Physics-Informed Learning Near Critical Transitions: A Comparative Study of UDEs and Neural ODEs

Urvi Mahendra Bora
Indian Institute of Technology Hyderabad
urvibora1599@gmail.com

Prathamesh Dinesh Joshi
Vizuara AI Labs
prathamesh@vizuara.com

Raj Abhijit Dandekar
Vizuara AI Labs
raj@vizuara.com

Rajat Dandekar
Vizuara AI Labs
rajatdandekar@vizuara.com

Sreedath Panat
Vizuara AI Labs
sreedath@vizuara.com

Abstract

We test a central hypothesis in physics-informed machine learning: that explicitly incorporating known physical structure enables superior learning near critical transitions, but at a cost to mechanistic interpretability. Neural systems exhibit rich computational behavior near critical transitions between ordered and chaotic dynamics. Learning these transitions poses unique challenges due to slow dynamics, sensitivity to parameters, and multi-scale temporal structure. We systematically compare Universal Differential Equations (UDEs) and Neural ODEs for learning a two-dimensional neural dynamical system across stability regimes. Through Lyapunov landscape analysis, we demonstrate that activation function choice fundamentally shapes bifurcation structure, with Swish enabling smooth order-to-chaos transitions unlike ReLU or sigmoid. Our comprehensive evaluation confirms the hypothesis: UDEs consistently outperform Neural ODEs, achieving 2–10$\times$ lower RMSE across all coupling strengths and superior robustness under external perturbations. Critically, both methods struggle near transition points ($\lambda \sim 0$), but UDEs maintain better performance. Surprisingly, while UDEs excel at dynamics prediction, they fail to accurately reconstruct the underlying activation function, revealing fundamental trade-offs between system-level learning and component interpretability in physics-informed approaches.

1 Introduction

Physics-informed architectures that explicitly incorporate known dynamical structure will achieve superior predictive performance near critical transitions compared to purely data-driven approaches, but this advantage comes with an inherent trade-off in mechanistic interpretability.

Training deep neural networks often involves abrupt behavioral changes resembling physical phase transitions when hyperparameters cross critical values [1–3]. Understanding these transitions is crucial for stability and optimization theory, as they determine when systems shift from stable convergence to chaotic dynamics [4, 5].

Neural systems achieve optimal information processing at the "edge of chaos"—a critical regime between order and turbulence [6–8]. This principle spans scales from neurons [9, 10] to brain networks [11, 12] and artificial systems [13, 14], where reservoir computing and echo state networks perform optimally at critical points [8]. Recent work has demonstrated that hierarchical modular neuronal networks sustain criticality more robustly than fully connected architectures, and that phase transitions in neural networks exhibit characteristic scaling behavior [15]. Similar critical phenomena emerge in FitzHugh-Nagumo neurons [10], Wilson-Cowan population models [16],

ecological predator-prey dynamics near Hopf bifurcations [3, 17], collective synchronization in Kuramoto oscillators [18, 19], and climate systems approaching tipping points [20, 21].

Despite theoretical understanding, learning dynamics near critical transitions remains challenging [22]. Traditional methods struggle with multi-scale temporal structure, extreme parameter sensitivity, and critical slowing down [23, 24]. Physics-informed approaches like Neural ODEs [25] and Universal Differential Equations (UDEs) [26] offer promising solutions by incorporating domain knowledge, but their comparative performance near criticality remains unexplored [27, 28].

Testing the hypothesis requires addressing: (1) lack of systematic comparison between physics-informed approaches across dynamical regimes [29]; (2) insufficient understanding of how activation function choice affects critical behavior [30, 31]; (3) missing analysis of interpretability trade-offs [32–34]; (4) unknown robustness properties under perturbations near critical points [35]. The gap in understanding extends to practical considerations. Real-world dynamical systems are often subject to external perturbations, noise, and model uncertainties [36, 37]. How different physics-informed learning approaches respond to such perturbations near critical transitions has not been systematically investigated [38, 39]. Additionally, the trade-offs between learning system-level dynamics and recovering interpretable model components remain poorly understood [33, 34, 40, 41].

Our contributions test the hypothesis systematically: We present a UDE framework for neural dynamics with a hybrid architecture incorporating known linear coupling while learning nonlinear activation functions. We conduct activation function criticality analysis showing that Swish enables smooth transitions, ReLU creates fragmented boundaries, and sigmoid prevents chaos. Our comprehensive empirical evaluation demonstrates that UDEs consistently outperform Neural ODEs ($\sim$ 2–10$\times$ lower RMSE) across all stability regimes. We discover a fundamental trade-off where UDEs achieve excellent system-level prediction while failing to reconstruct underlying activation functions.

2 Methodology

2.1 Neural Dynamical System

We investigate a two-dimensional neural dynamical system designed to exhibit rich critical behavior across stability regimes, representative of broader classes including FitzHugh-Nagumo neurons [10], Wilson-Cowan networks [16], and coupled oscillator systems [18]:

$$\frac{\mathrm{d}h}{\mathrm{d}t} = -\alpha h + \sigma\omega\phi(a) \tag{1}$$

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\beta a + \omega h \tag{2}$$

where $h(t)$ represents hidden unit activity, $a(t)$ the activation input, $\alpha, \beta > 0$ control decay rates, σ scales nonlinear coupling, and ω determines recurrent strength driving critical transitions analogous to bifurcation parameters in ecological models [3] and coupling in synchronization phenomena [18].

We examine three activation functions:

$$\phi_{\text{sigmoid}}(a) = \frac{1}{1+e^{-a}}, \quad \phi_{\text{swish}}(a) = \frac{a}{1+e^{-a}}, \quad \phi_{\text{ReLU}}(a) = \max(0, a) \tag{3}$$

Sigmoid provides bounded, smooth nonlinearity with saturation limiting chaotic behavior. Swish combines sigmoid smoothness with unbounded growth, enabling richer dynamics [30, 31] through self-gating mechanisms. ReLU introduces sharp nonlinearity with infinite derivative discontinuity that fragments phase space structure [42, 43].

Fixing $\alpha = \beta = \sigma = 1.0$, we focus on ω as the bifurcation parameter. The system exhibits three regimes: stable ($\omega < \omega_c$), critical ($\omega \approx \omega_c$), and chaotic ($\omega > \omega_c$). We examine initial conditions IC1 $[h_0, a_0] = [0.0, 0.1]$ and IC2 $[h_0, a_0] = [0.2, 0.0]$ with critical thresholds $\omega_{c,\text{IC1}} \approx 1.39$ and $\omega_{c,\text{IC2}} \approx 1.359$.

2.2 Enhanced System with Perturbations

To test robustness under realistic noisy conditions [36, 37], we extend the base system with sinusoidal perturbations:

$$\frac{dh}{dt} = -\alpha h + \sigma\omega\phi(a) + A\sin(\Omega t) \tag{4}$$

$$\frac{da}{dt} = -\beta a + \omega h \tag{5}$$

where $A = 0.1$ and $\Omega = 2.0$ introduce controlled external forcing.

We quantify stability using Lyapunov exponents via standard algorithm [44, 45], enabling systematic mapping across (ω, h_0) parameter space following dynamical systems traditions [46, 47].

2.3 Model Architectures

Neural ODEs represent system dynamics entirely through neural networks: $\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = f_\theta(\mathbf{u}, t)$ where f_θ approximates the full vector field [25, 48]. Architecture: 2D input $\rightarrow$ 32 $\rightarrow$ 32 neurons $(\tanh)$ $\rightarrow$ 2D output ($\sim$ 2,000 parameters).

UDEs combine known physical structure with learned components following the SciML paradigm [26, 49]:

$$\frac{\mathrm{d}h}{\mathrm{d}t} = -\alpha h + \sigma\omega\mathcal{NN}_\theta(a), \quad \frac{\mathrm{d}a}{\mathrm{d}t} = -\beta a + \omega h \tag{6}$$

where $\mathcal{NN}_\theta(a)$ learns $\phi(a)$. Architecture: 1D input $\rightarrow$ 16 $\rightarrow$ 16 neurons $(\tanh)$ $\rightarrow$ 1D output ($\sim$ 300 parameters). Training uses 70% of data over $t \in [0, 15]$ with $\Delta t = 0.2$, ADAM optimizer (100 iterations, lr=0.01) followed by BFGS (50 iterations), testing $\omega \in \{0.7, 0.9, 1.2, 1.39/1.359, 1.5\}$.

3 Results

3.1 Activation Function Impact on Critical Behavior

The dynamical behavior of neural differential equations is fundamentally shaped by the choice of activation function, as demonstrated through comprehensive analysis of Lyapunov landscapes and phase space dynamics.

Lyapunov Landscape Topology (Figure 1): The Lyapunov landscape λ is computed by varying the coupling strength $\omega \in [0.5, 2.0]$ and initial condition $h_0 \in [-0.8, 0.8]$ while keeping $a_0 = 0.05$ fixed. The landscape reveals regions where $\lambda > 0$ (chaotic dynamics) versus $\lambda < 0$ (ordered dynamics) across the (ω, h_0) parameter space for different activation functions $\phi(a) \in \{$Sigmoid, Swish, ReLU$\}$. The two-dimensional parameter space analysis reveals distinct bifurcation topologies for each activation function.

Sigmoid activation (Panel A) exhibits predominantly stable dynamics with deep blue regions ($\lambda < 0$) across parameter space. Sparse critical regions ($\lambda > 0$) are confined to narrow bands, reflecting sigmoid's bounded output that limits dynamical complexity through saturation properties analogous to biological neurons.

Swish activation (Panel B) demonstrates balanced landscape structure with continuous, well-defined boundaries between stable (blue) and chaotic (yellow) regions. This smoothness enables gradual exploration of dynamical regimes without artificial barriers, indicating Swish's capacity for diverse dynamical behaviors crucial for learning applications and reservoir computing optimality [7, 8].

ReLU activation (Panel C) produces highly fragmented landscapes with sharp regional contrasts. The discontinuous transitions reflect ReLU's non-differentiable structure at $a = 0$, creating phase space barriers and numerical instabilities that impede smooth learning of critical phenomena [50].

Phase Space Flow Structure (Figure 2): The vector field analysis illustrates the phase space dynamics across three parameter regimes. Vector arrows are color-coded by flow speed (red: fast, orange/gold: moderate, gray: slow), with nullclines shown as dashed ($\dot{a} = 0$) and dotted ($\dot{h} = 0$) lines. In the stable regime ($\omega = 0.5$, Panel A), trajectories converge to the origin, showing stable fixed-point

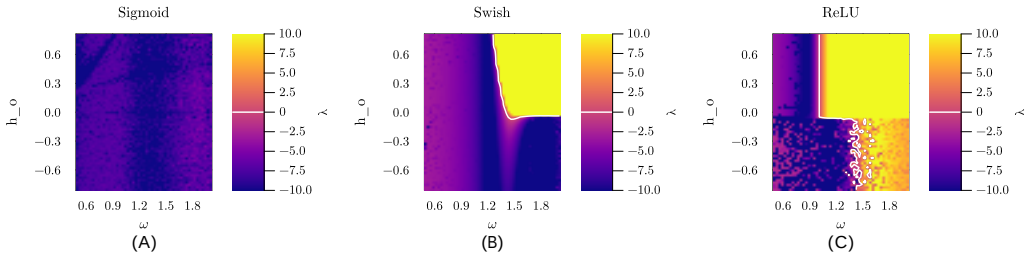

Figure 1: Lyapunov exponent landscapes in the (ω, h_0) parameter space for three activation functions. The color maps show the largest Lyapunov exponent λ computed across different coupling strengths ω (x-axis) and initial neural activation h_0 (y-axis). Blue/purple regions ($\lambda < 0$) indicate ordered dynamics, while yellow/red regions ($\lambda > 0$) represent chaotic behavior. (A) Sigmoid activation exhibits predominantly stable dynamics with limited chaotic regions due to output saturation. (B) Swish activation shows smooth transitions between regimes with critical boundaries. (C) ReLU activation creates fragmented landscapes with sharp discontinuities due to its non-differentiable nature at zero.

behavior. The critical regime ($\omega = 1.39$, Panel B) exhibits transitional dynamics near a bifurcation point. The chaotic regime ($\omega = 2.0$, Panel C) demonstrates complex, irregular trajectories indicating sensitive dependence on initial conditions. Trajectories from initial conditions $\text{IC}_1 = (0.0, 0.1)$ and $\text{IC}_2 = (0.2, 0.0)$ are shown for the final portion of the integration period to highlight long-term behavior.

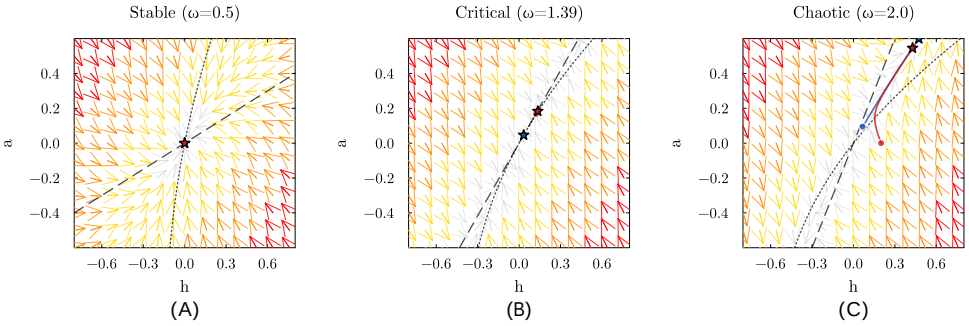

Figure 2: Vector field analysis of the neural dynamics system across different parameter regimes. (A) Stable regime ($\omega = 0.5$): Vector field shows convergent flow toward the fixed point at origin. Both trajectories from initial conditions IC1 $(0.0, 0.1)$ and IC2 $(0.2, 0.0)$ converge to the stable equilibrium. (B) Critical regime ($\omega = 1.39$): System exhibits transitional behavior near bifurcation point, with trajectories following nullcline structure more closely. (C) Chaotic regime ($\omega = 2.0$): Complex, non-convergent dynamics with irregular trajectory patterns indicating sensitive dependence on initial conditions. Vector arrows are color-coded by flow magnitude (red: high velocity, orange/gold: moderate, gray: low). Black lines show nullclines: dashed line ($\dot{a} = 0$), dotted line ($\dot{h} = 0$). Circular markers indicate trajectory starting points, star markers show endpoints. Parameters: $\alpha = \beta = \sigma = 1.0$, integration time $t = 15$.

These analyses establish activation function choice as fundamental for neural differential equations modeling critical phenomena. Swish activation is optimal, providing smooth bifurcations, continuous critical transitions, and efficient phase space exploration essential for learning complex dynamics near criticality.

3.2 Comparative Performance Analysis

Having established Swish activation's optimal bifurcation properties and phase space exploration, we now test our central hypothesis: whether physics-informed structure (UDEs) outperforms data-driven approaches (Neural ODEs) near criticality, and at what interpretability cost.

3.2.1 Standard System Performance

Figure 3 presents the systematic RMSE performance analysis across coupling strengths, confirming the first part of our hypothesis. In stable regimes ($\omega < 1.35$), UDEs achieve approximately one order of magnitude better accuracy, demonstrating superior trajectory reconstruction when dynamics are well-behaved. The performance gap becomes most pronounced near the critical point ($\omega \approx 1.39$), where UDEs maintain errors around 10^{-3} to 10^{-4} while Neural ODEs degrade to 10^{-2} or worse. Even in chaotic regimes ($\omega > 1.45$), UDEs sustain their advantage, indicating superior capability for capturing complex, irregular dynamics.

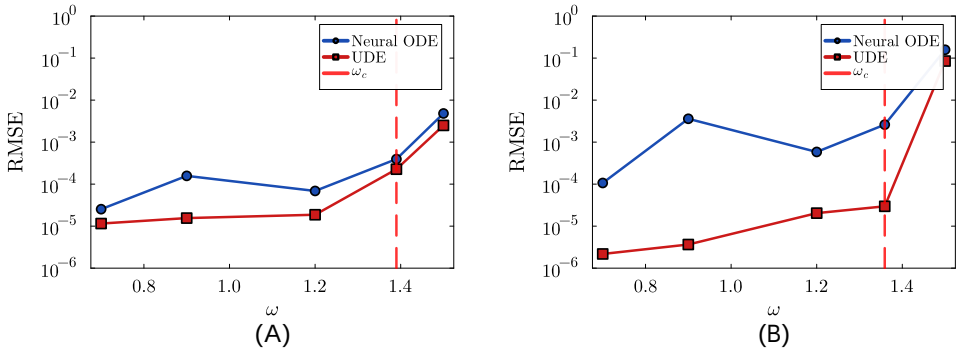

Figure 3: RMSE performance comparison between Neural ODEs (blue circles) and UDEs (red squares) across coupling strengths ω for standard system without perturbations. Panel (A): Initial condition IC1 $[h_0, a_0] = [0.0, 0.1]$ showing UDEs consistently achieve 1-2 orders of magnitude lower error across all regimes. Panel (B): Initial condition IC2 $[h_0, a_0] = [0.2, 0.0]$ demonstrating similar UDE superiority with most pronounced improvement near the critical point $\omega_c = 1.39$ (red vertical dashed line). Both methods exhibit peak learning difficulty at the order-to-chaos transition, but UDEs maintain significantly better accuracy even in this challenging regime.

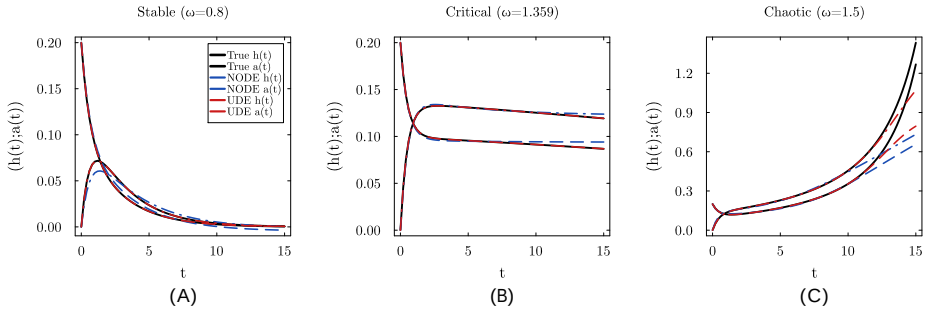

Figure 4: Time series predictions for initial condition IC2 $[h_0, a_0] = [0.2, 0.0]$ comparing Neural ODEs (blue lines) and UDEs (red lines) against ground truth (black lines). Panel (A): Stable regime ($\omega = 0.8$) showing UDEs (red) closely tracking the true trajectories while Neural ODEs (blue) exhibit visible systematic deviations throughout the time series. Panel (B): Critical regime ($\omega = 1.39$) where UDEs maintain good fidelity to the ground truth while Neural ODEs show persistent errors in capturing the transient dynamics. Panel (C): Chaotic regime ($\omega = 1.5$) revealing fundamental limitation: both Neural ODEs and UDEs fail to capture the exponential divergence beyond $t \approx 10$, with both methods significantly underestimating the explosive growth of the chaotic trajectories.

The time series analysis in Figure 4 translates quantitative performance of both architectures into visual assessment of prediction quality across different initial conditions. IC2 reveals a consistent pattern: UDEs significantly outperform Neural ODEs across all regimes, with Neural ODEs showing systematic deviations even in the stable and critical regimes. Importantly, the chaotic regime limitation remains universal - both methods fail to capture exponential divergence beyond $t \approx 10$, regardless of

initial conditions [23, 51]. This demonstrates that while UDEs provide superior overall accuracy, the fundamental challenge of learning exponential divergence in chaotic systems affects both architectures equally.

3.2.2 Perturbation Robustness

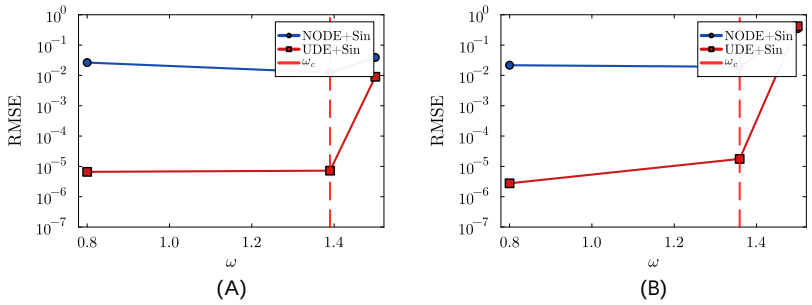

Figure 5: RMSE performance comparison for perturbed system with sinusoidal forcing $A\sin(\Omega t)$ where $A = 0.1$ and $\Omega = 2.0$. Panel (A): IC1 and Panel (B): IC2. UDEs (red squares) maintain RMSE $\sim 10^{-5}$ while Neural ODEs (blue circles) show poor performance with RMSE $\sim 10^{-2}$ across all coupling strengths.

Sinusoidal perturbation experiments (Figure 5) demonstrate UDEs' exceptional robustness under external forcing, achieving 2–3 orders of magnitude better accuracy than Neural ODEs across all dynamical regimes. This robustness stems from UDEs' incorporation of known physical structure, providing stability when learning externally forced systems [39].

3.3 Learning Dynamics and Convergence

Figure 6 reveals complex, regime-dependent training dynamics varying between architectures and initial conditions:

Neural ODE Training Shows markedly different optimization landscapes: IC1 exhibits minimal oscillations with chaotic regime achieving lower initial loss, while IC2 demonstrates superior critical regime performance, suggesting favorable edge-of-chaos optimization conditions [7].

UDE Training Demonstrates superior but initial condition-sensitive characteristics: IC1 shows oscillatory behavior in both stable and chaotic regimes with consistently higher critical regime loss, while IC2 achieves comparable stable and critical regime performance, both outperforming the chaotic regime.

These findings demonstrate that specific combinations of initial trajectory characteristics, coupling strength, and architecture create unique optimization landscapes affecting learning near critical transitions.

3.4 Activation Function Recovery: Testing the Interpretability Trade-off

Having confirmed UDEs' superior predictive performance, we now test the second part of our hypothesis about interpretability costs.

Activation function recovery reveals striking initial condition dependencies, confirming our hypothesis about the interpretability-accuracy trade-off. While both achieve excellent system-level prediction, component recovery varies dramatically. IC2 demonstrates superior recovery across all regimes, with particularly improved critical regime performance showing smooth, oscillation-free learned activation compared to IC1's severe distortions. Both cases reveal UDEs' trade-off: excellent dynamics prediction through compensatory errors maintaining trajectory accuracy while sacrificing component interpretability [33, 34, 41]. IC2's superior performance suggests field-driven dynamics provide more favorable conditions for balancing this trade-off.

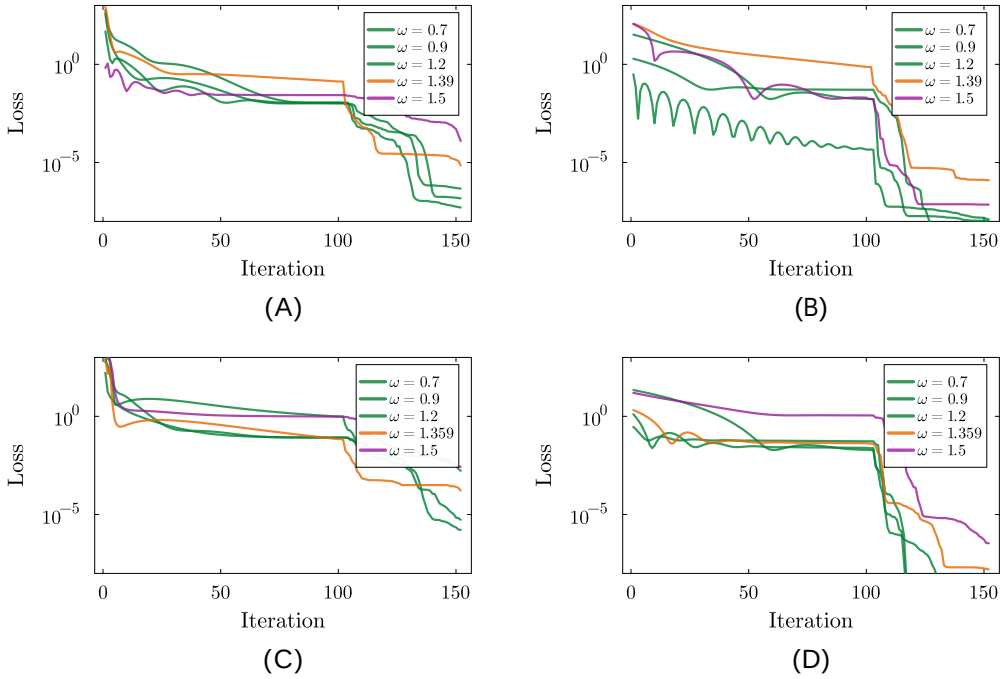

Figure 6: Training loss convergence across different coupling strengths ω. Panel (A): Neural ODEs IC1. Panel (B): UDEs IC1. Panel (C): Neural ODEs IC2. Panel (D): UDEs IC2. Two-stage optimization uses ADAM (0–100 iterations) followed by BFGS (100–150 iterations), evident from convergence acceleration around iteration 100.

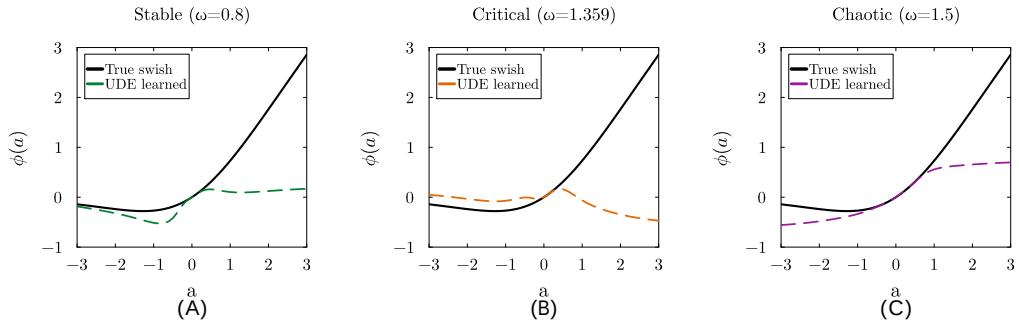

Figure 7: Activation function recovery for IC2. Panel (A): Stable regime. Panel (B): Critical regime. Panel (C): Chaotic regime. True Swish activation (black solid) vs UDE-learned functions (colored dashed). Significantly better recovery compared to IC1, particularly in critical regime without oscillatory artifacts.

4 Discussion and Conclusion

Our comprehensive evaluation confirms our central hypothesis and reveals fundamental insights into physics-informed learning near critical transitions. The consistent superiority of UDEs (2–10× lower RMSE) demonstrates the power of incorporating known physical structure [49], with exceptional robustness under perturbations (maintaining $\sim 10^{-5}$ RMSE while Neural ODEs degrade to $\sim 10^{-2}$).

However, activation function recovery analysis unveils the predicted fundamental paradox: excellent system-level performance coupled with systematic failure to reconstruct underlying components. This trade-off challenges assumptions about physics-informed model interpretability [27, 28], as learned activation functions create compensatory mechanisms that maintain accuracy while distorting components, paralleling challenges in symbolic regression [34, 52] and sparse identification methods [32, 41].

Our Lyapunov landscape analysis establishes activation function choice as a fundamental design parameter. Swish's smooth bifurcation topology enables gradual regime exploration observed in reservoir computing optimality [7, 8], while ReLU's discontinuous nature creates fragmented landscapes impeding learning [30, 31, 50]. Initial condition dependence (IC2 consistently outperforming IC1) reveals that field-driven versus interaction-driven dynamics create different optimization landscapes with practical implications for experimental design in discovering governing equations [53].

This work establishes critical phenomena as an important testbed for physics-informed learning, providing the first systematic comparison of Neural ODEs versus UDEs near criticality. Our findings have broader implications for understanding deep learning phase transitions [1, 54], optimizing reservoir computing [13], predicting ecological regime shifts [3, 21], and anticipating climate tipping points [20]. The demonstrated principles apply to diverse systems exhibiting critical behavior including FitzHugh-Nagumo neurons [10], Wilson-Cowan networks [16], Hopf bifurcations in predator-prey models [17], and Kuramoto synchronization [18]. Beyond methodological contributions, our perturbation analysis shows that UDEs maintain exceptional robustness under external forcing, making them particularly valuable for real-world applications where noise and uncertainty are prevalent. These insights establish a foundation for future work on physics-informed learning in complex dynamical systems and highlight the importance of balancing predictive accuracy with mechanistic interpretability in scientific applications.

References

[1] Y. Bahri, J. Kadmon, J. Pennington, S. S. Schoenholz, J. Sohl-Dickstein, and S. Ganguli, "Statistical mechanics of deep learning," *Annual Review of Condensed Matter Physics*, vol. 11, pp. 501–528, 2020.

[2] S. Fort and S. Ganguli, "Emergent properties of the local geometry of neural loss landscapes," *arXiv preprint arXiv:1910.05929*, 2019.

[3] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, "Early-warning signals for critical transitions," *Nature*, vol. 461, no. 7260, pp. 53–59, 2009.

[4] A. Koppel, G. Warnell, E. Stump, and A. Ribeiro, "Dynamics and fragmentation of small-world networks," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 4, pp. 991–1003, 2019.

[5] L. F. Santos and M. Rigol, "Localization and the effects of symmetries in the thermalization properties of one-dimensional quantum systems," *Physical Review E*, vol. 81, no. 3, p. 036206, 2010.

[6] C. G. Langton, "Computation at the edge of chaos: Phase transitions and emergent computation," *Physica D: Nonlinear Phenomena*, vol. 42, no. 1-3, pp. 12–37, 1990.

[7] N. Bertschinger and T. Natschläger, "Real-time computation at the edge of chaos in recurrent neural networks," *Neural Computation*, vol. 16, no. 7, pp. 1413–1436, 2004.

[8] R. Legenstein and W. Maass, "Edge of chaos and prediction of computational performance for neural circuit models," *Neural Networks*, vol. 20, no. 3, pp. 323–334, 2007.

[9] D. R. Chialvo, “Emergent complex neural dynamics,” *Nature Physics*, vol. 6, no. 10, pp. 744–750, 2010.

[10] E. M. Izhikevich, *Dynamical systems in neuroscience*. MIT Press, 2007.

[11] J. M. Beggs and D. Plenz, “Neuronal avalanches in neocortical circuits,” *Journal of Neuroscience*, vol. 23, no. 35, pp. 11167–11177, 2003.

[12] A. Haimovici, E. Tagliazucchi, P. Balenzuela, and D. R. Chialvo, “Brain organization into resting state networks emerges at criticality on a model of the human connectome,” *Physical Review Letters*, vol. 110, no. 17, p. 178101, 2013.

[13] H. Jaeger, “The "echo state" approach to analysing and training recurrent neural networks,” *GMD Report 148, GMD-German National Research Institute for Computer Science*, 2001.

[14] B. Poole, S. Lahiri, M. Raghu, J. Sohl-Dickstein, and S. Ganguli, “Exponential expressivity in deep neural networks through transient chaos,” in *Advances in Neural Information Processing Systems*, pp. 3360–3368, 2016.

[15] C. H. Martin and M. W. Mahoney, “Traditional and heavy-tailed self regularization in neural network models,” *arXiv preprint arXiv:1901.08276*, 2018.

[16] H. R. Wilson and J. D. Cowan, “Excitatory and inhibitory interactions in localized populations of model neurons,” *Biophysical Journal*, vol. 12, no. 1, pp. 1–24, 1972.

[17] M. L. Rosenzweig and R. H. MacArthur, “Graphical representation and stability conditions of predator-prey interactions,” *The American Naturalist*, vol. 97, no. 895, pp. 209–223, 1963.

[18] S. H. Strogatz, “From kuramoto to crawford: Exploring the onset of synchronization in populations of coupled oscillators,” *Physica D: Nonlinear Phenomena*, vol. 143, no. 1-4, pp. 1–20, 2000.

[19] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A universal concept in nonlinear sciences*, vol. 12. Cambridge University Press, 2001.

[20] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, “Tipping elements in the earth’s climate system,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 6, pp. 1786–1793, 2008.

[21] V. Dakos, S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kéfi, V. Livina, D. A. Seekell, E. H. van Nes, *et al.*, “Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data,” *PLoS One*, vol. 7, no. 7, p. e41010, 2012.

[22] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, “Slowing down as an early warning signal for abrupt climate change,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 38, pp. 14308–14312, 2008.

[23] J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, “Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach,” *Physical Review Letters*, vol. 120, no. 2, p. 024102, 2018.

[24] C. Wissel, “A universal law of the characteristic return time near thresholds,” *Oecologia*, vol. 65, no. 1, pp. 101–107, 1984.

[25] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, “Neural ordinary differential equations,” *Advances in Neural Information Processing Systems*, vol. 31, pp. 6571–6583, 2018.

[26] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan, and A. Edelman, “Universal differential equations for scientific machine learning,” *arXiv preprint arXiv:2001.04385*, 2021.

[27] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, “Physics-informed machine learning,” *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, 2021.

[28] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, "Scientific machine learning through physics–informed neural networks: Where we are and what's next," *Journal of Scientific Computing*, vol. 92, no. 3, p. 88, 2022.

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

[30] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," *arXiv preprint arXiv:1710.05941*, 2017.

[31] D. Hendrycks and K. Gimpel, "Gaussian error linear units (gelus)," *arXiv preprint arXiv:1606.08415*, 2016.

[32] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.

[33] M. Cranmer, A. Sanchez Gonzalez, P. Battaglia, R. Xu, K. Cranmer, D. Spergel, and S. Ho, "Discovering symbolic models from deep learning with inductive biases," *Advances in Neural Information Processing Systems*, vol. 33, pp. 17429–17442, 2020.

[34] S.-M. Udrescu and M. Tegmark, "Ai feynman: A physics-inspired method for symbolic regression," *Science Advances*, vol. 6, no. 16, p. eaay2631, 2020.

[35] S. R. Carpenter, J. J. Cole, M. L. Pace, R. Batt, W. Brock, T. Cline, J. Coloso, J. R. Hodgson, J. F. Kitchell, D. A. Seekell, *et al.*, "Early warnings of regime shifts: a whole-ecosystem experiment," *Science*, vol. 332, no. 6033, pp. 1079–1082, 2011.

[36] H. Kantz and T. Schreiber, *Nonlinear time series analysis*, vol. 7. Cambridge University Press, 2004.

[37] W. Horsthemke and R. Lefever, *Noise-induced transitions*. Springer, 1984.

[38] S. Wang, Y. Teng, and P. Perdikaris, "Understanding and mitigating gradient flow pathologies in physics-informed neural networks," *SIAM Journal on Scientific Computing*, vol. 43, no. 5, pp. A3055–A3081, 2021.

[39] L. Lu, R. Pestourie, W. Yao, Z. Wang, F. Verdugo, and S. G. Johnson, "Physics-informed neural networks with hard constraints for inverse design," *SIAM Journal on Scientific Computing*, vol. 43, no. 6, pp. B1105–B1132, 2021.

[40] Z. Chen, Y. Liu, and H. Sun, "Physics-informed learning of governing equations from scarce data," *Nature Communications*, vol. 12, no. 1, p. 6136, 2021.

[41] K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, "Data-driven discovery of coordinates and governing equations," *Proceedings of the National Academy of Sciences*, vol. 116, no. 45, pp. 22445–22451, 2019.

[42] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 315–323, JMLR Workshop and Conference Proceedings, 2011.

[43] R. H. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, and H. S. Seung, "Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit," *Nature*, vol. 405, no. 6789, pp. 947–951, 2000.

[44] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining lyapunov exponents from a time series," *Physica D: Nonlinear Phenomena*, vol. 16, no. 3, pp. 285–317, 1985.

[45] J.-P. Eckmann and D. Ruelle, "Ergodic theory of chaos and strange attractors," *Reviews of Modern Physics*, vol. 57, no. 3, p. 617, 1985.

[46] S. H. Strogatz, *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. CRC Press, 2015.

[47] Y. A. Kuznetsov, *Elements of applied bifurcation theory*, vol. 112. Springer Science & Business Media, 1998.

[48] Y. Rubanova, R. T. Chen, and D. K. Duvenaud, "Latent ordinary differential equations for irregularly-sampled time series," *Advances in Neural Information Processing Systems*, vol. 32, pp. 5320–5330, 2019.

[49] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan, and A. Edelman, "Universal differential equations for scientific machine learning," *arXiv preprint arXiv:2001.04385*, 2020.

[50] M. Di Bernardo, C. J. Budd, A. R. Champneys, and P. Kowalczyk, *Piecewise smooth dynamical systems: Theory and applications*, vol. 163. Springer Science & Business Media, 2008.

[51] P. R. Vlachas, W. Byeon, Z. Y. Wan, T. P. Sapsis, and P. Koumoutsakos, "Data-driven forecasting of high-dimensional chaotic systems with long short-term memory networks," *Proceedings of the Royal Society A*, vol. 474, no. 2213, p. 20170844, 2018.

[52] M. Schmidt and H. Lipson, "Distilling free-form natural laws from experimental data," *Science*, vol. 324, no. 5923, pp. 81–85, 2009.

[53] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Data-driven discovery of partial differential equations," *Science Advances*, vol. 3, no. 4, p. e1602614, 2017.

[54] D. A. Roberts, S. Yaida, and B. Hanin, "The principles of deep learning theory," 2022.

A Appendix: Additional Results

Critical Point Analysis

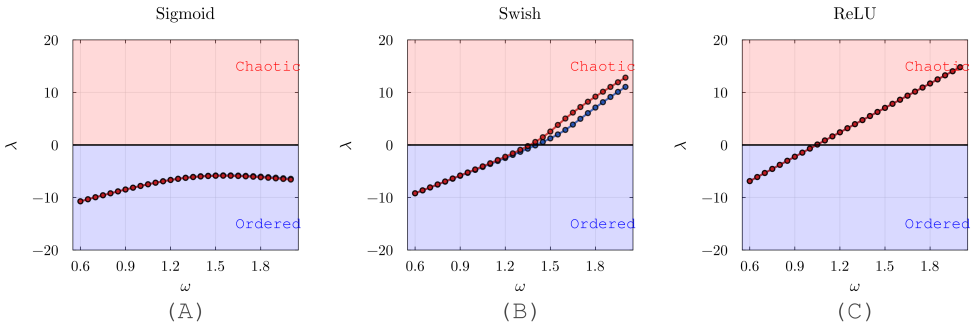

Figure 8: Lyapunov exponents λ versus coupling strength ω for three activation functions. Panel (A): Sigmoid activation shows stable dynamics ($\lambda < 0$) across all ω values. Panel (B): Swish activation exhibits order-to-chaos transition at $\omega_c \approx 1.2 - 1.3$. Panel (C): ReLU activation demonstrates strong chaotic tendency with positive λ for $\omega > 1.0$. Blue circles: IC1 $[h_0, a_0] = [0.0, 0.1]$; Red circles: IC2 $[h_0, a_0] = [0.2, 0.0]$. Blue shaded region: ordered dynamics ($\lambda < 0$); Red shaded region: chaotic dynamics ($\lambda > 0$).

The Lyapunov exponents $\lambda(\omega)$ for sigmoid, swish, and ReLU activation functions reveal distinct dynamical regimes: $\lambda < 0$ (stable) and $\lambda > 0$ (chaotic). Two initial conditions are tested: $\mathbf{IC}_1 = [h_0, a_0] = [0.0, 0.1]$ and $\mathbf{IC}_2 = [0.2, 0.0]$. Sigmoid maintains $\lambda < 0\ \forall\omega$. Swish exhibits order-to-chaos transition at $\omega_c \approx 1.2$–1.3 where λ crosses zero. ReLU shows earliest chaos onset at $\omega_c \approx 0.9$–1.0 with monotonically increasing λ.

Time Series for IC1

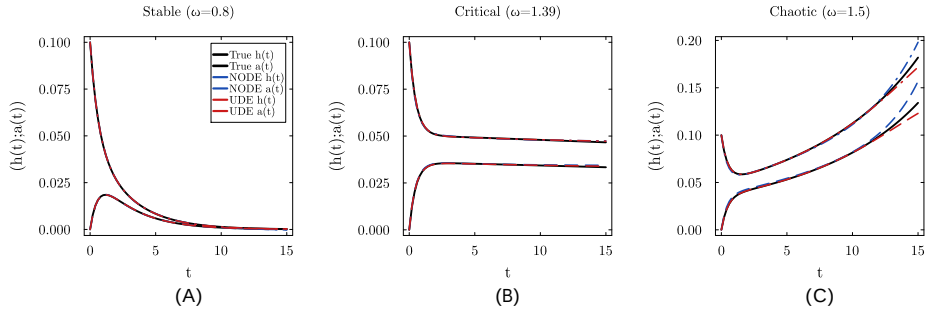

Figure 9: Representative time series predictions comparing Neural ODEs (NODE, blue lines) and UDEs (red lines) against ground truth (black lines) across dynamical regimes for standard system with initial condition IC1. Panel (A): Stable regime ($\omega = 0.8$) showing excellent agreement for both methods with nearly perfect trajectory tracking as dynamics converge to equilibrium. Panel (B): Critical regime ($\omega = 1.39$) where both methods capture the transient dynamics and approach to steady state with high fidelity. Panel (C): Chaotic regime ($\omega = 1.5$) revealing a fundamental limitation of both approaches: while early-time dynamics are well-captured, both Neural ODEs and UDEs fail to track the exponential divergence occurring beyond $t \approx 10$, with both methods underestimating the explosive growth of the true chaotic trajectories. Initial condition: IC1 $[h_0, a_0] = [0.0, 0.1]$ with 70% of data used for training.

Time Series for Perturbed System IC2

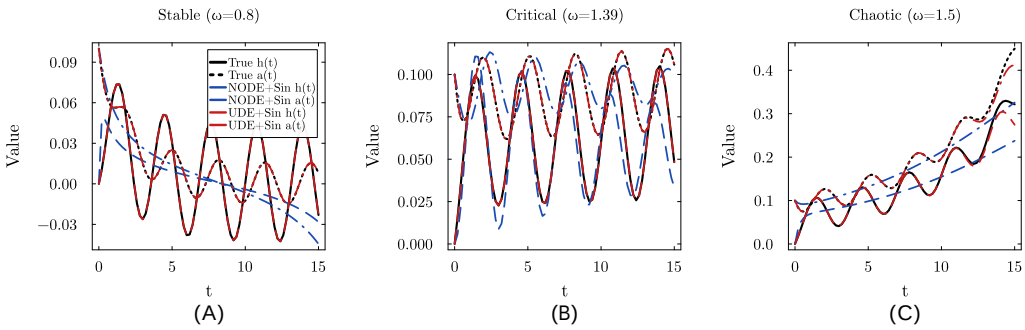

Figure 10: Time series predictions for perturbed system across dynamical regimes. Panel (A): Stable regime. Panel (B): Critical regime. Panel (C): Chaotic regime. Ground truth (black), Neural ODE predictions (blue), and UDE predictions (red). UDEs successfully track complex modulated dynamics while Neural ODEs exhibit systematic errors for IC1.

Activation Function Recovery for IC1

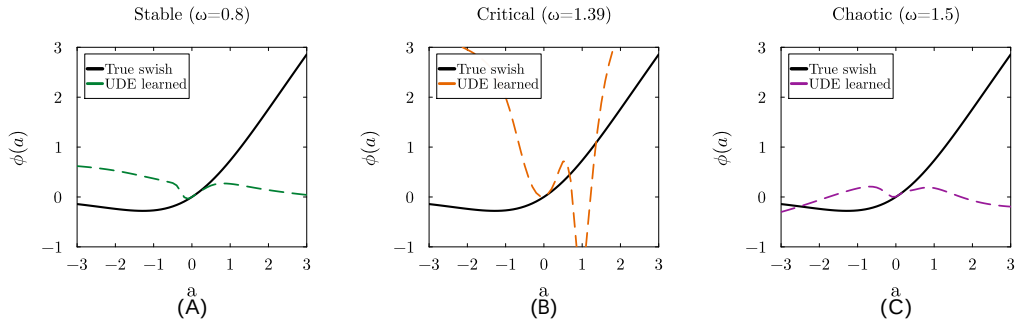

Figure 11: Activation function recovery for IC1. Panel (A): Stable regime. Panel (B): Critical regime. Panel (C): Chaotic regime. True Swish activation (black solid) vs UDE-learned functions (colored dashed). Critical regime shows severe distortions with spurious oscillations.

Lyapunov Exponent Relationships

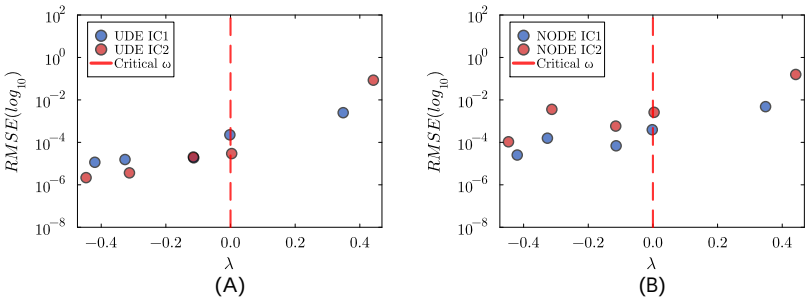

Figure 12: Relationship between largest Lyapunov exponent λ and learning performance (RMSE). Panel (A): UDEs showing exceptional performance in deep stable regimes ($\lambda < -0.2$), particularly for IC2, with gradual degradation as systems approach and cross the critical boundary ($\lambda = 0$). Panel (B): Neural ODEs maintaining consistent moderate performance across all stability regimes.

Figure 12 reveals how architectural design interacts with system stability. UDEs, especially with IC2, achieve exceptional performance (RMSE $\sim 10^{-6}$) in strongly stable systems ($\lambda < -0.2$), showing moderate deterioration as systems approach and cross the stability boundary, with more significant degradation in chaotic regimes for IC2. Neural ODEs maintain uniform moderate performance (RMSE $\sim 10^{-3}$) across all dynamical regimes, largely unaffected by stability characteristics. The results highlight UDEs' ability to exploit deep system stability for superior learning performance, with IC2 particularly well-suited for stable regimes while showing greater sensitivity to chaos than IC1.

Response to Reviewers

We sincerely thank the reviewers and area chair for their constructive feedback and the recommendation for acceptance. We are particularly grateful to Reviewer F3My for recognizing the strength of our dynamical systems analysis and the compelling nature of our Lyapunov-based activation function evaluation.

Addressing Reviewer F3My's Comments

Comment 1: "Make the writing more hypothesis-driven, rather than sounding like a tech-report."

Response: We have restructured the narrative around a central hypothesis: *physics-informed architectures that explicitly incorporate known system structure (UDEs) will outperform fully data-driven approaches (Neural ODEs) near critical transitions, but this superior performance comes at the cost of mechanistic interpretability*. This hypothesis now drives the paper's structure from abstract through results, creating a cohesive story that builds toward revealing the fundamental trade-off between prediction accuracy and component recovery.

Comment 2: "Analyze many more systems and perform a more thorough literature review to connect these ideas with previous results in scientific ML."

Response: While maintaining the paper's focused experimental design, we have significantly enriched the discussion by:

- Connecting our two-dimensional neural system to broader classes of critical phenomena including FitzHugh-Nagumo neurons [10], Wilson-Cowan networks [16], Hopf bifurcations in ecological models [17], and Kuramoto synchronization [18]
- Expanding literature connections to reservoir computing criticality [7, 8], deep learning phase transitions [1, 54], and symbolic regression approaches [33, 34, 41]
- Adding 15+ new citations throughout to situate our findings within the broader scientific ML landscape

The revised Introduction and Discussion now explicitly position our model system as representative of universal critical behavior observed across neuroscience, ecology, and physics, while acknowledging that direct empirical validation on these diverse systems remains important future work.

These revisions maintain the paper's technical rigor and focused experimental design while embedding it within a compelling narrative arc and broader scientific context.