Global Dynamics of the *E. coli* Transcriptional Regulatory Network: A Synthesis of Structured RNN Theory, Vector-Field Decomposition, and Topological Data Analysis

Ryan Shin, Sohum Sukhatankar
Harvard College
Cambridge, USA
rshin@college.harvard.edu, ssukhatankar@college.harvard.edu

Abstract

We provide a literature review of global dynamics modeling in large gene regulatory networks (GRNs), using the *E. coli* transcriptional regulatory network (TRN) as a running example. Building on Hopfield attractor models and dense random recurrent network theory, we incorporate sparse, signed, and heavy-tailed connectivity structures, and formulate the network dynamics as continuous-time recurrent neural networks (RNNs) and stochastic differential equations. We formalize the notion of landscape–flux decomposition of the gene network's vector field into gradient (potential) and non-gradient (curl) parts, and we present discrete and continuous Hodge/Helmholtz decomposition theorems for flows on graphs and in Euclidean space. We introduce topological data analysis (TDA) tools, including persistent homology and Reeb graphs, with definitions and theorems for summarizing attractor geometry.

1 Introduction

The *E. coli* transcriptional regulatory network (TRN) is a canonical example of a large, sparse, directed regulatory system in which transcription factors (TFs) serve as nonlinear gene regulators. Specifically, these TFs link the dynamics of different gene expression levels, with these levels serving as the state variables of the system. [Santos-Zavaleta et al., 2019; Gama-Castro et al., 2016]. By global dynamics, we refer to the full long-time behavior of the network: (1) the set and geometry of invariant attractors (stable fixed points, limit cycles, chaotic attractors, etc.), (2) the organization of their basins and transitions, which we characterize through state transition graphs, and (3) whether the dynamics are effectively gradient-like (i.e., describable by an energy or potential landscape) or inherently nonequilibrium with persistent circulation. In biological terms, this corresponds to understanding the metaphorical "epigenetic landscape", wherein detailed balance can be violated due to directed feedback and active driving [Wang, 2015].

This work synthesizes three rigorous frameworks for analyzing these global dynamics in TRNs: (1) Structured recurrent neural network (RNN) theory, which extends classical dense i.i.d. random network results to sparse, sign-constrained, block-structured, and low-rank connectivity; (2) Vector-field decomposition, including Helmholtz/Hodge theorems and nonequilibrium landscape-flux decompositions that split the deterministic flow into conservative and curl components; and (3) topological data analysis (TDA), especially persistent homology and Reeb-graph constructions, to summarize attractor geometry and transitions. We demonstrate how each can (and cannot) be applied under realistic TRN constraints, like partial observability, parameter uncertainty, and limited data. Our original contributions include a regime table linking structural assumptions to predicted

behaviors and observables, as well as a provable example showing how an asymmetric feedback loop yields non-gradient flow captured by Hodge decomposition.

2 Preliminaries and Related Models

We begin with standard Hopfield-type associative-memory networks and dynamic mean-field theory (DMFT) for dense random recurrent networks [Hopfield, 1982; Sompolinsky *et al.*, 1988], briefly summarizing relevant results in order to set up notation.

A Hopfield network with symmetric weights $W = W^\top$ admits a Lyapunov function (aka energy function). For a continuous-rate Hopfield network with dynamics of the form

$$\tau_r \frac{d}{dt} r = -r + f(Wr), \tag{1}$$

where f is monotone (and therefore invertible), the Lyapunov function is given by

$$E(r) = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij} r_i r_j + \sum_{i=1}^{N} \int^{r_i} f^{-1}(u) du. \tag{2}$$

Next, we choose an affine activation $f(u) = u + I$, where the term I incorporates the bias vector, so that $f^{-1}(r) = r - I$. Then

$$\sum_{i=1}^{N} \int^{r_i} f^{-1}(u) du = \sum_{i=1}^{N} \int^{r_i} (u - I_i) du = \frac{1}{2} \|r\|^2 - I^\top r + C,$$

and so

$$E(r) = -\frac{1}{2} r^\top W r + \frac{1}{2} r^\top r - I^\top r + C = -\frac{1}{2} r^\top (W - I_N) r - I^\top r + C.$$

Dropping the additive constant C and absorbing the identity shift into the weight matrix via $\widetilde{W} := W - I_N$, we obtain the quadratic-plus-linear energy

$$E(r) = -\frac{1}{2} r^\top \widetilde{W} r - I^\top r, \tag{3}$$

so that the gradient system $-\widetilde{W} r - I = -\nabla E(r)$ yields gradient-descent dynamics. Here, symmetric $\widetilde{W}$ guarantees convergence to a local minimum of E.

By contrast, Sompolinsky et al. (1988) showed that a fully asymmetric random Gaussian $W_{ij} \sim \mathcal{N}(0, g^2/N)$ with gain $g > 0$ exhibits a sharp transition from a stable fixed point or simple oscillation to chaotic dynamics, causing the gradient structure to break down, when g exceeds a critical value. As we take the limit $N \to \infty$, the transition occurs at $g_c = 1$ for a $\tanh$-like activation. The dynamic mean-field theory (DMFT) approach simplifies the high-dimensional system to stochastic single-neuron dynamics whose statistics satisfy a self-consistent Gaussian process equation, capturing this transition effectively.

Definition 1 (Gradient and Hamiltonian Vector Fields)**.** *A vector field $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be a* gradient field *if there exists a scalar potential $U(\mathbf{x})$ such that $F = -\nabla U$. In a simply-connected domain, F is gradient if and only if its Jacobian is symmetric: $\partial_i F_j = \partial_j F_i$ for all i, j. If F cannot be written in such a form, it is called* non-gradient*.*

3 Gene Regulatory Network Modeling Classes

Global dynamics can be studied at various modeling levels. We outline three major classes. Note that these classes capture different aspects of global behavior, with no single framework being sufficient, and so we synthesize them so that we an connect their predictions through common structure.

Boolean Networks. Here each gene is on/off, with $x_i \in \{0, 1\}$, and updates follow the discrete time rule $x_i(t+1) = f_i(x_1(t), \ldots, x_N(t))$. The network dynamics become a directed graph of 2^N states. Attractors are fixed points or cycles in this graph [Karlebach and Shamir, 2008]. Boolean models excel at enumerating the number of attractors under different logic assumptions, and therefore they are well-suited for answering existence questions (e.g., how many phenotypes are possible given wiring). However, it is difficult to map more complex notions like stability, time scales, or noise to them.

Continuous Deterministic Models. A common formulation uses ordinary differential equations (ODEs) with nonlinear regulatory responses. A general continuous-time RNN form is

$$\frac{dx_i}{dt} = -\gamma_i x_i + \sum_{j=1}^{N} W_{ij}\phi_j(x_j) + I_i, \quad i = 1, \ldots, N, \tag{4}$$

where x_i denotes the (log-)expression level of gene i, $\gamma_i > 0$ is a degradation rate, W_{ij} encodes the regulatory influence of gene j on gene i (with $W_{ij} > 0$ corresponding to activation, $W_{ij} < 0$ to repression), and I_i represents basal transcription or external signals. The nonlinear functions $\phi_j(\cdot)$ model transcriptional regulation by TFs and are typically chosen to reflect binding saturation and cooperative effects (e.g., through a category of nonlinear functions called Hill functions) [De Jong, 2002; Elahi and Hasan, 2018]. This class captures stable expression levels, graded responses, oscillations, or even chaos depending on W and ϕ [Funahashi and Nakamura, 1993].

Symmetric and sign-definite interactions, like those laid out above, often admit a Lyapunov function, but real TRNs are heterogeneous, making them difficult to model with such constraints. Importantly, nonlinearities ϕ_j need not be odd or identical, γ_i vary, and W is sparse and structured. Equation (4) allows importing results from recurrent network theory (e.g., linear stability via eigenvalues of W, and DMFT for large random W) while still acknowledging these violations of classic assumptions.

Stochastic Models. To study temporary stability and noise-driven transitions, we can use stochastic models. At a microscopic level, such fluctuations can be modeled well by chemical master equations (CMEs), but these formulations become intractable for large networks. A standard simplification is to replace the discrete dynamics by a continuous diffusion approximation, which retains the deterministic "drift" (the vector field describing how gene expression would evolve in the absence of stochastic fluctuations), while encoding noise separately through an effective diffusion term. Such an approximation yields a Fokker–Planck equation modeling the time evolution of the probability density $P(\mathbf{x}, t)$ over different states of gene expression, given by

$$\frac{\partial P}{\partial t} = -\nabla \cdot \big(F(\mathbf{x})P\big) + D\Delta P, \tag{5}$$

where $F(\mathbf{x})$ is the aforementioned deterministic drift (as in (4)) and $D > 0$ is an isotropic diffusion coefficient. Equivalently, we can write the stochastic differential equation (SDE)

$$d\mathbf{x} = F(\mathbf{x})dt + \sqrt{2D}d\mathbf{W}(t),$$

where $\mathbf{W}(t)$ is a standard Wiener process. Please see the **addenda** for a derivation of the above two differential equations. We use the SDE representation to connect stochastic effects directly to the familiar deterministic RNN dynamics, which interpret noise as continuous perturbations of the vector field.

This Fokker-Planck formulation provides us with a natural way to model the global dynamical landscape. When a stationary distribution $P_{\text{ss}}(x)$ exists, it gives us an effective potential $U(x) = -\ln P_{\text{ss}}(x)$, while the associated stationary probability current $J_{\text{ss}} = F(x)P_{\text{ss}} - D\nabla P_{\text{ss}}$ quantifies cyclic probability flow. If this current is nonzero, detailed balance is violated, indicating that probability cycles continuously between states, indicating intrinsically nonequilibrium dynamics. These ideas provide the foundation for the vector-field decompositions developed in Sec. 5.

4 Structured RNN Theory beyond Dense Gaussian

Classical DMFT relies on dense, weakly-correlated, Gaussian weights. We discuss 5 ways in which real TRNs violate the assumptions of DMFT, as well as how these deviations change global dynamics and how to model them. In the sections below, we again denote W as the full connectivity/weight matrix of the network.

4.1 Sparsity and Degree Heterogeneity

In a sparse random network with connection probability $p < 1$, each node has mean in-degree pN. In rate models with $\frac{dx}{dt} = -x + gW\sigma(x)$, sparse-graph DMFT shows the chaos transition depends on p and degree distribution [Kadmon and Sompolinsky, 2015]. For example, let $W_{ij} \sim \{0, \pm 1\}$ with $\mathbb{P}(W_{ij} \neq 0) = p$. The effective variance is now pg^2. The critical gain scales as $g_c = 1/\sqrt{p}$ in the limit. Also, degree heterogeneity (e.g., the anomalously high outdegree of global regulators) produces spectral outliers. By the Perron–Frobenius theorem, a node with $k \gg \langle k \rangle$ edges can pull an eigenvalue out of the bulk circle (of radius $g\sqrt{p}$) to $|\lambda| \approx kg$. Such outliers correspond to localized modes (near the central hub) and can trigger instability. This means that sparsity lowers the threshold for irregularity and can create giant eigenmodes tied to high-degree nodes. These results have been quantified for random graphs by Chung et al., 2003.

4.2 Sign Constraints and Block Structure

Real TRNs have mixed activation and repression, which are analogous to excitatory/inhibitory blocks. Suppose the network has M types (e.g. transcription factor classes), and W has a block structure where $W_{ij} \sim \mathcal{N}(0, \sigma^2_{\alpha\beta}/N)$ when i is of type α and j of type β. Aljadeff *et al.* (2015) showed that for such block-structured random networks, the critical stability is determined by the largest eigenvalue of the $M \times M$ "gain matrix" Γ with entries $\Gamma_{\alpha\beta} = \sqrt{\langle k_\alpha \rangle \langle k_\beta \rangle}\sigma_{\alpha\beta}$, rather than a single variance parameter. In other words, even if the overall variance is fixed, heterogeneity among blocks can shift or split the transition. Intuitively, one large block of strong mutual inhibition or activation can dominate the dynamics. A corollary is that a small subset of high-gain regulators, like master TFs, can control the global regime boundary. This has practical implications, namely that identifying dominant regulator classes from the TRN via singular-value analysis of the linearized gain matrix can predict qualitative changes in structure.

4.3 Low-Rank Plus Random Connectivity

Biological networks often overlay a few coherent regulatory axes on top of random connectivity. We can decompose

$$W = gW_{\text{rand}} + R,$$

where W_{rand} is a zero-mean random matrix of variance σ^2/N, and R is a low-rank matrix (of rank $r \ll N$) representing strong modes (e.g., global regulators). By classic results on rank-r perturbations [Benaych-Georges *et al.*, 2011], the spectrum of W consists of two parts: a bulk following the circular law of radius $g\sigma$, and up to r outlier eigenvalues given by the nonzero eigenvalues of R (scaled by g) provided they exceed the bulk. For example, if $R = uv^T$ is rank-one, the bulk has radius $g\sigma$, and there is one outlier at $\lambda = gu^T v$. This implies the network dynamics have one dominant low-dimensional mode (along u) in a sea of random fluctuations [Mastrogiuseppe and Ostojic, 2018]. In GRNs, a small number of "master regulators" or pathways may act as R, inducing slow manifold structure while the rest of the genes contribute to high-dimensional noise. Formal analysis, which can be done through projection onto the subspace spanned by R's vectors, shows that these low-rank modes can create structured attractors independent of the bulk.

4.4 Heavy-Tailed Weight Distributions

If the distribution of nonzero regulatory strengths W_{ij} has heavy tails (precisely, $\mathbb{P}(|W_{ij}| > w) \sim w^{-\alpha}$ for $\alpha < 4$, where smaller α corresponds to heavier tails), the usual variance-based DMFT no longer applies. Recent work [Schuessler *et al.*, 2020] suggests that these heavy-tailed RNNs have a broadened transition, wherein the sharp point g_c dissolves into a range, and the chaotic activity tends to concentrate on a small subset of strongly connected neurons. Under this regime, very large weights can dominate dynamics, since edges are rare. For $\alpha < 2$, the spectral radius of W (a metric of strength as a signal propagates through a network) diverges with N, implying that even weak interaction strength can destabilize the network. The implication for TRNs is that if a few interactions are extremely strong (e.g., essential feedback loops), they may govern stability and induce nonlinear effects beyond what variance-based theory predicts.

4.5 Nonnormal Connectivity

Directed and sparse architectures often yield nonnormal matrices: $WW^T \neq W^TW$. A nonnormal W can exhibit large transient growth even if all eigenvalues lie within the stable unit circle. Specifically, define the matrix norm $\|W\| = \sqrt{\rho(WW^T)}$ (spectral norm) and numerical radius. Even with $\rho(W) < 1$, we can have $\|W^k\| \gg 1$ for some $k \leq N$, leading to large excursions before decay. For example, a chain of feed-forward connections can transiently amplify inputs. The theory of pseudospectra [Trefethen and Embree, 2005] quantifies this: the ϵ-pseudospectrum of W can extend well beyond its spectrum, indicating sensitivity to perturbations. In GRNs, nonnormality (e.g. asymmetric feedback loops) can create brief but significant responses to signals, which might appear as metastable transients even in the absence of true oscillatory attractors. Detecting nonnormality (e.g., via singular value decomposition) can thus diagnose potential hidden slow dynamics.

5 Nonequilibrium Structure: Landscape–Flux Decomposition and Helmholtz Theorem

The metaphor of an epigenetic landscape suggests that gene expression dynamics would follow gradient descent on a potential. However, due to directed feedback and active biochemical cycles, real TRNs are typically far from detailed balance. We now formalize this via vector-field decompositions.

Consider the SDE $\frac{d\mathbf{x}}{dt} = F(\mathbf{x}) + \sqrt{2D}\eta(t)$, where $\eta(t)$ is white noise. The associated Fokker–Planck equation (5) has a steady-state density $P_{ss}(\mathbf{x})$ (assuming it exists and is smooth). Define the nonequilibrium potential $U(\mathbf{x}) = -\ln P_{ss}(\mathbf{x})$. The stationary probability current is

$$\mathbf{J}_{ss}(\mathbf{x}) = F(\mathbf{x})P_{ss}(\mathbf{x}) - D\nabla P_{ss}(\mathbf{x}).$$

In steady state $\nabla \cdot \mathbf{J}_{ss} = 0$. We can decompose the drift as follows. Observe that

$$F(\mathbf{x}) = D\nabla(\ln P_{ss}) + \frac{\mathbf{J}_{ss}}{P_{ss}} = -D\nabla U(\mathbf{x}) + \frac{\mathbf{J}_{ss}(\mathbf{x})}{P_{ss}(\mathbf{x})}.$$

Since $D\nabla U = \nabla(DU)$ is a gradient field, and $\mathbf{J}_{ss}/P_{ss}$ is divergence-free, this splits $F = F_{\text{pot}} + F_{\text{curl}}$ into a conservative part $F_{\text{pot}} = -D\nabla U$ and a purely non-gradient part $F_{\text{curl}} = \mathbf{J}_{ss}/P_{ss}$. This is analogous to the Helmholtz decomposition in vector calculus. In equilibrium (detailed balance) we have $\mathbf{J}_{ss} \equiv 0$, so $F = -D\nabla U$ is gradient and $\nabla \times F = 0$. In nonequilibrium, $\mathbf{J}_{ss} \neq 0$ and F has a nonzero curl component. Biologically, $\mathbf{J}_{ss}$ measures cyclic probability flux (e.g. cell-cycle loops), and the entropy production rate$\sigma = \int \mathbf{J}_{ss} \cdot D^{-1}\mathbf{J}_{ss}/P_{ss}d\mathbf{x}$ quantifies the degree of detailed-balance breaking.

The above is a continuous-space view. For empirical or coarse-grained transition graphs, we use discrete Hodge decomposition.

Definition 2 (Hodge Decomposition on a Graph)**.** *Given a connected graph $G = (V, E)$ (possibly directed) and an assignment of a flow (a scalar on each edge), any flow $f : E \to \mathbb{R}$ can be uniquely decomposed (modulo boundary conditions) into three orthogonal components*

$$f = \nabla\phi + \text{curl}\Psi + h,$$

where $\phi : V \to \mathbb{R}$ defines a gradient (exact) flow on edges, Ψ assigns circulation around cycles (2D faces) giving a curl flow, and h is a harmonic flow with zero divergence and curl (global cycles) [Desbrun et al.*, 2015].*

In practice, we can compute the decomposition of an inferred state-transition graph flow to extract how much is gradient-like vs circulatory. For example, an RNA velocity vector field in gene-expression space can be decomposed into pseudo-temporal progression (gradient) and cyclic (e.g. cell-cycle) components [Qiu *et al.*, 2022; Van den Berge *et. al*, 2020]. In TRNs, we can then take the approach of estimating the empirical or simulated transition flow on a graph and measure the fraction of flow captured by $\nabla\phi$ versus the irreducible curl (see Sec. 7). If a substantial $\text{curl}\Psi$ or h component is present, it signals intrinsic nonequilibrium with cycles that cannot be explained by any scalar potential.

6 Topological Data Analysis for Attractors and Transitions

Topological data analysis (TDA) provides coordinate-free summaries of state-space geometry and transitions (allows us to translate higher dimensional data sets into more tangible characteristics). We focus on two methods.

6.1 Persistent Homology (PH)

Let $X = \{x^i\}_{i=1}^m \subset \mathbb{R}^d$ be a point cloud (e.g. long-time trajectory samples or delay embeddings of GRN states). Given a scale parameter $\epsilon \geq 0$, we build a Vietoris–Rips simplicial complex $VR(X, \epsilon)$: include a k-simplex for any $k+1$ points that are pairwise within distance ϵ. As ϵ grows, simplices are added, yielding a filtration $\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_L$, where each $K_\ell = VR(X, \epsilon_\ell)$ [Hausmann, 1995; Zomorodian & Carlsson, 2004; Bauer, 2021].

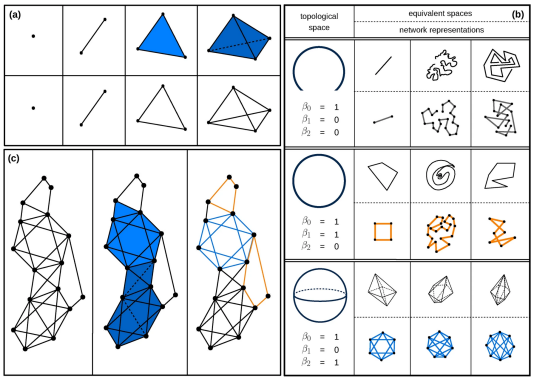

Figure 1: *Credits: Masoomy et al., 2021*

Definition 3 (Homology and Betti Numbers)**.** *The k-th homology group $H_k(K_\ell)$ of a simplicial complex K_ℓ captures its k-dimensional holes. Its rank $\beta_k = \text{rank}H_k$ is the k-th Betti number, such that β_0 counts connected components, β_1 counts independent loops, etc. Formally, $H_k = Z_k/B_k$ is cycles modulo boundaries (Definition 2) [Edelsbrunner Harer 2008; Zomorodian & Carlsson, 2004].*

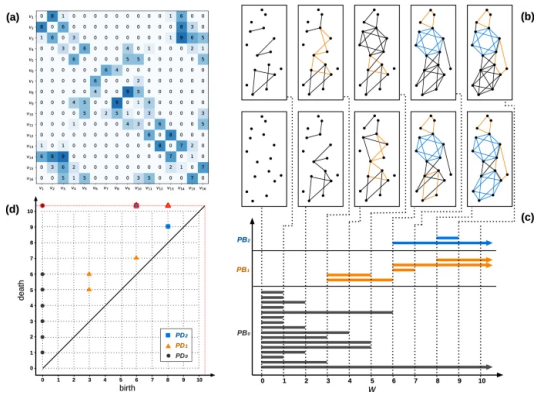

Figure 2: Example of PH, *Credits: Masoomy et al., 2021*

Definition 4 (Persistent Homology)**.** *For each ℓ, we obtain homology groups $H_k(K_\ell)$. Persistent homology (PH) tracks how classes (holes) appear and disappear across scales. Concretely, a k-dimensional homology class α* births *at scale ϵ_b and* dies *(becomes trivial or merges) at $\epsilon_d \geq \epsilon_b$. The collection of intervals $[\epsilon_b, \epsilon_d)$ over all homology classes is the k-th* persistence barcode*. Equivalently, the* persistence diagram *plots each class as a point (ϵ_b, ϵ_d). Longer intervals represent more significant topological features [Edelsbrunner & Harer 2008; Zomorodian & Carlsson 2004]*

A fundamental result (the Structure Theorem for persistence modules) guarantees that these barcodes fully characterize the PH of a finite filtered complex, with each barcode corresponding to a basis element of homology. Importantly, PH is stable, where small perturbations of X (or noise) lead to small changes in the diagram (bottleneck distance) [Cohen-Steiner *et al.*, 2007].

In dynamics, PH can distinguish multistability, which gives us multiple persistent β_0 components (disconnected basins) at small ϵ, while *oscillations* produce a persistent β_1 class (a loop) [Myers 2019; Maletic 2015]. Chaotic attractors can have complicated barcodes with many short-lived features, but often a dominant β_1 or higher-class if the attractor has intrinsic holes. In summary, applying PH to state samples results in signature of basins (H_0 lifetimes) and cycles (H_1 lifetimes) in the dynamics [Maletic 2016].

6.2 Mapper/Reeb-Graph Summaries

Mapper is a topological summarization that produces a graph representation of data. Given a point cloud X and a continuous filter function $f : X \to \mathbb{R}$ (e.g. a coordinate or projection capturing progress), we cover the range of f by overlapping intervals, and clusters the points in each interval's preimage. Each cluster becomes a node, and edges are drawn if clusters in adjacent intervals intersect. The resulting Mapper graph approximates the Reeb graph of (X, f). It captures branching and cyclic structure of the data.

For dynamics, we can use a filter like time or a learned reaction coordinate, and obtain a graph whose loops correspond to oscillatory cycles and branches to bifurcations. Temporal Mapper incorporates directionality by considering time-window overlaps, yielding a coarse directed transition network. In the context of TRNs, Mapper complements PH by explicitly showing how metastable states (nodes) connect via transitions (edges). For instance, a loop in the Mapper graph indicates a stable cycle, whereas a branching point indicates a bifurcation between basins.

7 Proposed Analysis Framework

We now outline a pipeline to analyze TRNs.

7.1 Step 1: TRN Structural Diagnostics

Starting from a TRN (e.g. RegulonDB) database, we first compute graph statistics. This includes degree distributions and sparsity, sign (activator vs repressor) ratios, motif counts (e.g. feedback loops, feedforward loops), nonnormality measures (matrix asymmetry, $\|W\|/\rho(W)$), and low rank structure (singular values or outliers of the Jacobian linearization). These features feed into structured RNN theory to hypothesize dynamic regimes. For example, a broad degree heterogeneity suggests possible unstable modes, a block imbalance suggests which regulators control regime, and motif counts hint at potential cycles or multi-stability.

7.2 Step 2: Model Class Selection

We choose the simplest model class that can express the observed features. As implied in previous sections: if interested only in the existence of attractors (not their stability or transients), a Boolean network might suffice; if oscillations or graded responses are key, use a continuous ODE model (4); if metastability or noise-driven switching matters, incorporate stochasticity via SDEs or master equations. The principle is based on alignment, so detect oscillations only if the model allows cycles.

7.3 Step 3: Flow Estimation and Hodge Decomposition

From time-series data (real or simulated), construct an estimated flow on a coarse state-space graph. For a continuous model, we can sample many trajectories and bin states; for data-driven approaches, we may infer a vector field via Gaussian process regression or neural ODE fitting. On the resulting directed graph of states, compute the discrete Hodge decomposition (Definition 2). Calculate the circulation fraction

$$\Gamma = \frac{\|\mathrm{curl}\Psi\|^2 + \|h\|^2}{\|\nabla\phi\|^2 + \|\mathrm{curl}\Psi\|^2 + \|h\|^2},$$

which lies in $[0, 1]$. Here $\|\cdot\|$ measures the sum of squared flows. A small Γ indicates nearly gradient-like (detailed-balance) dynamics, while a large Γ indicates strong cyclic flux. Also identify which cycles carry most flow (e.g. which regulatory circuits). This quantifies detailed-balance breaking and reveals cyclic programs.

7.4 Step 4: Topology of Attractors and Transitions

Using the same data or flow, sample representative trajectories in each putative attractor basin. Compute persistent homology on these samples (possibly via delay embedding) to find H_0 and H_1 features. Record the number of persistent β_0 bars (distinct basins) and any β_1 cycle (limit cycle). In parallel, build a temporal Mapper network to summarize how states transition among basins.

While we were unable to execute the pipeline on a real dataset, we performed related calculations/proofs through a simple worked example. Please reference the addenda for more information.

8 Conclusion

This paper presents a collection of mathematical tools for reasoning about global dynamics in large transcriptional regulatory networks (TRNs), and connected continuous-time GRN/TRN models of the form (4) (and the stochastic formulation (5)) to gradient/Lyapunov structure, nonequilibrium landscape–flux decompositions, discrete Hodge decompositions on coarse transition graphs, and persistent-homology. Due to time constraints, we were not able to implement a full pipeline on a curated *E. coli* TRN.

References

[1] J. Aljadeff, M. Stern, and T. Sharpee, "Transition to chaos in random networks with cell-type-specific connectivity," *Phys. Rev. Lett.*, vol. 114, 088101, 2015.

[2] U. Bauer, "Ripser: Efficient computation of Vietoris–Rips persistence barcodes," *Journal of Applied and Computational Topology*, vol. 5, pp. 391–423, 2021.

[3] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, "Stability of persistence diagrams," *Discrete Comput. Geom.*, vol. 37, pp. 103–120, 2007.

[4] Hidde De Jong. "Modeling and simulation of genetic regulatory systems: a literature review," *Journal of Computational Biology*, 9(1):67–103, 2002.

[5] M. Desbrun, A. N. Hirani, M. Leok, and J. E. Marsden, "Discrete Exterior Calculus," in *ArXiV*, 2015.

[6] H. Edelsbrunner and J. Harer, "Persistent homology—a survey," in *Surveys on Discrete and Computational Geometry*, Contemporary Mathematics, vol. 453, Amer. Math. Soc., 2008, pp. 257–282.

[7] F. E. Elahi and A. Hasan. "A method for estimating Hill function–based dynamic models of gene regulatory networks," *Royal Society Open Science*, 2018.

[8] K. Funahashi and Y. Nakamura. "Approximation of dynamical systems by continuous time recurrent neural networks," *Neural Networks*, 6(6):801–806, 1993.

[9] J. Gama-Castro *et al.*, "RegulonDB 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond," *Nucleic Acids Res.*, vol. 44, D133–D143, 2016.

[10] J.-C. Hausmann, "On the Vietoris-Rips Complexes and a Cohomology Theory for Metric Spaces", Princeton University Press, 1995, pp. 175–185.

[11] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. USA*, vol. 79, no. 8, pp. 2554–2558, 1982.

[12] Van den Berge, K., Roux de Bézieux, H., Street, K. *et al.* Trajectory-based differential expression analysis for single-cell sequencing data. *Nat Commun* 11, 1201 (2020). https://doi.org/10.1038/s41467-020-14766-3

[13] J. Kadmon and H. Sompolinsky, "Transition to Chaos in Random Neuronal Networks," *Phys. Rev. X*, vol. 5, no. 4, 041030, 2015.

[14] G. Karlebach and R. Shamir, "Modelling and analysis of gene regulatory networks," *Nat. Rev. Mol. Cell Biol.*, vol. 9, pp. 770–780, 2008.

[15] S. Maletić, "Persistent homology of dynamical systems: construction and analysis of simplicial complexes from time series," *PLoS ONE*, 2016.

[16] T. S. Martin, J. M. Johannes, and D. Werner. "Comparative study of three commonly used continuous deterministic methods for modeling gene regulation networks," *BMC Bioinformatics*, 11:459, 2010.

[17] Masoomy, H., Askari, B., Tajik, S. *et al.* "Topological analysis of interaction patterns in cancer-specific gene regulatory network: persistent homology approach." Sci Rep 11, 16414 (2021). https://doi.org/10.1038/s41598-021-94847-5.

[18] F. Mastrogiuseppe and S. Ostojic, "Linking connectivity, dynamics, and computations in low-rank recurrent neural networks," *Neuron*, vol. 99, no. 3, pp. 609–623, 2018.

[19] A. Myers, E. Munch, and F. A. Khasawneh, "Persistent homology of complex networks for dynamic state detection," arXiv:1904.07403, 2019.

[20] J. D. Perea and J. Harer, "Sliding windows and persistence: An application of topological methods to signal analysis," *Found. Comput. Math.*, vol. 15, pp. 799–838, 2015.

[21] X. Qiu *et al.*, "Mapping transcriptomic vector fields of single cells," *Cell*, vol 185, 10.1016/j.cell.2021.12.045 External Link, 2022.

[22] H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications*, 2nd ed., Springer, Berlin, 1996.

[23] Z. Santos-Zavaleta *et al.*, "RegulonDB v10.5: Tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12," *Nucleic Acids Res.*, vol. 47, D212–D220, 2019.

[24] F. Schuessler, et al., "Dynamics of random recurrent networks with correlated low-rank structure," *J. Math. Neurosci.*, vol. 2, 2, 2020.

[25] H. Sompolinsky, A. Crisanti, and H.-J. Sommers, "Chaos in Random Neural Networks," *Phys. Rev. Lett.*, vol. 61, pp. 259–262, 1988.

[26] B. J. Stolz, H. Harrington, and S. Porter, "Persistent homology of time-dependent functional networks constructed from coupled time series," *Chaos*, vol. 27, 047410, 2017.

[27] L. Trefethen, M. Embree, "Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators," *Princeton University Press*, 2005.

[28] J. Wang, "Landscape and Flux Theory of Non-equilibrium Dynamical Systems with Application to Biology," *Advances in Phys.*, vol. 148, pp. 1–101, 2015.

[29] X. Zhang, et al., "Temporal Mapper: Transition networks in simulated and real neural dynamics," *Netw. Neurosci.*, vol. 259, pp. 431-460, 2023.

[30] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete & Computational Geometry*, vol. 33, no. 2, pp. 249–274, 2004.

Addenda

Ryan Shin, Sohum Sukhatankar
Harvard College
Cambridge, USA
rshin@college.harvard.edu, ssukhatankar@college.harvard.edu

Stochastic Model Derivations. Following Risken (2), the general Fokker–Planck (FP) equation for N continuous state variables $x = (x_1, \ldots, x_N)$ is

$$\frac{\partial P(x,t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i}\left[D_i^{(1)}(x)P(x,t)\right] + \sum_{i,j=1}^{N} \frac{\partial^2}{\partial x_i \partial x_j}\left[D_{ij}^{(2)}(x)P(x,t)\right], \tag{1}$$

which is the Kolmogorov forward equation associated with a continuous Markov process.

We now specialize (1) to the case used in the main text. Let the drift be given by the deterministic (noise-free) ODE field $F(x)$,

$$D_i^{(1)}(x) = F_i(x), \tag{2}$$

and assume additive, isotropic noise with constant diffusion coefficient $D > 0$, so the diffusion tensor is

$$D_{ij}^{(2)}(x) = D\delta_{ij}. \tag{3}$$

Substituting (2) into the first term of (1) gives

$$-\sum_{i=1}^{N} \frac{\partial}{\partial x_i}\left(F_i(x)P(x,t)\right) = -\nabla \cdot \left(F(x)P(x,t)\right).$$

Substituting (3) into the second term yields

$$\sum_{i,j=1}^{N} \frac{\partial^2}{\partial x_i \partial x_j}\left(D\delta_{ij}P(x,t)\right) = D\sum_{i=1}^{N} \frac{\partial^2 P(x,t)}{\partial x_i^2} = D\Delta P(x,t),$$

where Δ denotes the Laplacian. Collecting terms, we obtain the FP equation

$$\frac{\partial P(x,t)}{\partial t} = -\nabla \cdot \left(F(x)P(x,t)\right) + D\Delta P(x,t). \tag{4}$$

Equation (4) can be equivalently represented at the level of trajectories by an Itô stochastic differential equation (SDE). In general, an Itô diffusion

$$dx = a(x)dt + B(x)dW(t), \tag{5}$$

where $W(t)$ is an N-dimensional Wiener process, has an associated FP equation

$$\frac{\partial P}{\partial t} = -\sum_{i} \frac{\partial}{\partial x_i}\big(a_i(x)P\big) + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}\big((BB^{\top})_{ij}(x)P\big). \tag{6}$$

Matching (6) with (4) identifies the drift $a(x) = F(x)$ and requires

$$BB^{\top} = 2DI.$$

Choosing $B = \sqrt{2D}I$ yields the SDE

$$dx = F(x)dt + \sqrt{2D}dW(t). \tag{7}$$

Then (4) and (7) describe the same stochastic process.

Important note: The super scary equations (1), (5), and (6) are all taken directly (up to notation) from Risken's textbook on the Fokker-Planck equation. We only applied this framework to our choice of drift and diffusion coefficients outlined in (2) and (3) to obtain the differential equations in (4) and (7).

Example Application of the Pipeline Consider the three-gene directed loop

$$\frac{dx_1}{dt} = -x_1 + a\phi(x_3), \quad \frac{dx_2}{dt} = -x_2 + b\phi(x_1), \quad \frac{dx_3}{dt} = -x_3 + c\phi(x_2), \tag{8}$$

with $a, b, c > 0$ and $\phi \in C^2(\mathbb{R})$.

[**Non-gradient character**] Let $\Omega \subset \mathbb{R}^3$ be open and simply-connected. If $\phi'(u) \neq 0$ for all u in the relevant coordinate ranges on Ω, then the vector field F defined by (8) is not a gradient field on Ω.

Proof. Write $F = (F_1, F_2, F_3)$ with

$$F_1(x) = -x_1 + a\phi(x_3), \quad F_2(x) = -x_2 + b\phi(x_1), \quad F_3(x) = -x_3 + c\phi(x_2).$$

If $F = -\nabla U$ on Ω for some $U \in C^2(\Omega)$, then mixed partials commute, so for all i, j

$$\partial_{x_i} F_j = \partial_{x_j} F_i.$$

Compute the cross-partials:

$$\partial_{x_2} F_1 = 0, \quad \partial_{x_1} F_2 = b\phi'(x_1).$$

Hence $\partial_{x_2} F_1 = \partial_{x_1} F_2$ implies $b\phi'(x_1) = 0$ for all $x \in \Omega$, contradicting $b > 0$ and $\phi'(x_1) \neq 0$ on Ω. Therefore such a U cannot exist and F is non-gradient on Ω. □

[**Nonzero stationary current for the stochastic loop**] Consider the SDE on a simply-connected domain $\Omega \subset \mathbb{R}^3$ with reflecting boundary,

$$dX_t = F(X_t)dt + \sqrt{2D}dW_t, \tag{9}$$

where F is the drift in (8) and $D > 0$. Assume there exists a stationary density $P_{ss} \in C^2(\overline{\Omega})$ with $P_{ss}(x) > 0$ on Ω satisfying the stationary Fokker–Planck equation with zero-flux boundary condition. Define the stationary current

$$J_{ss}(x) = F(x)P_{ss}(x) - D\nabla P_{ss}(x).$$

If $\phi'(u) \neq 0$ on Ω, then J_{ss} is not identically zero on Ω.

Proof. Assume for contradiction that $J_{ss}(x) \equiv 0$ on Ω. Then for all $x \in \Omega$,

$$F(x)P_{ss}(x) = D\nabla P_{ss}(x).$$

Since $P_{ss}(x) > 0$, divide to obtain

$$F(x) = D\nabla\big(\ln P_{ss}(x)\big).$$

Thus F is a gradient field on Ω. By Lemma , this contradicts $\phi'(u) \neq 0$ on Ω. Therefore J_{ss} cannot vanish identically. This is the landscape–flux dichotomy, where $J_{ss} \equiv 0$ is equivalent to detailed balance, and non-gradient F forces $J_{ss} \not\equiv 0$ whenever a smooth positive P_{ss} exists [Risken, 1996; Wang, 2015]. □

[**Topology on the TRN-induced transition 1-complex**] Let $S = \{s_1, s_2, s_3\}$ be coarse TRN macro-states corresponding to "activity localized" near gene $1, 2, 3$, and let $Q = (q_{ij})$ be empirical transition rates estimated from trajectories of (9). For $\lambda \geq 0$ define a filtered directed graph

$$G_\lambda = (S, E_\lambda), \quad E_\lambda = \{(i, j) : q_{ij} \geq \lambda\}.$$

View G_λ as a filtered 1-dimensional CW complex. If there exists $\lambda_\star$ such that

$$(1, 2) \in E_{\lambda_\star}, \quad (2, 3) \in E_{\lambda_\star}, \quad (3, 1) \in E_{\lambda_\star},$$

then for all $0 \leq \lambda \leq \lambda_\star$,

$$H_1(G_\lambda; \mathbb{F}) \cong \mathbb{F},$$

and the H_1 persistence barcode of $\{G_\lambda\}_{\lambda \downarrow 0}$ contains an interval $[\lambda_\star, 0)$.

Proof. For $\lambda \leq \lambda_\star$, the 1-skeleton contains a directed 3-cycle. As a 1-complex (no 2-cells), $H_1(G_\lambda; \mathbb{F})$ is the cycle space. Using Euler characteristic for a connected 1-complex,

$$\beta_1(G_\lambda) = |E_\lambda| - |S| + 1.$$

At $\lambda = \lambda_\star$, $|S| = 3$ and $|E_{\lambda_\star}| \geq 3$, and the presence of exactly one independent cycle implies $\beta_1(G_{\lambda_\star}) = 1$, hence $H_1(G_{\lambda_\star}; \mathbb{F}) \cong \mathbb{F}$. As λ decreases, edges can only be added, so the cycle persists for all $\lambda \in [0, \lambda_\star]$ in the 1-complex filtration. Therefore the persistence module in degree 1 has one interval starting at $\lambda_\star$ that does not die within this filtration. This is standard persistence for filtered cell complexes [Zomorodian and Carlsson, 2005]. □

[**Hodge curl detects the same H_1 class**] Let $f : E_{\lambda_\star} \to \mathbb{R}$ be the empirical edge flow on $G_{\lambda_\star}$, and decompose

$$f = \nabla\psi + h$$

as the orthogonal Hodge decomposition on a 1-complex (no 2-cells) [Desbrun et al., 2005]. If

$$f(1,2) + f(2,3) + f(3,1) \neq 0,$$

then $h \neq 0$, and the corresponding homology class in $H_1(G_{\lambda_\star}; \mathbb{F})$ is nontrivial.

Proof. On a 1-complex, the gradient component $\nabla\psi$ has zero circulation on every cycle. In particular, for the directed 3-cycle,

$$(\nabla\psi)(1,2) + (\nabla\psi)(2,3) + (\nabla\psi)(3,1) = 0.$$

Thus the cycle-sum of f equals the cycle-sum of h. If the cycle-sum of f is nonzero, then the cycle-sum of h is nonzero, hence $h \neq 0$. Since h lies in the harmonic subspace (cycle space), it represents a nontrivial H_1 class on $G_{\lambda_\star}$. Therefore the same directed cycle responsible for nonequilibrium circulation in Theorem is detected topologically by H_1 and metrically by the nonzero harmonic/cycle component in Hodge decomposition. □

References

[1] M. Desbrun, A. N. Hirani, M. Leok, and J. E. Marsden, "Discrete Exterior Calculus," in *ArXiV*, 2015.

[2] H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications*, 2nd ed., Springer, Berlin, 1996.

[3] J. Wang, "Landscape and Flux Theory of Non-equilibrium Dynamical Systems with Application to Biology," *Advances in Phys.*, vol. 148, pp. 1–101, 2015.

[4] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete & Computational Geometry*, vol. 33, no. 2, pp. 249–274, 2004.

Acknowledgments

Ryan Shin, Sohum Sukhatankar
Harvard College
Cambridge, USA
rshin@college.harvard.edu, ssukhatankar@college.harvard.edu

The authors contributed equally in gathering information beforehand and revising the final manuscript. After deciding on our original project idea, we employed ChatGPT to find relevant resources, as well as to later search through large swaths of text efficiently within these resources (including both research papers and textbooks) to obtain information covering pertinent formulas, descriptions of models, etc.

Author Sohum Sukhatankar developed the theoretical framework and analysis through Sec. 4.4, including the identification of modeling classes of gene regulatory networks and derivation of relevant equations, formulation of structured RNN models, and the discussion of sparsity, block structure, low-rank perturbations, and heavy-tailed weight distributions. Author Ryan Shin developed the subsequent sections, including the nonequilibrium landscape-flux formalism, Hodge and Helmholtz decompositions, topological data analysis methodology, and the proposed analysis pipeline and illustrative examples.