Probing coexistence of robust threshold and ultrasensitivity in molecular switches and cascades

Anonymous Authors[1]

Abstract

Molecular switches are fundamental building blocks in many biophysical processes. These switches arise from dissimilar molecular interactions, affected by intrinsic and extrinsic noise, but exhibit a common ON-OFF behavior that suffices for signal transduction, amplification, and relay. Such repertoires of different mechanisms pose a fundamental question: which mechanisms in a noisy environment confer ultrasensitivity as a switch while also ensuring robust decoding of the concentration threshold required for information interpretation? In this paper, we design molecular switches based on molecular-exchange mechanisms (MEM) and dimeric ligand formation, comprising reversible interactions, and compare them stochastically with other switch-forming mechanisms. Our results reveal that the MEM and Dimer mechanisms exhibit improved robustness of the threshold against noise and appear less uncertain if assessed information-theoretically. This superiority is reflected in their switch-like behavior, making them less prone to flipping states in noisy environments. Extending beyond single-switch comparisons, we also explore alternative switch cascades and assess their performance against noise. Alongside this, the continuing expansion of this work is transformed to a Physics-Informed Neural Network (PINN) form that can predict switching behavior for variation in input signal. More importantly, the PINN-switch can ensure that any agentic-AI driven design adheres to the biological and physical constraints indispensable in a molecular switch. Insights from this study may be helpful in drug design and synthetic circuits aimed at therapeutics and other aqueous applications.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Submitted to the 2026 Workshop on Generative and Agentic AI for Biology (ICML 2026). Do not distribute.

1. Introduction

Molecular switches are indispensable in biological functions (Simons et al., 2005; Furriols & Bray, 2001; Fishel, 1998) that require cell-cell communication, information transmission, and cells' responses to intrinsic and extrinsic changes. For instance, ligand-receptor interaction activates downstream phosphorylation (Furriols & Bray, 2001; Weiss & Attisano, 2013; Umulis et al., 2009), which subsequently contributes to the cascades of switches in transcriptional and post-transcriptional events. In gene regulation, switches turn on and off transcription and translation processes based on a threshold of a transcription factor, with subsequent regulation by other factors, including microRNA (Zhang et al., 2013). Evidence suggests that aberrant functioning of molecular switches is responsible for diseases and can affect essential processes, including bistability, oscillations and roles in circadian clocks (Del Olmo et al., 2024).

Molecular switches are also crucial in drug design and smart therapeutics. A specific example is the BCR-tyrosine kinase switch in chronic myeloid leukemia, where correcting the underlying switch mechanism can reduce disease severity (Deininger et al., 2000). A good switch is ultrasensitive to small changes in concentration or stimuli (Ferrell & Ha, 2014). However, low-concentration interactions are subject to intrinsic noise (Raser & O'shea, 2005; Plesa et al., 2018; Blake et al., 2003) that a molecular mechanism must mitigate—a counterintuitive problem for the good molecular switch, which remains elusive and is of primary interest in this study. Besides, switch-forming interactions at different layers of cascades differ considerably (Ghusinga et al., 2021), suggesting a probable evolutionary role for context-specific switch preferences and their in vivo functional requirements.

Among alternative switching mechanisms, such as covalent modification through phosphorylation (Goldbeter & Koshland Jr, 1981; Dasgupta et al., 2014), is prevalent in controlling diverse biological functions. Molecular titration—in the activation of transcription upon binding of a transcription factor (Brewster et al., 2014) and the subsequent activation of post-transcriptional regulation of messenger RNA (mRNA) through microRNA (miRNA) (Mukherji et al., 2011)—is a standard switch that turns gene expression

on and off. Titration mechanism (Buchler & Louis, 2008) is also common in synthetic switch design, as used in the case of hinge protein design (Praetorius et al., 2023), which, in the absence or presence of a ligand, switches between alternative conformations. These mechanisms also have limitations in explaining switch-like behavior in other contexts or in satisfying specific performance objectives. For instance, Michaelis–Menten kinetics may be insufficient when substrate binding to the enzyme itself governs the kinetics of subsequent substrate recruitment (Owen et al., 2023). In addition, the assumption of irreversibility in the Michaelis–Menten model is unrealistic, as it fails to account for the physiological relevance of reversible reactions in biological functions (Ortega et al., 2002). The reversibility of reaction is not the only issue; another caveat is the noise arising from intrinsic and extrinsic factors (Thattai & Van Oudenaarden, 2001; Plesa et al., 2018; Bialek & Setayeshgar, 2005). Noise may unexpectedly turn a molecular switch on or off, rendering it ineffective and erroneous (Hu et al., 2011). Specifically, synthetic circuits used in therapeutics must demonstrate precision in their stipulated tasks. In gene regulatory networks, signaling pathways, and neural activity, noise is detrimental but can also facilitate coordinated gene expression or support evolutionary transitions (Eldar & Elowitz, 2010). Many signaling regulators mitigate noise via feedback or additional interactions, but these circuits have limits on their ability to reduce noise (Lestas et al., 2010). Therefore, reversible molecular interactions that reduce noise while exhibiting robust, reproducible thresholds and switch-like behavior are of significant interest.

To address the aforestated limitations, we design molecular switches based on mechanisms observed in dimeric ligand formation (Little & Mullins, 2009; Karim et al., 2021; Shimmi et al., 2005) and the exchange between a receptor (R) and an exchanger (E) via a tripartite complex (ligand:R:E), which is common in many signaling pathways (Umulis et al., 2009; Balemans & Van Hul, 2002). In silico interrogation, using both stochastic and deterministic approaches, reveals the proposed switches' ability to trade off between the ultrasensitivity of switch-like responses and noise mitigation for robust threshold formation. These traits make the dimeric and MEM mechanisms potential alternatives for synthetic switch design. Beyond this, we also design switch cascades comprising layers that follow distinct mechanisms from titration, covalent modification, MEM, or a dimeric switch architecture. Specifically, it searches for potential reasons for the evolutionary choices that underlie the control of many biological functions via order-specific switch cascades.

While these models, starting from a simple molecular switch to cellular systems that capture the multiscale response have significantly strengthened our understanding of cells and their functions and interactions, they fall short in terms of predictive power and analysis speed. For instance, computational models that combine agent-based, stochastic, and deterministic approaches, though can capture intricate details of the orchestration happening at different scales, often suffer from the curse of dimensionality, high nonlinearity, and more importantly, lack predictive power necessary for quick hypothesis testing. Specifically, in-silico interrogation, defined as virtual experiments, screens an enormous hypothesis space comprising order-of-magnitude variations in one or more perturbagens(Miladinovic et al., 2025). Neural networks can step in and provide predictability, generative power, and queryable models—the trio that orchestrates new experimental designs at an incredibly faster pace. In line with this, the PINN-switch we design here, when interfaced with agentic AI, ensures that machine learning models obey physical and biological constraints and laws imbued in synthetic circuit design.

2. Models and Methods

All the deterministic simulations conducted here use Ordinary Differential Equation (ODE) models based on mass-action kinetics. The stochastic analysis uses Gillespie's Stochastic Simulation Algorithm (SSA) (Gillespie, 1976) for approximation of the continuous-time Markov chain (Van Kampen, 1992), as implemented in Stochkit (Li et al., 2008) package available in GillesPy2 (Matthew et al., 2023). The PINN implementation is done using the DeepXDE backend (Lu et al., 2021). We provide a concise representation of the derivation of the switch formulation, and a detailed description of the underlying ODEs and other derivation steps is available in detail in Appendix A.1 and A.3.

2.1. Proposed Molecular Switches

Among the four mechanisms introduced following extracellular and intracellular interactions in a signaling pathway, the titration and covalent modification cycle switches are well-studied. We derive the necessary condition and expression for switch formation using both the dimer and MEM mechanisms. The details are summarized below (see Appendix A.1 for details).

2.1.1. MEM Switch

$$\begin{aligned}
&1.\ \mathrm{N} + \mathrm{R} \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \mathrm{NR}, \qquad 2.\ \mathrm{N} + \mathrm{E} \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} \mathrm{NE},\\
&3.\ \mathrm{NR} + \mathrm{E} \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} \mathrm{NRE}, \qquad 4.\ \mathrm{NE} + \mathrm{R} \underset{k_{-4}}{\overset{k_4}{\rightleftarrows}} \mathrm{NRE},\\
&5.\ \mathrm{NR}, \mathrm{NRE}, \mathrm{NE} \xrightarrow{\delta} \text{null} \quad \text{and} \quad \mathrm{R}_{\mathrm{tot}} = \mathrm{R} + \mathrm{NR} + \mathrm{NRE}
\end{aligned} \tag{1}$$

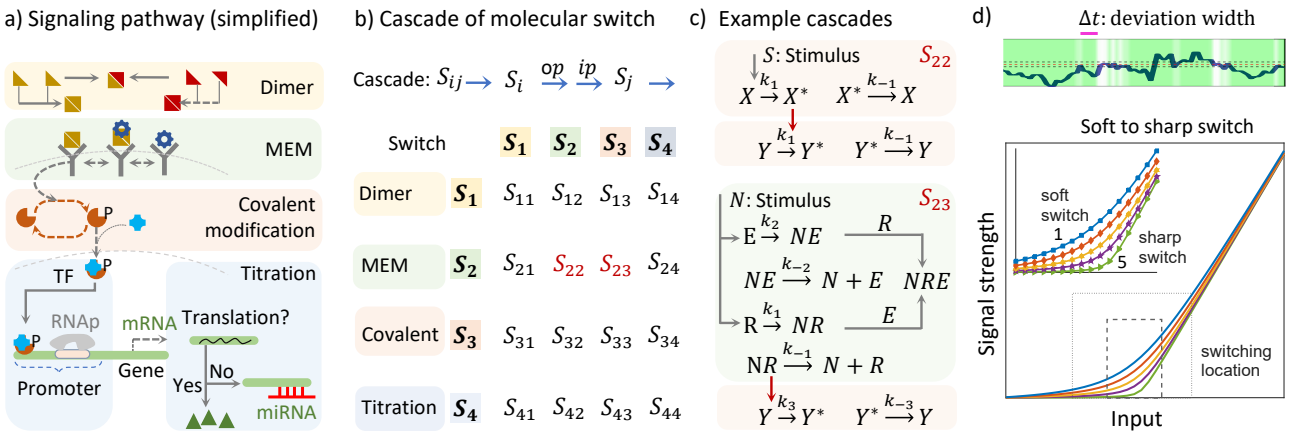

Figure 1. a) Molecular circuitry in signaling pathway: Dimeric ligands produced from monomers are sensed by extracellular receptors. Activated receptors, regulated by extracellular factors, trigger downstream phosphorylation following covalent modifications. A transcription factor, once bound to the promoter region, initiates transcription, exhibiting switch-like behavior. In the post-transcriptional phase, miRNAs regulate mRNA translation, exhibiting a switch-like response. b) Mechanisms sandwiched to form switch cascades. c) Molecular interactions of cascades S_{22} and S_{23}. d) Soft to sharp switch that transforms from soft to hard nonlinearity in switching region.

Using mass-action kinetics, the obtained ODEs of MEM were solved for steady state to devise the switch forming equation as follows

$$[NR] = \frac{-b + \sqrt{b^2 + 4a[R_{\text{TOT}}]}}{2a} \text{where a, b are}$$
$$a = \frac{k_{\text{on}}^{\text{NER}}\mu}{\Delta(\phi + \xi)},$$
$$b = \frac{1}{(\phi + \xi)} + 1 + \frac{k_{\text{on}}^{\text{NRE}}[E]}{\Delta} + \frac{k_{\text{on}}^{\text{NER}}\lambda[N]}{\Delta(\phi + \xi)}, \text{and}$$
$$\Delta = k_{\text{off}}^{\text{NRE}} + k_{\text{off}}^{\text{NER}} + \delta \tag{2}$$

For the detailed derivation on how the Eq. 2 arises and the exact replaceable for ϕ and ξ, see appendix A.1.1.

2.1.2. Dimer Switch

$$1.X + X \underset{k_1}{\overset{k_{-1}}{\rightleftharpoons}} XX, \qquad 2.X + Y \underset{k_2}{\overset{k_{-2}}{\rightleftharpoons}} XY,$$
$$3.Y + Y \underset{k_3}{\overset{k_{-3}}{\rightleftharpoons}} YY \qquad [X_{\text{TOT}}] = [X] + 2[XX] + XY,$$
$$[Y_{\text{TOT}}] = [Y] + 2[YY] + XY \tag{3}$$

From steady state analysis, and subsequent simplifications, we obtain the following quadratic form

$$XX = \frac{-b + \sqrt{(b^2 + 4 \cdot a \cdot 2\gamma[XY]^2)}}{2a} \tag{4}$$

where $a = 2$ and $b = [Y_{\text{TOT}}] - [X_{\text{TOT}}] - [Y] + [X]$. For a case $Y_{\text{TOT}} = X_{\text{TOT}}$, the threshold concentration of $[XX]$ formulates as

$$[XX] = \frac{([Y] - [X]) + \sqrt{([Y] - [X])^2 + 16 \cdot \gamma \cdot [XY]^2}}{4} \tag{5}$$

2.2. Stochastic Simulation

To study how particles evolve in a spatially homogeneous chemical system, we apply Gillespie's Stochastic Simulation Algorithm (SSA) to each of the four switch models and assess their fluctuations. The SSA describes the time-evolution of N chemical species, $S = \{S_1, S_2, \ldots, S_N\}$, interacting over M different unidirectional chemical reactions, $R = \{R_1, R_2, \ldots, R_M\}$, by approximating the Chemical Master Equation (CME) in Eq. 6.

$$\frac{dP(\mathbf{n}, t)}{dt} = \sum_{j=1}^{R} \left[a_j(\mathbf{n} - \boldsymbol{\nu}_j, t)P(\mathbf{n} - \boldsymbol{\nu}_j, t) - a_j(\mathbf{n}, t)P(\mathbf{n}, t)\right] \tag{6}$$

The number of molecules of each species S_i is a set of random variables $X = \{X_1, X_2, \ldots, X_N\}$. For example, at any time, t, the number of molecules of S_1 is X_1.

In Eq. 6, $a_j(\mathbf{n}, t)$ is the propensity function and $\boldsymbol{\nu}_j$ is the stoichiometric vector. The CME determines the probability that the species $S_1, \ldots, S_N$ will have certain values $X_1, \ldots, X_N$ at any given time. So, $X_i(t)$ describes the temporal evolution of the species S_i over time. The Markovian (memoryless) property of the SSA ensures that at any point in time, the prediction of the system's next state depends only on the current number of molecules present (Van Kampen, 1992).

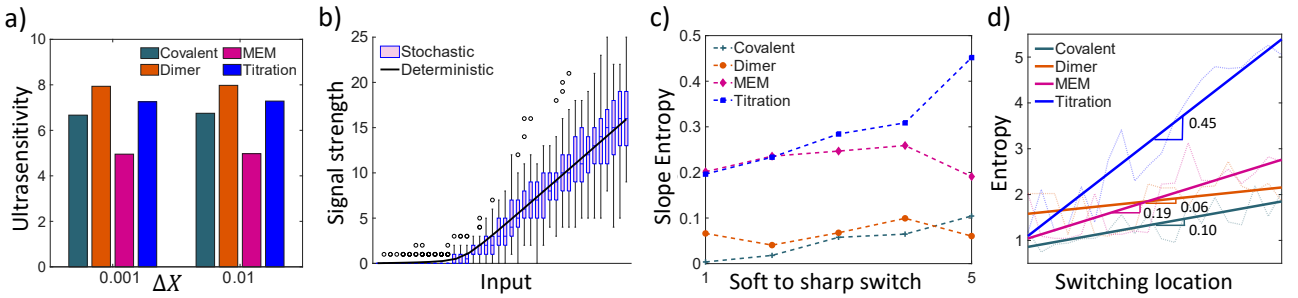

Figure 2. a) Mean ultrasensitivity for each model averaged over 5 different switch sharpness. b) Spontaneous flip of switch state due to fluctuations. c) Slope of the line that linearly fits entropy change in the switching region for soft to sharp switches. Here, MEM and Dimer switch maintain steady slope. d) Entropy change in the switching region and the line fit for the different switches.

2.3. Threshold Assessment: Robustness and Fidelity Analysis

We tune the error definition between exact and approximate distributions, as used in (Rijal & Mehta, 2025), to assess the fidelity of the dynamics of stochastic trajectories. For the pdfs P_t and $Q_{t+\Delta t}$ of signaling molecule S at time t and $t+\Delta t$, respectively, the ratio between Jensen-Shannon divergence JSD $(P_t||Q_{t+\Delta t})$ and entropy $H(Q_{t+\Delta t})$ captures the uncertainty evolution in time. Precisely, the ratio quantifies whether uncertainty grows or shrinks within a Δt interval, and the formulation goes as follows:

$$\text{JSD}(P_t||Q_{t+\Delta t}) = \frac{1}{2}D_{KL}(P||R) + \frac{1}{2}D_{KL}(Q||R),$$
$$\text{where } R = \frac{P+Q}{2}$$
$$H(P) = \sum_{x\in\mathbf{X}} p(x)\log\frac{1}{p(x)}, \text{and}$$
$$D_{\text{KL}}(A||B) = \sum_i A(i)\log\frac{A(i)}{B(i)} \tag{7}$$

We evaluated the fluctuation of the steady-state threshold after allowing the dynamics to evolve for long time and used: i) Coefficient of variations, measured as the ratio between standard deviation σ of species S and its mean μ, ii) deviation width (Δt) distribution that captures a molecular mechanism's tendency to commit smaller Δt and avoid larger Δt in the presence of noise. Here, t_i, evaluated every second, is a flag that is 1 if S deviates beyond $\mu_S \pm \beta\mu_S$, and 0 otherwise. A deviation width of $\Delta t = 5s$ needs $t_i = 1, \forall i = \{1,2,3,4,5\}$.

2.4. Ultrasensitivity and Dynamic Range of Switch

In a molecular switch, ultrasensitivity (U) captures the fractional change in the output (Y) induced by a fractional change in the input (X). The formulation goes as follows

$$\text{U(switch)} = \lim_{\Delta X\to\infty}\frac{(\Delta Y/Y)}{(\Delta X/X)} = \frac{d\ln X}{d\ln Y} \tag{8}$$

The response is ultrasensitive if $|U(\text{switch})| > 1$, and remains linear for $|U(\text{switch})| = 1$. Here, ultrasensitivity analysis consider the steady-state data of the species of interest. For the switch cascade comparison, we calculate the slope $\Delta Y/\Delta X$ and the analysis on switching region evaluate ratio $\text{Stimulus}_{\text{at }\beta}/\text{Stimulus}_{\text{at }\alpha}$, where $\alpha = 0.1$ and $\beta = 0.9$ denote the fractional occupancy Y^*/Y_{T} at the output of a switch cascade (see Fig. 1b).

2.5. PINN Implementation

A Physics-Informed Neural Network (PINN) obeys the laws of physics when solving a set of partial differential equations, supported by the initial and boundary conditions, by incorporating the physical laws directly into the loss functions (Raissi, 2018; Raissi et al., 2019; Cuomo et al., 2022). While the PINN has gained popularity in solving both forward and inverse problems, their utility in predicting biophysical switching behavior remains limited (Ji et al., 2021; Ahmadi Daryakenari et al., 2024a). In contrast, a reliable, robust PINN implementation of synthetic biological circuits has the potential to support further experimental design and computationally efficient exploration of large kinetic spaces (Liu et al., 2025). For instance, PINN implementation of Turing Patterns can be used to extract parameters, that are otherwise difficult to track, for various patterns (Kho et al., 2022). Thus, the neural network approach can provide viable predictions for a large array of problems and free up the computational cost of having to look exhaustively through large parameter spaces (Ahmadi Daryakenari et al., 2024b; Zhai et al., 2023). As such, we design a PINN aimed at predicting the switching behavior arising out of a Chemical Reaction Network (CRN) of a Dimer switch (see Section 2.1.2). Specifically, the PINN model is trained for the Dimer ODEs with conservation conditions imposed on the monomer species X_{TOT} and Y_{TOT}, as shown in Eq. 3. Another factor of key interest when designing a neural network is the search for an optimal architecture (Cai et al., 2018). Methods may constitute the automated search for number of layers, hyperparameters such as the learning rate or number of units of fully-connected layers (Elsken et al., 2019). A small scale search for an optimal network architecture in

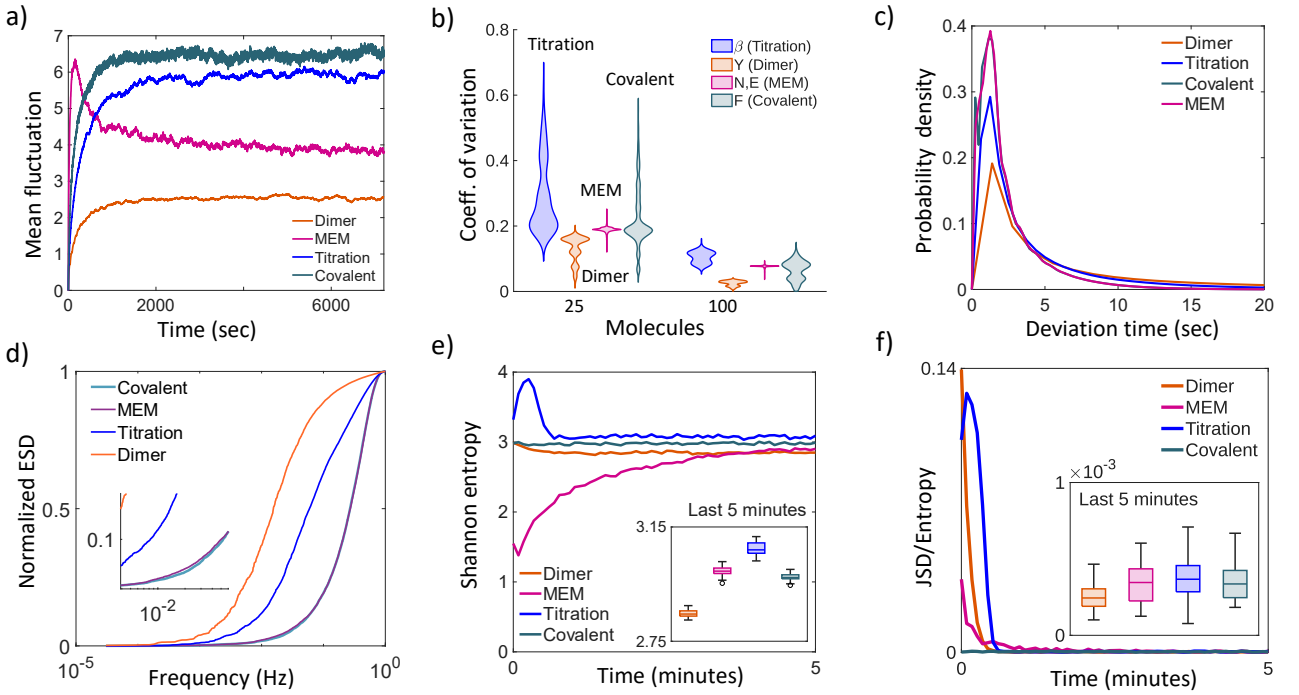

Figure 3. a) Mean-subtracted dynamic fluctuation over 2000 trajectories. d) For target threshold 25, 100 molecules, the CV of 5 hours of steady state data. b) Probability distribution of deviation (Δt) over 200 hours of steady state data, where Δt captures continuing deviation from the target threshold (mean). d) Applying DFT on the steady state signaling obtained using kinetics of the most probable CV calculated in Fig. 3b. e) Shannon entropy calculated over 2000 trajectories of first 5 minutes and last 5 minutes (inset). f) Fidelity assessment through JSD/Shannon entropy, anchored at t using dynamic data from 2000 trajectories between t and $t + \Delta t$.

Table 1. Comparison of training time and GPU memory usage between time-marching models and single PINN model evolved over one minute. For the time-marching model, we show the aggregated time of all individual models.

Model	Time/hours	Memory/MiB
One-Minute Model	1.11	14378
Time-Marching	6.66	86208

modeling the PINN is shown in the schematic in Fig. 4a. We screen through layer sizes of 3, 4, 5, 6 and 0.6, 0.8 and 1 million training points.

We take 1 million training points selected through DeepXDE's default pseudo-random sampling. As the training points are sampled once before training starts, training samples larger than this are memory-intensive. The model predicts species concentrations with respect to two input dimensions: the time domain and the initial amount of the conserved species X_{TOT} added to the system, i.e., the input signal to the switch. A 10 minute PINN-predicted trajectory and the ODE solution of species $[XX]$ are shown in Fig. 4d.

Switch design requires the underlying CRN to equilibrate, which often becomes computationally intensive due to slow kinetics. In PINNs, simulations spanning over large time domains lead to error accumulation and inaccurate predictions, and remains a pertinent issue (Chen et al., 2024). Solutions such as time-marching (Wight & Zhao, 2020) propose dividing the temporal domain into smaller sub-domains but require training the PINN separately over each sub-unit. However, as the prediction of a preceding model is used as the initial condition to the next during training, errors in predictions can still accumulate over time. We run a small-scale experiment between time-marching models and a single PINN to assess the cost incurred. The single PINN model simply predicts the trajectory of the species in a time domain of 1 minute. Alternatively, in time-marching the one minute interval is divided into segments of 10 seconds, and a single time-marching model M_i predicts the dynamics within that segment. The output of model M_{i-1} becomes intial condition to model M_i. The additional costs incurred are summarized in Table 1.

We use a *tanh* activation function and Adam optimizer, and take a time-domain of 10 minutes, observed from MATLAB's ode15s solver to be sufficient for the Dimer reaction network's specific kinetics to reach steady-state to compute our switch. Through a network architecture search as shown in Fig. 4, we select optimal layer sizes and number of training points. To investigate the effect of increasing layer size and training points on the predictions, we calculate the L_2

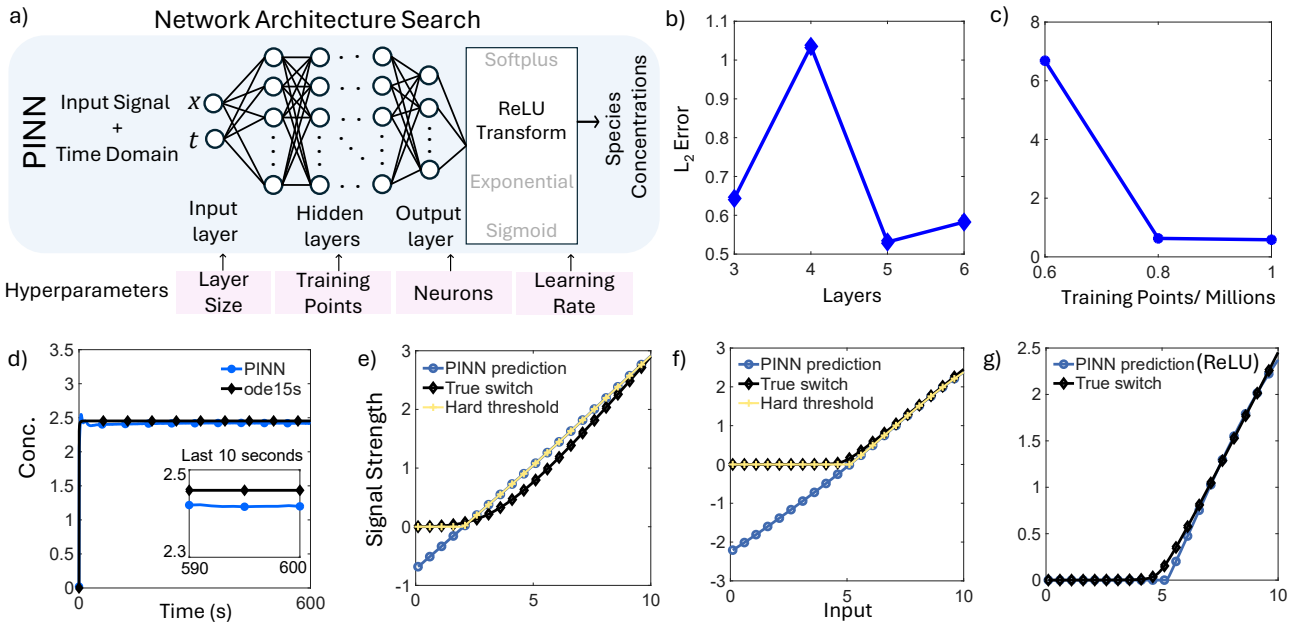

Figure 4. a) Network architecture search for PINN model. L_2 error for varying b) number of layers in the neural network and c) number of training points. d) PINN predicted trajectory and ode15s solution for 600 seconds, $X_{\text{TOT}} = 5$ and $Y_{\text{TOT}} = 5$. Inset shows the difference in species concentration in the last 10 seconds. Hard threshold condition applied on PINN for e) soft switch, f) sharp switch. g) PINN predicted switch for ReLU transformed output layer. X,Y-axis label for e, f, g are same.

Error following Eq. 9:

$$L_2 \text{ Error} = \sqrt{\sum_{i=1}^{n} (X_i - X_i')^2} \tag{9}$$

where X_i and X_i' are the true output level from ode15s and predicted output level from PINN for each input signal, respectively.

3. Results

3.1. Mechanisms that confer robust threshold

In an initial assessment of dynamically evolving fluctuations (Fig. 3a), the dimer and MEM mechanisms restrict fluctuations more effectively, as evidenced by subtracting the mean across 2000 trajectories to isolate mean fluctuations at each time point. Subsequently, across an ample parameter space (Appendix A.2), we calculated the coefficient of variation (CV) from 5 hrs of steady-state data. We found that the dimer and MEM mechanisms maintain a narrow CV range across target molecular counts (25 and 100 molecules), as in Fig. 3b. This suggests MEM's pronounced noise-mitigating ability and better reproducibility at the target threshold. Stochastic fluctuations drive the signaling level away from the mean μ, leading to diminished control of robustness. Precisely, for reliable functioning of biological processes (Wright & Mansour, 2003), threshold mechanisms that activate downstream processes must maintain mean signaling μ within an acceptable range ($\mu \pm 0.1\mu$). Failure leads to errors or process malfunctions (Mariani et al., 2008; Zinkle & Mohammadi, 2018), necessitating the deviation from threshold in a noisy environment to be under exquisite control.

Stochastic interrogation of steady-state data of the switch-forming mechanisms reveals the superiority of the MEM mechanism, as shown in Fig. 3c. Precisely, MEM restricts prolonged deviations and shows a comparatively high probability for shorter deviations (Fig. 3c), with the covalent mechanism performing equally well. Using Fourier analysis for the kinetics that yield the most probable CV, as in Fig. 3b, we assessed the frequency component of noise associated with achieving target thresholds. As observed from the progression of the energy spectral density (ESD) in Fig. 3d, the MEM and Covalent mechanisms intrinsically exhibit a substantial high-frequency component while restricting the prevalence of low-frequency noise. High-frequency noise arising from a molecular mechanism, if cascaded down to a relatively slow-timescale molecular process, averages out fluctuations (Lauffenburger & Linderman, 1996; Karim et al., 2012). Also, MEM mitigates uncertainty of trajectories, as quantified by the lowest Shannon entropy shown in Fig. 3e for every time point t_i. Though MEM dynamics were comparatively slower, the starting entropy immediately after $t = 0$ in all other mechanisms surpasses that of MEM and holds at steady state as well. Extending the uncertainty

to fidelity, measured as $JSD/H(P)$ given in Eq. 7, shows that MEM attains lower error (i.e. higher fidelity) compared to the rest.

3.2. Switch Performance : Information-theoretic assessment

Maintaining a switch state unchanged in the presence of stochastic fluctuations is crucial for the accurate functioning of biological processes. For instance, an active switch may be reset to an inactive state due to stochastic fluctuation, as shown using a box plot in Fig. 2b. Synthetic circuits used in drug design to control molecular switches must demonstrate extreme precision to avoid unintended resets. We shed light on molecular switches designed using MEM, Dimer, Covalent modification, and titration, focusing on their switching regions, from the sharpest to the softest responses (Fig. 2; see figures in Appendix A.4). A mean ultrasensitivity, calculated using Eq. 8 across five switch cases, shows that the dimer is superior and performs equally well when compared by slope, capturing the rate of change in the Shannon entropy for all switches in the switching region. As observed for the sharpest switch case, dimer, MEM, and Covalent switches demonstrate a steady slope of the line fitted for entropy progression, suggesting a maintenance of order in contrast to that of titration (blue lines in Fig. 2b, c), exhibiting a sharp increase in entropy in the switching region.

3.3. Robust and reliable switch cascade

Stochastic fluctuations can flip switch states spontaneously, affecting downstream processes when switches are in a network cascade. We design a set of hypothetical switch cascades ($S_{i,j}$, as in Fig. 1b) to assess whether robust control and greater reliability in a noisy environment require specific cascade architectures. Comparing with an activation-type switch cascade architecture in which receptor activation turns on a downstream covalent-modification switch, we show that the MEM mechanism, which turns on the switch, confers greater sensitivity, control, and uncertainty reduction. Precisely, the cascade comprising MEM at its outer layer achieves a sharper response ($\Delta Y/\Delta X = 1.38$) than the cascade S_{22} ($\Delta Y/\Delta X = 0.94$) using covalent modification in its outer layer. The S_{23} cascade exhibits a greater dynamic range (107 fold) than S_{22} (81 fold), but is amenable to even larger fold-increase through kinetic tuning via *NRE* formation (see Fig. 5a), inset), suggesting the possibility of greater control yet with higher sensitivity.

For stochastic trajectories (50 runs) of the rising part of the sigmoidal, the CV and Shannon entropy at each input point suggest qualitative performance difference when the MEM mechanism replaces (S_{23}) the input layer in the switch cascade S_{22}, designed in (Ghusinga et al., 2021). MEM at the initial layer introduces comparatively conservative fluctuations (Fig. 5b), measured as the CV, capable of minimizing unintended turn-on of the output layer. This may contribute to reliable downstream information transduction, which is necessary for signaling pathways and synthetic circuits. A comparison of the rate of change of entropy in switching region (SR) of S_{22} and S_{23} reveal that MEM (S_{23}) may help mitigate sharp entropy changes (relates to disorder), quantified in Fig. 5c by the slope of the line fits to the rate of entropy change. Also, insertion of MEM in S_{23} tends to restrict the maximum amplitude of the fluctuations of entropy at the output switch, as depicted in Fig. 5d.

3.4. PINN-based dimer switch

The PINN switch exhibits predictability of the output signal strength for a variation at the input end. Analysis reveals that as the number of hidden layers increase, the error between the predicted output and true output values generally decreases (see Fig. 4b). Upon selecting a layer size of 6 subsequent exploration searches for the ideal number of training points, optimizing time, memory usage and L_2 error. For instance, increasing the number of training points leads to improved accuracy (Fig. 4c).

With accurate ODE-based predictability of the dimer CRN via the PINN-based implementation, we extend it and design a PINN dimer switch over the input range X_{Tot}. As observed, the PINN-switch's concentration predictions deviate in the switching region, specifically the 'OFF' region, where the PINN's predictions extend linearly downward into physiologically unfeasible negative values (see Fig. 4e,f). Precisely, PINN predictions violate the physical mass-action properties of the CRN, which we absorb by enforcing a hard threshold on the switching species' concentration. We apply two different approaches: i) Forceful zero-adjustment to negative values, and ii) Restricting PINN's predictive domain using a hard constraint through the ReLU function. Forcefully converting negative values to 0 post-prediction leads to a comparable switch formation in Fig. 4e,f (pale yellow line). Alternatively, to address the negative concentration issue, we add a hard constraint in the form of a ReLU function at the neural network's output layer. The ReLU strictly maps output values in the range $[0, \infty)$, mitigating negative values for species concentration in the output (Fig. 4g).

4. Discussion

Ultrasensitivity and reliability in a noisy environment are critical performance criteria for a molecular switch, with noise playing a paradoxical role in context-specific ways (Eldar & Elowitz, 2010). Spurred by this, we designed switches based on MEM and dimer-formation mechanisms and compared them with switches based on covalent modification and titration mechanisms. Stochastic analysis reveals interesting differences among the mechanisms, with MEM

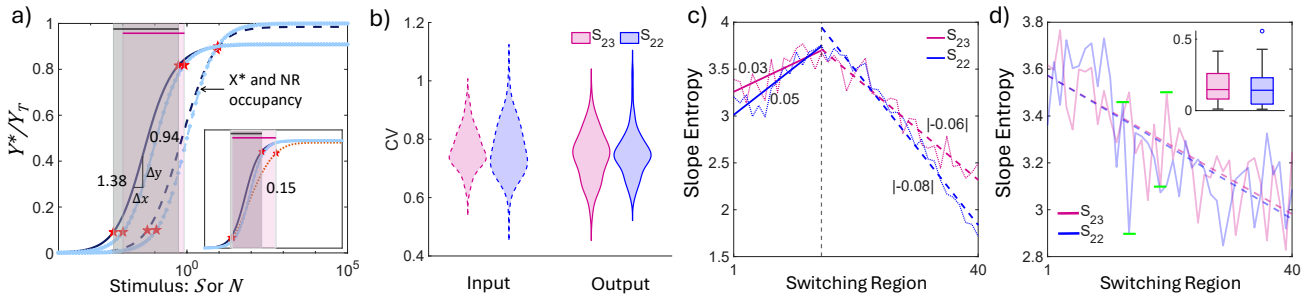

Figure 5. a) Switch-like role of from the switch cascades S_{22} (blue) and S_{23} (black), along with the dynamic range (shadow width) and sensitivity quantification (slope). Broken lines are the response to the stimulus, captured at the intermediate layers. b) Coeff. variation of S_{22}, S_{23}. c-d) Rate of entropy change and its amplitude comparison of S_{22}, S_{23}.

Table 2. Quantitative ranking of switch performance metrics: the mechanisms demonstrate context-specificity, but the proposed MEM and Dimer switches exhibit their strength– specifically, MEM as a mechanism for noise-resilience exhibits its potential. Here, 1 denotes superior performance, while 0 is inferior. We convert the values to a scale of [0 1] using the formulation $(X_i - \min(X))/(\max(X) - \min(X))$, where X is the vector of all values for a particular property for all switches. When lower values indicate better performance, we subtract each value from 1 to treat 1 as the best and 0 as the worst. The final switch-performance score is the average of all values column-wise. Here, $H(P)$ is Shannon Entropy, and SR is switching region.

Property	**Covalent**	**Titration**	**MEM**	**Dimer**	**Rank 1 and Rank 2**
Coeff. of Variation	0.365	0	0.312	1	Dimer and Covalent
Deviation Length	0.85	0.98	1	0	MEM and Titration
Frequency Analysis	1	0.5	1	0	MEM and Covalent
Entropy	0.431	0	0.333	1	Dimer and Covalent
JSD/H(P)	0.251	0	0.175	1	Dimer and Covalent
		Switch behavior			
Ultrasensitivity	0.591	0.769	0	1	Dimer and Titration
Slope of $H(P)$ in SR	0.897	0	0.666	1	Dimer and Covalent
Average Score	0.6	0.3	0.5	0.7	Dimer and Covalent

forming a robust threshold that appears competitive over different performance objectives. Also, MEM exhibits its potential for reliable information transduction in a switch cascade, demanding a thorough exploration of possible switch cascades comprising different switch mechanisms. Our ongoing work considers other combinations of mechanisms as switch cascades (see Fig. 1b) and currently pursues a more intensive in silico analysis of individual switches and their network cascades. Moreover, our initial PINN implementation of the dimeric switch has shown predictive abilities for variation at the input signal end. Switch predictability for customizations in dissociation constant (k_d), and multi-parameter PINNs that enable scaling and adjustment to switching location, are part of our ongoing effort. Finally, a robust PINN-guided agentic-AI-driven implementation of molecular switches, abiding by biological constraints and physical laws, aligns perfectly with the AIVC (AI Virtual Cell) (Bunne et al., 2024) initiative, which aims to develop a neural network framework to simulate diverse, dynamic cell behavior, on a multiscale, multimodal level at a faster pace, however, with necessary precision and accuracy in predictions.

References

Ahmadi Daryakenari, N., De Florio, M., Shukla, K., and Karniadakis, G. E. Ai-aristotle: A physics-informed framework for systems biology gray-box identification. *PLOS Computational Biology*, 20(3):e1011916, 2024a.

Ahmadi Daryakenari, N., De Florio, M., Shukla, K., and Karniadakis, G. E. Ai-aristotle: A physics-informed framework for systems biology gray-box identification. *PLOS Computational Biology*, 20(3):e1011916, 2024b.

Balemans, W. and Van Hul, W. Extracellular regulation of bmp signaling in vertebrates: a cocktail of modulators. *Developmental biology*, 250(2):231–250, 2002.

Bialek, W. and Setayeshgar, S. Physical limits to biochemical signaling. *Proceedings of the National Academy of Sciences*, 102(29):10040–10045, 2005.

Blake, W. J., Kærn, M., Cantor, C. R., and Collins, J. J. Noise in eukaryotic gene expression. *Nature*, 422(6932): 633–637, 2003.

Brewster, R. C., Weinert, F. M., Garcia, H. G., Song, D., Rydenfelt, M., and Phillips, R. The transcription factor

titration effect dictates level of gene expression. *Cell*, 156 (6):1312–1323, 2014.

Buchler, N. E. and Louis, M. Molecular titration and ultrasensitivity in regulatory networks. *Journal of molecular biology*, 384(5):1106–1119, 2008.

Bunne, C., Roohani, Y., Rosen, Y., Gupta, A., Zhang, X., Roed, M., Alexandrov, T., AlQuraishi, M., Brennan, P., Burkhardt, D. B., et al. How to build the virtual cell with artificial intelligence: Priorities and opportunities. *Cell*, 187(25):7045–7063, 2024.

Cai, H., Chen, T., Zhang, W., Yu, Y., and Wang, J. Efficient architecture search by network transformation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

Chen, Z., Lai, S.-K., and Yang, Z. At-pinn: Advanced time-marching physics-informed neural network for structural vibration analysis. *Thin-Walled Structures*, 196:111423, 2024.

Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F. Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022.

Dasgupta, T., Croll, D. H., Owen, J. A., Vander Heiden, M. G., Locasale, J. W., Alon, U., Cantley, L. C., and Gunawardena, J. A fundamental trade-off in covalent switching and its circumvention by enzyme bifunctionality in glucose homeostasis. *Journal of Biological Chemistry*, 289(19):13010–13025, 2014.

Deininger, M. W., Goldman, J. M., and Melo, J. V. The molecular biology of chronic myeloid leukemia. *Blood*, 96(10):3343–3356, 2000.

Del Olmo, M., Legewie, S., Brunner, M., Höfer, T., Kramer, A., Blüthgen, N., and Herzel, H. Network switches and their role in circadian clocks. *Journal of Biological Chemistry*, 300(5), 2024.

Eldar, A. and Elowitz, M. B. Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–173, 2010.

Elsken, T., Metzen, J. H., and Hutter, F. Neural architecture search: A survey. *Journal of Machine Learning Research*, 20(55):1–21, 2019.

Ferrell, J. E. and Ha, S. H. Ultrasensitivity part i: Michaelian responses and zero-order ultrasensitivity. *Trends in biochemical sciences*, 39(10):496–503, 2014.

Fishel, R. Mismatch repair, molecular switches, and signal transduction. *Genes & development*, 12(14):2096–2101, 1998.

Furriols, M. and Bray, S. A model notch response element detects suppressor of hairless–dependent molecular switch. *Current Biology*, 11(1):60–64, 2001.

Ghusinga, K. R., Jones, R. D., Jones, A. M., and Elston, T. C. Molecular switch architecture determines response properties of signaling pathways. *Proceedings of the National Academy of Sciences*, 118(11):e2013401118, 2021.

Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

Gillespie, D. T. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, 2007.

Goldbeter, A. and Koshland Jr, D. E. An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences*, 78(11):6840–6844, 1981.

Hu, B., Kessler, D. A., Rappel, W.-J., and Levine, H. Effects of input noise on a simple biochemical switch. *Physical review letters*, 107(14):148101, 2011.

Ji, W., Qiu, W., Shi, Z., Pan, S., and Deng, S. Stiff-pinn: Physics-informed neural network for stiff chemical kinetics. *The Journal of Physical Chemistry A*, 125(36): 8098–8106, 2021.

Karim, M. S., Buzzard, G. T., and Umulis, D. M. Secreted, receptor-associated bone morphogenetic protein regulators reduce stochastic noise intrinsic to many extracellular morphogen distributions. *Journal of The Royal Society Interface*, 9(70):1073–1083, 2012.

Karim, M. S., Madamanchi, A., Dutko, J. A., Mullins, M. C., and Umulis, D. M. Heterodimer-heterotetramer formation mediates enhanced sensor activity in a biophysical model for bmp signaling. *PLOS Computational Biology*, 17(9): e1009422, 2021.

Kho, J., Koh, W., Wong, J. C., Chiu, P.-H., and Ooi, C. C. Design of turing systems with physics-informed neural networks. In *2022 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1180–1186. IEEE, 2022.

Lauffenburger, D. A. and Linderman, J. J. *Receptors: models for binding, trafficking, and signaling*. Oxford University Press, 1996.

Lestas, I., Vinnicombe, G., and Paulsson, J. Fundamental limits on the suppression of molecular fluctuations. *Nature*, 467(7312):174–178, 2010.

Li, H., Cao, Y., Petzold, L. R., and Gillespie, D. T. Algorithms and software for stochastic simulation of biochemical reacting systems. *Biotechnology progress*, 24 (1):56–61, 2008.

Little, S. C. and Mullins, M. C. Bone morphogenetic protein heterodimers assemble heteromeric type i receptor complexes to pattern the dorsoventral axis. *Nature cell biology*, 11(5):637–643, 2009.

Liu, Y., Zhang, Y., Liu, J., and Ma, Z. Physics-informed neural networks for pharmacodynamic compartment models of animal cell culture. In *2025 19th International Conference on Complex Medical Engineering (CME)*, pp. 135–138. IEEE, 2025.

Lu, L., Meng, X., Mao, Z., and Karniadakis, G. E. Deepxde: A deep learning library for solving differential equations. *SIAM review*, 63(1):208–228, 2021.

Mariani, F. V., Ahn, C. P., and Martin, G. R. Genetic evidence that fgfs have an instructive role in limb proximal–distal patterning. *Nature*, 453(7193):401–405, 2008.

Matthew, S., Carter, F., Cooper, J., Dippel, M., Green, E., Hodges, S., Kidwell, M., Nickerson, D., Rumsey, B., Reeve, J., et al. Gillespy2: A biochemical modeling framework for simulation driven biological discovery. *Letters in biomathematics*, 10(1):87, 2023.

Miladinovic, D., Höppe, T., Chevalley, M., Georgiou, A., Stuart, L., Mehrjou, A., Bantscheff, M., Schölkopf, B., and Schwab, P. In silico biological discovery with large perturbation models. *Nature Computational Science*, pp. 1–12, 2025.

Mukherji, S., Ebert, M. S., Zheng, G. X., Tsang, J. S., Sharp, P. A., and Van Oudenaarden, A. Micrornas can generate thresholds in target gene expression. *Nature genetics*, 43 (9):854–859, 2011.

Ortega, F., Acerenza, L., Westerhoff, H. V., Mas, F., and Cascante, M. Product dependence and bifunctionality compromise the ultrasensitivity of signal transduction cascades. *Proceedings of the National Academy of Sciences*, 99(3):1170–1175, 2002.

Owen, J. A., Talla, P., Biddle, J. W., and Gunawardena, J. Thermodynamic bounds on ultrasensitivity in covalent switching. *Biophysical Journal*, 122(10):1833–1845, 2023.

Plesa, T., Zygalakis, K. C., Anderson, D. F., and Erban, R. Noise control for molecular computing. *Journal of the Royal Society Interface*, 15(144):20180199, 2018.

Praetorius, F., Leung, P. J., Tessmer, M. H., Broerman, A., Demakis, C., Dishman, A. F., Pillai, A., Idris, A., Juergens, D., Dauparas, J., et al. Design of stimulus-responsive two-state hinge proteins. *Science*, 381(6659): 754–760, 2023.

Raissi, M. Deep hidden physics models: Deep learning of nonlinear partial differential equations. *Journal of Machine Learning Research*, 19(25):1–24, 2018.

Raissi, M., Perdikaris, P., and Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

Raser, J. M. and O'shea, E. K. Noise in gene expression: origins, consequences, and control. *Science*, 309(5743): 2010–2013, 2005.

Rijal, K. and Mehta, P. A differentiable gillespie algorithm for simulating chemical kinetics, parameter estimation, and designing synthetic biological circuits. *ELife*, 14: RP103877, 2025.

Shimmi, O., Umulis, D., Othmer, H., and O'Connor, M. B. Facilitated transport of a dpp/scw heterodimer by sog/tsg leads to robust patterning of the drosophila blastoderm embryo. *Cell*, 120(6):873–886, 2005.

Simons, M., Gloy, J., Ganner, A., Bullerkotte, A., Bashkurov, M., Krönig, C., Schermer, B., Benzing, T., Cabello, O. A., Jenny, A., et al. Inversin, the gene product mutated in nephronophthisis type ii, functions as a molecular switch between wnt signaling pathways. *Nature genetics*, 37(5):537–543, 2005.

Thattai, M. and Van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 98(15):8614–8619, 2001.

Umulis, D., O'Connor, M. B., and Blair, S. S. The extracellular regulation of bone morphogenetic protein signaling. *Development*, 136(22):3715–3728, 2009.

Van Kampen, N. G. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

Weiss, A. and Attisano, L. The tgfbeta superfamily signaling pathway. *Wiley Interdisciplinary Reviews: Developmental Biology*, 2(1):47–63, 2013.

Wight, C. L. and Zhao, J. Solving allen-cahn and cahn-hilliard equations using the adaptive physics informed neural networks. *arXiv preprint arXiv:2007.04542*, 2020.

Wright, T. J. and Mansour, S. L. Fgf3 and fgf10 are required for mouse otic placode induction. 2003.

Zhai, W., Tao, D., and Bao, Y. Parameter estimation and modeling of nonlinear dynamical systems based on runge–kutta physics-informed neural network. *Nonlinear Dynamics*, 111(22):21117–21130, 2023.

Zhang, Q., Bhattacharya, S., and Andersen, M. E. Ultrasensitive response motifs: basic amplifiers in molecular signalling networks. *Open biology*, 3(4):130031, 2013.

Zinkle, A. and Mohammadi, M. A threshold model for receptor tyrosine kinase signaling specificity and cell fate determination. *F1000Research*, 7:F1000–Faculty, 2018.

A. Appendix

A.1. Models, Reactions and Derivations

A.1.1. MEM Chemical Reaction Network

$$\text{r1} : N + R \underset{k_{\text{on}}^{\text{NR}}}{\overset{k_{\text{off}}^{\text{NR}}}{\rightleftharpoons}} NR, \qquad \text{r2} : N + E \underset{k_{\text{on}}^{\text{NE}}}{\overset{k_{\text{off}}^{\text{NE}}}{\rightleftharpoons}} NE,$$

$$\text{r3} : NR + E \underset{k_{\text{on}}^{\text{NRE}}}{\overset{k_{\text{off}}^{\text{NRE}}}{\rightleftharpoons}} NRE, \qquad \text{r4} : NE + R \underset{k_{\text{on}}^{\text{NER}}}{\overset{k_{\text{off}}^{\text{NER}}}{\rightleftharpoons}} NRE,$$

$$\text{r5} : NR, NE, NRE \xrightarrow{\delta} \emptyset$$

$$[R_{\text{TOT}}] = [R] + [NR] + [NRE]$$

In the MEM reaction network, N is the signaling agent, R is the receptor and E is the exchanger molecule. The system of mass-action based ODEs derived from the reaction network are as follows.

$$\begin{aligned}
[\dot{NR}] &= k_{\text{on}}^{\text{NR}}[N][R] + k_{\text{off}}^{\text{NRE}}[NRE] - k_{\text{off}}^{\text{NR}}[NR] - k_{\text{on}}^{\text{NRE}}[NR][E] - \delta[NR] \\
[\dot{NE}] &= k_{\text{on}}^{\text{NE}}[N][E] + k_{\text{off}}^{\text{NER}}[NRE] - k_{\text{off}}^{\text{NE}}[NE] - k_{\text{on}}^{\text{NER}}[NE][R] - \delta[NE] \\
[\dot{NRE}] &= k_{\text{on}}^{\text{NRE}}[NR][E] + k_{\text{on}}^{\text{NER}} - k_{\text{off}}^{\text{NRE}}[NRE] - k_{\text{off}}^{\text{NER}}[NRE] - \delta[NRE]
\end{aligned}$$

For the switch derivation we start from the quasi-steady-state of $[NRE]$.

$$[\dot{NRE}] = 0 \Rightarrow [NRE] = \frac{k_{\text{on}}^{\text{NRE}}[NR][E] + k_{\text{on}}^{\text{NER}}[NE][R]]}{k_{\text{off}}^{\text{NRE}} + k_{\text{off}}^{\text{NER}} + \delta} \tag{10}$$

From the steady state $[\dot{NE}] = 0$ we get:

$$[NE] = \frac{k_{\text{on}}^{\text{NE}}[N] + \alpha k_{\text{on}}^{\text{NRE}}[NR]}{\gamma} \tag{11}$$

where α and γ are unitless parameters defined as

$$\alpha = \frac{k_{\text{off}}^{\text{NER}}}{k_{\text{off}}^{\text{NRE}} + k_{\text{off}}^{\text{NER}} + \delta} \qquad \text{and} \qquad \gamma = \frac{k_{\text{off}}^{\text{NE}} + \delta - [R](\alpha k_{\text{on}}^{\text{NER}} - k_{\text{on}}^{\text{NER}})}{[E]}$$

Simplifying further, we get the following expression:

$$[NE] = \lambda[N] + \mu[NR] \tag{12}$$

where λ and μ are unitless parameters defined as:

$$\lambda = k_{\text{on}}^{\text{NE}}/\gamma \qquad \text{and} \qquad \mu = (\alpha k_{\text{on}}^{\text{NRE}})/\gamma$$

Plugging in Eq. 10 and 12 into $[\dot{NR}] = 0$, we get the expression:

$$[NR] = \frac{k_{\text{on}}^{\text{NR}}[R] + \pi\lambda k_{\text{on}}^{\text{NER}}[R]}{\theta}$$

The units π and θ are:

$$\pi = \frac{k_{\text{off}}^{\text{NRE}}}{k_{\text{off}}^{\text{NRE}} + k_{\text{off}}^{\text{NER}} + \delta} \qquad \text{and} \qquad \theta = \frac{\delta + k_{\text{on}}^{\text{NRE}}[E] + k_{\text{off}}^{\text{NR}} - \pi\mu k_{\text{on}}^{\text{NER}}[R] - \pi k_{\text{on}}^{\text{NRE}}[E]}{[N]}$$

Simplifying further,

$$[NR] = \frac{k_{\text{on}}^{\text{NR}}}{\theta}[R] + \frac{\pi\lambda k_{\text{on}}^{\text{NER}}}{\theta}[R] = [R](\phi + \xi) \Rightarrow [R] = [NR]/(\phi + \xi) \tag{13}$$

Finally, substituting 13 into the conservation condition, we get the following expression for $[NR]$,

$$[NR] = \frac{-b + \sqrt{b^2 + 4a[R_{\text{TOT}}]}}{2a}$$

where a, and b are as follows:

$$a = \frac{k_{\text{on}}^{\text{NER}}\mu}{\Delta(\phi + \xi)}$$
$$b = \frac{1}{(\phi + \xi)} + 1 + \frac{k_{\text{on}}^{\text{NRE}}[E]}{\Delta} + \frac{k_{\text{on}}^{\text{NER}}\lambda[N]}{\Delta(\phi + \xi)}, \qquad \Delta = k_{\text{off}}^{\text{NRE}} + k_{\text{off}}^{\text{NER}} + \delta$$

A.1.2. Titration Chemical Reaction Network

$$\text{r1} : \emptyset \xrightarrow{\beta} r, \qquad \text{r2} : m + r \xrightarrow{k_{\text{on}}} c \qquad \text{r3} : c \xrightarrow{k_{\text{off}}} m + r,$$
$$\text{r4} : r \xrightarrow{\gamma_{\text{r}}} \emptyset \qquad \text{r5} : c \xrightarrow{\gamma_{\text{c}}} \emptyset$$

$$[m_{\text{TOT}}] = [m] + [c]$$

The species involved in the miRNA-mRNA CRN are m(miRNA), r (mRNA), and c (mRNA-miRNA complex). The reactions are represented as the following system of ODEs.

$$[\dot{r}] = \beta + k_{\text{off}}[c] - k_{\text{on}}[r][m] - \gamma_{\text{r}}[r]$$
$$[\dot{c}] = k_{\text{on}}[m][r] - k_{\text{off}}[c] - \gamma_{\text{c}}[c]$$

Solving for the steady state $[\dot{r}] = [\dot{c}] = 0$, we get the following expressions for $[c]$:

$$[c] = \frac{\beta - \gamma_{\text{r}}[r]}{\gamma_{\text{c}}} \qquad \text{and} \qquad [c] = \frac{k_{\text{on}}[r][m_{\text{TOT}}]}{k_{\text{off}} + \gamma_{\text{c}} + k_{\text{on}}[r]} \tag{14}$$

Defining two new parameters, and equating the expressions in Eq. 14:

$$[r]^2 + [r](\lambda + \theta - r_0) - \lambda r_0 = 0 \tag{15}$$

Finally, solving the quadratic equation in Eq. 15 we obtain the following expression for the Titration switch.

$$[r] = \frac{(r_0 - \lambda - \theta) + \sqrt{(r_0 - \lambda - \theta)^2 + 4\lambda r_0}}{2}$$

A.1.3. Dimer Chemical Reaction Network

$$\text{r1} : X + X \underset{k_1}{\overset{k_{-1}}{\rightleftharpoons}} XX, \qquad \text{r2} : X + Y \underset{k_2}{\overset{k_{-2}}{\rightleftharpoons}} XY, \qquad \text{r3} : Y + Y \underset{k_3}{\overset{k_{-3}}{\rightleftharpoons}} YY$$

$$[X_{\text{TOT}}] = [X] + 2[XX] + XY$$
$$[Y_{\text{TOT}}] = [Y] + 2[YY] + XY$$

X, Y are monomers, XX, YY are homodimers, and XY is a heterodimer. The following ODEs represent the Dimer reaction network.

$$\dot{[XX]} = k_1[X][X] - k_{-1}[XX]$$
$$\dot{[XY]} = k_2[X][Y] - k_{-2}[XY]$$
$$\dot{[YY]} = k_3[Y][Y] - k_{-3}[YY]$$

From the conservation conditions, we obtain the following expression.

$$2[XX] = [X_{\text{TOT}}] - [Y_{\text{TOT}}] - [X] + [Y] + 2[YY] \tag{16}$$

At steady state, $\dot{[XX]} = \dot{[XY]} = 0$. Following this, we have the expressions for $[X]^2$ and $[Y]$.

$$[X]^2 = k_{1\text{D}}[XX] \tag{17}$$
$$[Y] = \tfrac{k_{2\text{D}}[XY]}{[X]} \tag{18}$$

Here, $k_{1\text{D}}$ and $k_{2\text{D}}$ are the dissociation rate constants for for $[XX]$, $[XY]$.

We plug in Eq. 17 into the steady state $\dot{[YY]} = 0$.

$$[YY] = k_{3\text{A}}[Y]^2, \ (k_{3\text{A}} = k_3/k_{-3}) \tag{19}$$
$$[YY] = \gamma[XY]^2/[XX]^2, \ (\gamma = k_{3\text{A}} \cdot k_{2\text{D}}^2/k_{1\text{D}}) \tag{20}$$

Substituting the expression in Eq. 19 into Eq. 16, we get the quadratic expression as follows.

$$XX = \frac{-b + \sqrt{(b^2 + 4 \cdot a \cdot 2\gamma[XY]^2)}}{2a} \tag{21}$$

where a and b are:

$$a = 2 \qquad\qquad b = [Y_{\mathrm{TOT}}] - [X_{\mathrm{TOT}}] - [Y] + [X]$$

When $Y_{\mathrm{TOT}} = X_{\mathrm{TOT}}$, the threshold concentration of $[XX]$ is as follows.

$$[XX] = \frac{([Y] - [X]) + \sqrt{([Y] - [X])^2 + 16 \cdot \gamma \cdot [XY]^2}}{4}$$

A.1.4. Covalent Modification Cycle Reaction Network

In the following reactions, E and F are enzymes, S_0 is the unmodified substrate, S_1 is the modified substrate, ES and FS are enzymatic intermediates.

$$\mathrm{r1} : E + S_0 \underset{k_1}{\overset{k_{-1}}{\rightleftharpoons}} ES \qquad\qquad \mathrm{r2} : ES \xrightarrow{k_2} E + S_1$$

$$\mathrm{r3} : F + S_1 \underset{k_3}{\overset{k_{-3}}{\rightleftharpoons}} FS \qquad\qquad \mathrm{r4} : FS \xrightarrow{k_4} F + S_0$$

The mass-action based ODEs for the Covalent Modification cycle is shown.

$$\begin{aligned}
[\dot{ES}] &= k_1[S_0][E] - k_{-1}[ES] - k_2[ES] \\
[\dot{FS}] &= k_3[F][S_1] - k_{-3}[FS] - k_4[FS] \\
[\dot{S_0}] &= k_{-1}[ES] + k_4[FS] - k_1[E][S_0] \\
[\dot{S1}] &= k_2[ES] + k_{-3}[FS] - k_3[F][S_1]
\end{aligned}$$

A.1.5. Switch Cascade Models

The ODEs for the S_{22} switch cascade are as follows.

$$\begin{aligned}
[\dot{X^*}] &= k_1[S][X] - k_2[X^*] \\
[\dot{Y^*}] &= (k_3 + k_5[X^*])[Y] - k_4[Y^*]
\end{aligned}$$

The conservation conditions $X_{\mathrm{TOT}} = X + X^*$ and $Y_{\mathrm{TOT}} = Y + Y^*$ are applied.

The ODEs for the S_{23} switch cascade are as follows.

$$\begin{aligned}
[\dot{R}] &= k_{-1}[NR] + k_{-4}[NRE] - k_1[N][R] - k_4[NE][R] \\
[\dot{NR}] &= k_1[N][R] + k_{-3}[NRE] - k_{-1}[NR] - k_3[NR][E \\
[\dot{NE}] &= k_2[N][E] + k_{-4}[NRE] - k_{-2}[NE] - k_4[NE][R] \\
[\dot{NRE}] &= k_3[NR][E] + k_4[NE][R] - (k_{-3} + k_{-4})[NRE] \\
[\dot{Y}] &= k_{-6}[Y] - (k_6 + k_5[NR])[Y] \\
[\dot{Y^*}] &= (k_6 + k_5[NR])[Y] - k_{-6}[Y^*]
\end{aligned}$$

A.2. Kinetic Screen

We conducted an extensive parameter screen to search for the optimal kinetic set for each of the models. The exact parameter space and corresponding models are as follows: Titration (144), Dimer (729), MEM (3600) and Covalent Modification (729).

Algorithm 1 Kinetic Screen for Threshold Analysis

Require: parameter grid, ODE model, target concentration K, optimization function
Ensure: Species count $S_i \geq 0$
  for parameter set in grid **do**
    optimized parameters ← `fmincon`(optimization function, parameter set, K)
    Solve system of ODEs using K and optimized parameters
    $S \leftarrow \{S_1, S_2, \ldots, S_N\}$ ▷ Save the end point molecular count of each species
    model ← `reaction_model`(optimized parameters, S, parameter set) ▷ Simulate stochastic model with species initial count as steady state molecular count S
    $T \leftarrow$ 5 hour trajectory from SSA
    $CV \leftarrow \sigma_T/\mu_T$ ▷ Calculate coefficient of variation
  end for

Parameter	Values	Unit
k_{on}	[0.01, 0.1, 1]	$nM^{-1}s$
k_{off}	[0.01, 0.1, 1]	s^{-1}
γ_{r}	[0.001, 0.01, 0.1, 1]	s^{-1}
γ_{c}	[0.001, 0.01, 0.1, 1]	s^{-1}

Table 3. Kinetics used in screen for Titration model. The production rate of mRNA, β, is searched for a target concentration.

Parameter	Value	Unit
$k_{\text{on}}^{\text{NR}}$, $k_{\text{on}}^{\text{NE}}$	[0.001, 0.01]	$nM^{-1}s^{-1}$
$k_{\text{off}}^{\text{NR}}$, $k_{\text{off}}^{\text{NE}}$	[0.0001]	s^{-1}
$k_{\text{on}}^{\text{NRE}}$, $k_{\text{on}}^{\text{NER}}$	[0.0001, 0.001, 0.01, 0.1, 1, 10]	$nM^{-1}s^{-1}$
$k_{\text{off}}^{\text{NRE}}$, $k_{\text{off}}^{\text{NER}}$	[0.0001, 0.001, 0.01, 0.1, 1]	s^{-1}
δ	[0.0001]	s^{-1}

Table 4. Kinetics used in screen for MEM model.

Parameter	Value	Unit
k_1, k_2, k_3	[0.01, 0.1, 1]	$nM^{-1}s^{-1}$
k_{-1}, k_{-2}, k_{-3}	[0.01, 0.1, 1]	s^{-1}

Table 5. Kinetics used in screen for Dimer model.

Parameter	Value	Unit
k_1, k_3	[0.01, 0.1, 1]	$nM^{-1}s^{-1}$
k_{-1}, k_2, k_{-3}, k_4	[0.01, 0.1, 1]	s^{-1}

Table 6. Kinetics used in screen for Covalent Modification model.

A.3. Stochastic Simulation Methods

The Chemical Master Equation (CME) (Van Kampen, 1992) of the underlying continuous-time Markov chain of reaction system comprises of N chemical species $S = \{S_1, S_2, \ldots, S_N\}$ interacting in M $R = \{R_1, R_2, \ldots, R_M\}$ formulates as follows

$$\frac{dP(\mathbf{n}, t)}{dt} = \sum_{j=1}^{R} \left[a_j(\mathbf{n} - \boldsymbol{\nu}_j, t) P(\mathbf{n} - \boldsymbol{\nu}_j, t) - a_j(\mathbf{n}, t) P(\mathbf{n}, t) \right] \tag{22}$$

where $a_j(\mathbf{n}, t)$ is the propensity function and $\boldsymbol{\nu}_j$ is the stoichiometric vector. For a reaction network, the joint probability density function (PDF) that maps the evolution of the number of species over time is as follows.

$$\mathbf{P}(X_1, X_2, \ldots, X_N; t) \equiv \text{probability that any } S_i \text{ will have } X_i \text{ molecules at time } t \tag{23}$$

For each of our chemical reaction networks, we set up a stochastic model in the GillesPy2 (Matthew et al., 2023) module with transformed rate constants for stochastic adaptation. To approximate the CME, we used the available Gillespie's SSA (Gillespie, 1976; 2007) to under Stochkit (Li et al., 2008) module of the GillesPy2.

A.3.1. Unit Conversions

Order of Reaction	Deterministic Rate Law	Stochastic Transformation
zeroth order : $\emptyset \xrightarrow{k} X$	kMs^{-1}	$c = n_A k V$
first order : $X \xrightarrow{k} \cdot$	$k[X]Ms^{-1}$	$c = k$
second order : $X + Y \xrightarrow{k} \cdot$	$k[X][Y]Ms^{-1}$	$c = k/(n_A V)$

Table 7. Unit conversions used to convert deterministic reaction rate (k) to stochastic rate constant (c).

In Table 7, n_A is Avogadro's constant, 6.022×10^{23}, and V is the volume. To convert concentration based values to molecular units, we used a cylindrical cell volume of radius $r = 0.5\mu m$ and length $l = 2\mu m$. This gives us a scaling volume of

$$V = \frac{\pi}{2} 10^{-15}$$

In this volume $10nM$ concentration corresponds to $9.45(= 9)$ molecules.

A.4. Soft to Sharp Nonlinearity

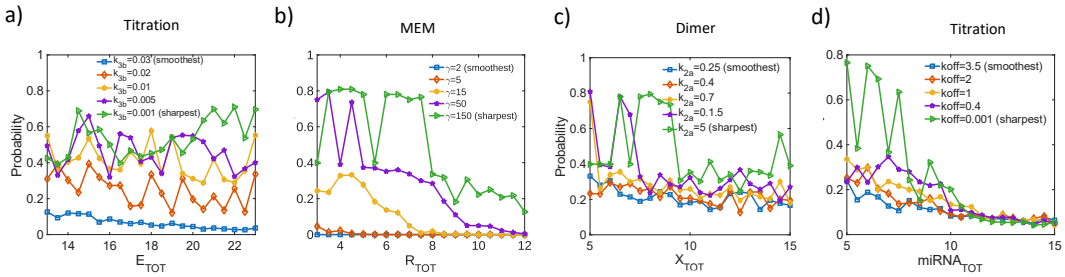

Figure 6. Probability measure of molecular count being equal to deterministic output in the switching region.

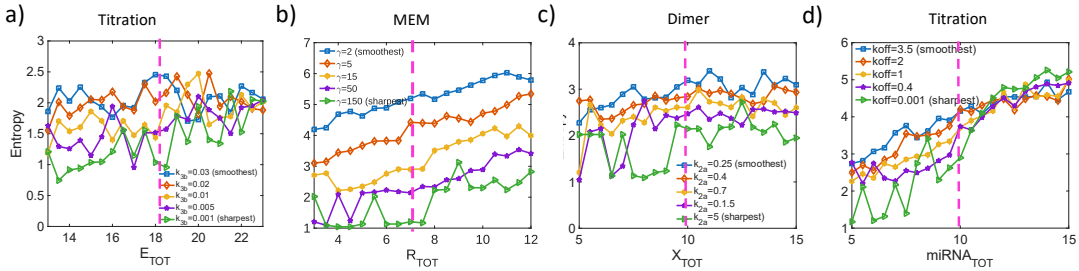

Figure 7. Entropy of the output in the switching region.

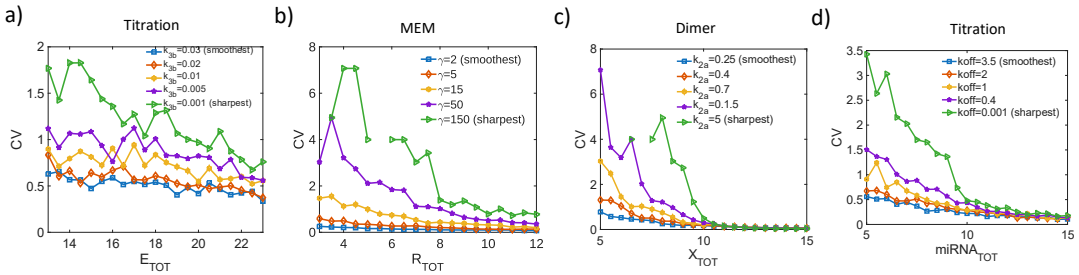

Figure 8. Coefficient of variation of the output in the switching region. Missing points correspond to zero mean.