, K. M. Quick, M. D. Golub, S. M. Chase, S. I. Ryu, E. C. Tyler-Kabara, B. M. Yu, and A. P. Batista. Neural constraints on learning. *Nature*, 512(7515):423–426, 2014. doi: 10.1038/nature13665.

G. Schiebinger, J. Shu, M. Tabaka, B. Cleary, V. Subramanian, A. Solomon, J. Gould, S. Liu, S. Lin, P. Berber, et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*, 176(4):928–943, 2019. doi: 10.1016/j.cell.2019.01.006.

D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers. Atomic-level characterization of the structural dynamics of proteins. *Science*, 330(6002):341–346, 2010. doi: 10.1126/science.1187409.

N. Tishby, F. C. Pereira, and W. Bialek. The information bottleneck method. In *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing*, pages 368–377, 1999.

M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna. Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *The Journal of Chemical Physics*, 143(1):010901, 2015. doi: 10.1063/1.4923066.

C. Trapnell. Defining cell types and states with single-cell genomics. *Genome Research*, 25 (10):1491–1498, 2015. doi: 10.1101/gr.190595.115.

A. M. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London B*, 237(641):37–72, 1952. doi: 10.1098/rstb.1952.0012.

L. Van der Maaten and G. Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

F. Varela, J.-P. Lachaux, E. Rodriguez, and J. Martinerie. The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2(4):229–239, 2001. doi: 10.1038/35067550.

C. Weinreb, S. Wolock, B. K. Tusi, M. Socolovsky, and A. M. Klein. Fundamental limits on dynamic inference from single-cell snapshots. *Proceedings of the National Academy of Sciences*, 115(10):E2467–E2476, 2018. doi: 10.1073/pnas.1714723115.

L. Wolpert. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, 25(1):1–47, 1969. doi: 10.1016/S0022-5193(69)80016-0.

Supplementary Figures

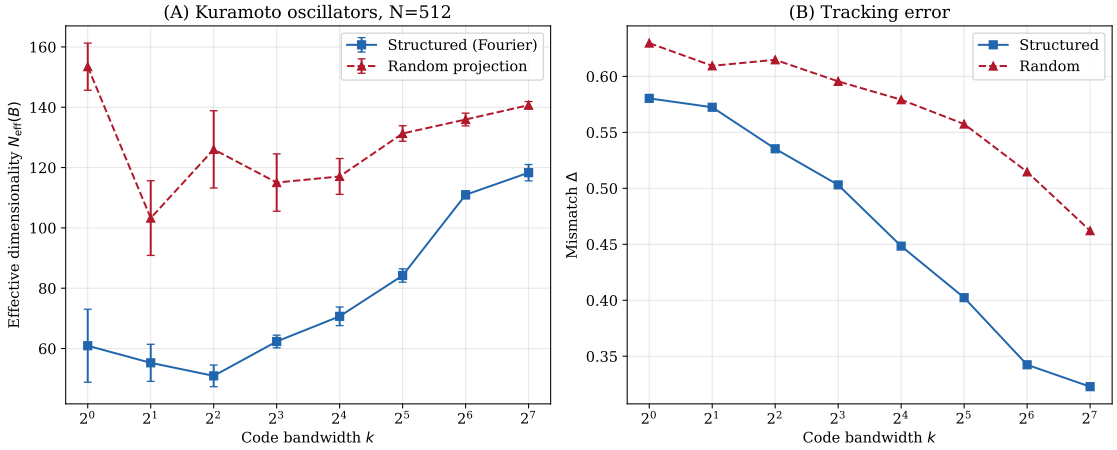

Figure S1. Large-scale Kuramoto simulations ($N = 512$). Complexity collapse persists at larger system sizes. Structured (low-frequency Fourier) projections induce systematic N_{eff} reduction in System B, while random projections show no collapse. The constraint signature is not a finite-size artifact; see Figure 6B for finite-size scaling across $N \in \{256, 512, 1024, 2048\}$.

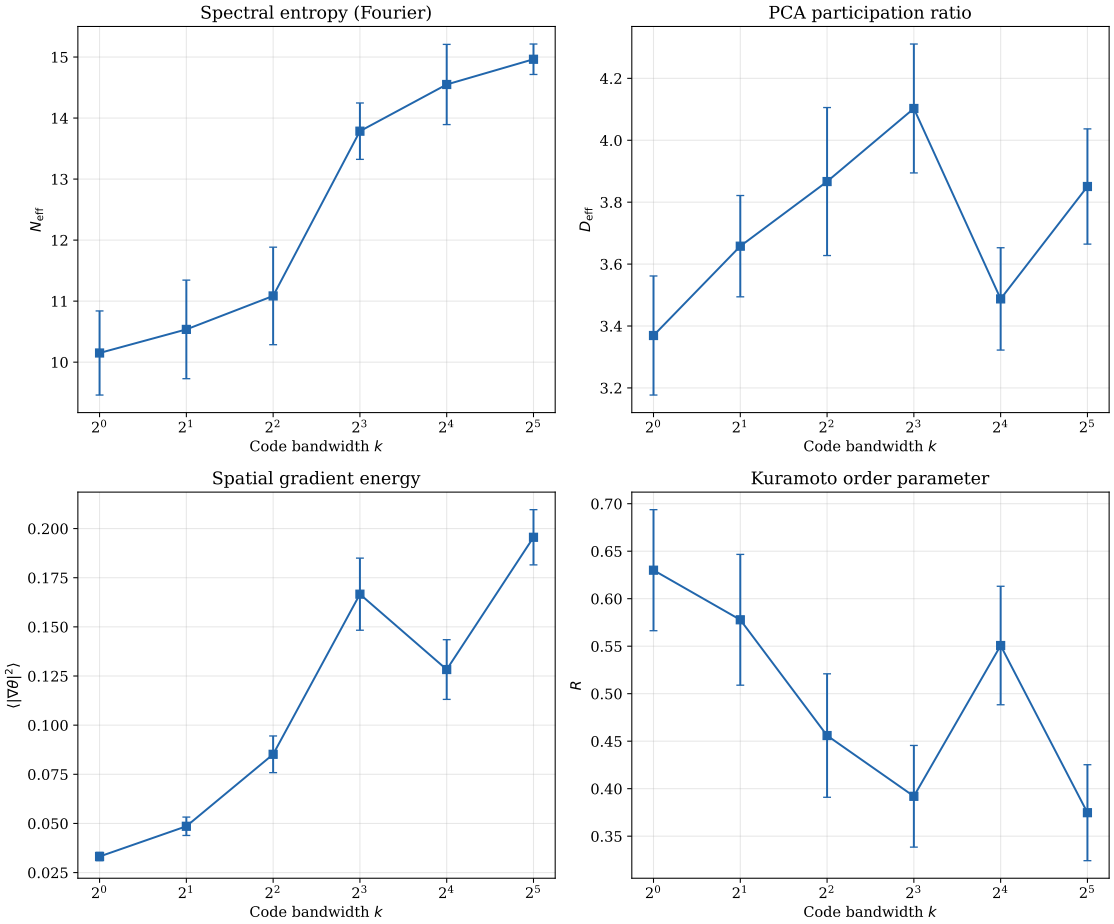

Figure S2. Alternative complexity metrics confirm the constraint signature using measures independent of the Fourier basis. (A) Spectral entropy (primary metric, for reference). (B) PCA participation ratio shows dimensionality collapse in System B. (C) Spatial gradient energy decreases at low k, indicating smoother phase fields. (D) Kuramoto order parameter increases at low k, indicating greater global synchrony. All four metrics show systematic change in System B with bandwidth while System A remains constant.

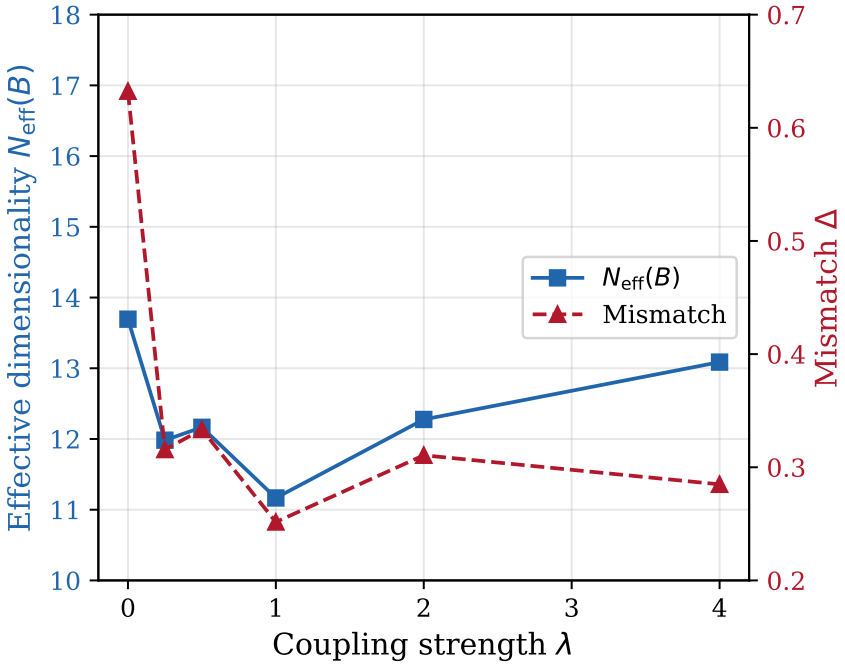

Figure S3. Coupling strength robustness. Effective dimensionality $N_{\text{eff}}(B)$ (blue) and mismatch (red) as a function of coupling strength λ at fixed $k = 8$. Stronger coupling produces greater complexity collapse and tighter tracking. The qualitative pattern persists across the tested range $\lambda \in [0.25, 4.0]$.

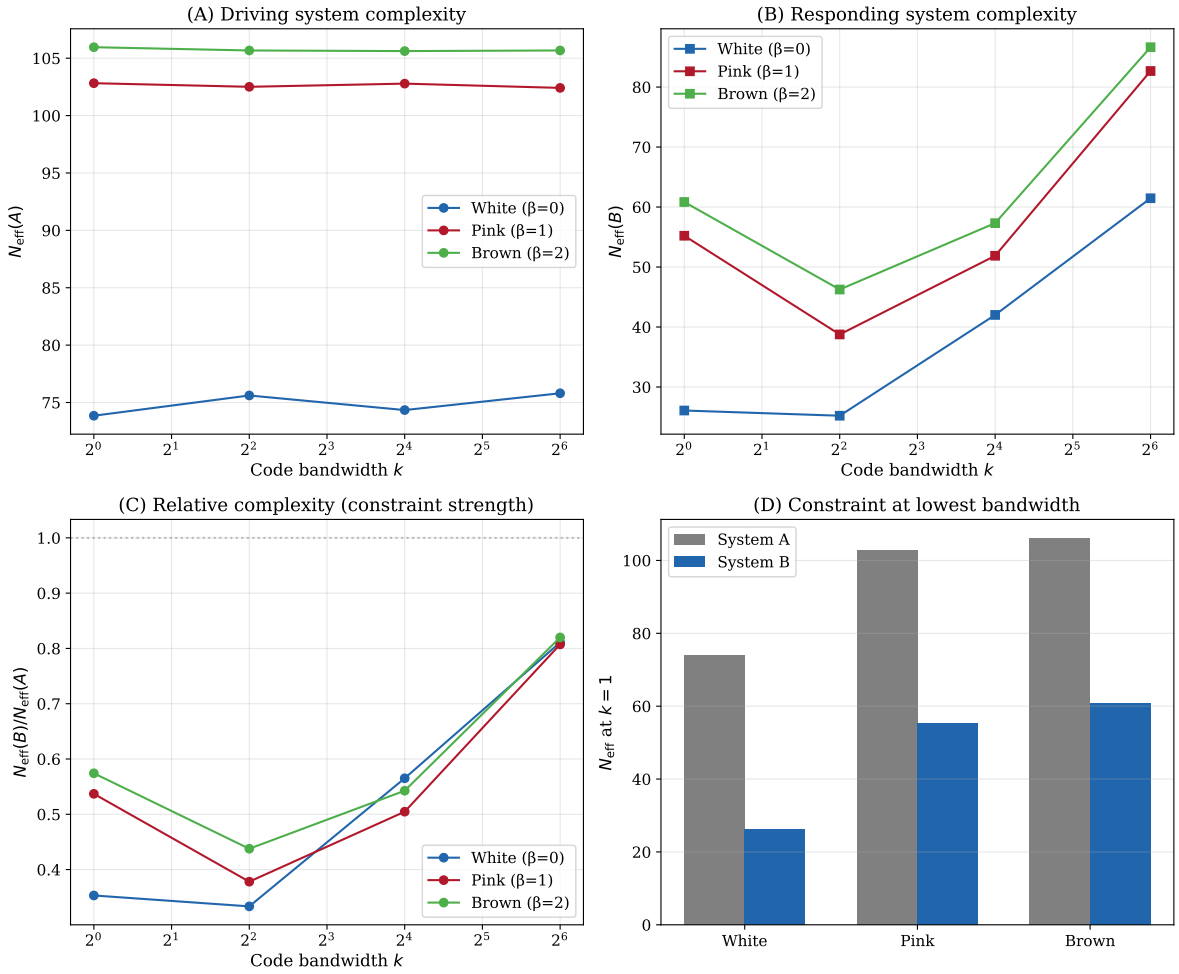

Figure S4. Constraint under fractal ($1/f^{\beta}$) noise. Real biological time series exhibit long-range temporal correlations characteristic of $1/f$ dynamics; we test whether the constraint signature persists under such driving. (A) Driving system complexity increases with noise exponent β (white $\rightarrow$ pink $\rightarrow$ brown), reflecting the richer mode structure of fractal noise. (B) Responding system shows constraint across all noise types, though the pattern becomes non-monotonic under fractal driving—likely because long-memory forcing injects coherent low-frequency structure into A, changing the mapping between k and coherent collective variables. (C) Relative complexity $N_{\text{eff}}(B)/N_{\text{eff}}(A)$ remains below unity at all bandwidths, confirming constraint persists. (D) At minimal bandwidth ($k = 1$), collapse is 65% (white), 46% (pink), and 43% (brown). The constraint mechanism is robust to realistic $1/f$-like temporal correlations, though its magnitude is modulated by noise structure.