Exploring Learnability in Dynamical Stochastic Networks: A Field-Theoretic Approach

Yibo Jacky Zhang YIBOZ@STANFORD.EDU and **Sanmi Koyejo** SANMI@STANFORD.EDU
Computer Science Department, Stanford University, USA

Editors: List of editors' names

Abstract

A persistent puzzle appears across multiple fields, yet its solution continues to elude full understanding. How can a network of simple nodes, each evolving with only local information and learning with local rules, collectively solve complex global tasks? Such *dynamical stochastic networks* generalize cellular automata and recurrent neural networks, model biological circuits, and can be interpreted as decentralized multi-agent systems. We identify three fundamental challenges in the efficient learning of dynamical stochastic networks: (1) constructing precise yet easy-to-use theoretical models; (2) designing mechanisms for local credit assignment aligned with global objectives; and (3) characterizing the regimes of configurations that enable efficient learning. To address these issues, we adopt a theoretical framework of objective-driven dynamical stochastic fields, referred to as the *intelligent field*, and propose theoretical quantities that capture learnability. Crucially, we show that efficient learning emerges when systems maximize their ability to retain information over time. Experiments demonstrate that local information retention is linked to global learnability, shedding light on the future practical design of effective dynamical stochastic networks.

Keywords: Dynamical Stochastic Networks, Learnability, Intelligent Field.

1. Introduction

A *dynamical stochastic network* can be abstracted as follows. The system is composed of many interacting nodes, each with its own hidden state and capable of exchanging signals with neighboring nodes. The local evolution at each node depends solely on its local observable information. Additionally, the system may be influenced by an external environment, which can either perturb the system or extract information from it.

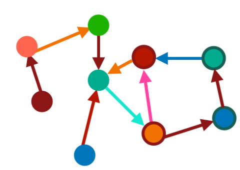


Figure 1: Illustration of a Dynamical Stochastic Net.

To make this less abstract, imagine a binary dynamical stochastic network, arranged in a peculiar configuration, reminiscent of a neural network, attempting to play blindfold chess in discrete time steps. At every move, the opponent's action is encoded as a binary pattern injected into designated input nodes. The network then evolves for a few internal steps, where each binary node updates its state based solely on the states of its local neighbors. A readout mechanism then accesses a small subset of nodes to decode the network's response move. Over time, the recurrent dynamics act as a form of memory, enabling the system to maintain a latent representation of the chessboard using only the sequential stream of moves.

Indeed, this abstract model of dynamical stochastic networks is shown to be capable of complex behaviors across disciplines. From a computer science perspective, its deterministic counterpart, cellular automata, has been shown to be Turing complete (Rendell, 2002; Cook et al., 2004), a property also shared by recurrent networks (Funahashi and Nakamura, 1993; Siegelmann and Sontag, 1992). Leveraging this computational power, researchers have extended the idea to stochastic artificial neural networks (Williams, 1992; Hopfield, 1982). From a neuroscience standpoint, locally coupled, recurrent dynamics are a well-supported account of computation in biological circuits (Maass, 1997; Mante et al., 2013). Finally, the same system can be viewed as a decentralized partially observable multi-agent system (Oliehoek et al., 2016; Jin et al., 2024; Zhang et al., 2018; Omidshafiei et al., 2017): local update rules correspond to agents acting on partial observations.

However, *efficient learning* of a dynamical stochastic network poses a substantial puzzle.

> ***How can individual nodes, operating on simple local rules, efficiently learn behaviors such that collectively they achieve a complex global objective?***

Important gaps remain in our understanding of dynamical stochastic networks. We identify three main challenges in understanding and designing such systems.

- **Challenge 1: Theoretical Modeling.** Choosing the right theoretical model is essential. It should be both foundational and easy to use for theoretical analysis while offering precise and concrete descriptions for algorithmic development.
- **Challenge 2: Credit Assignment.** The global objective should be decomposed and propagated within the network. Due to locality, credit assignment must be conveyed through the local propagation of reward or penalty signals, referred to as *objective signals*. With the existence of recurrent connections, each node both sends objective signals to others and adapts its behavior based on what it receives. This mutual influence creates complex feedback loops that largely complicate the system's dynamics.
- **Challenge 3: Regimes of Learnability.** Many system configurations can lead to inefficient learning when nodes rely only on local observations. Consider the following illustrative example: each node receives signals from its neighbors (and possibly the environment) and decides whether to propagate a signal onward. Such a system may operate in two regimes: (1) nodes are overly sensitive and fire signals chaotically, or (2) nodes fire too sparsely and remain largely unresponsive. In both cases, learning becomes ineffective for obvious reasons. What remains unclear, however, is the fundamental quantity that governs learnability in dynamical stochastic networks more generally. Identifying and exploiting the regimes that enable efficient learning is therefore another central challenge.

In this paper, we first show how a dynamical stochastic network evolving in continuous time can be formulated under the intelligent field framework (Zhang and Koyejo, 2025), where credit assignment is captured through objective propagation. We then introduce theoretical quantities that characterize learnability and, crucially, show that *efficient learning emerges when systems maximize their ability to retain information over time.* Based on this, we explore how local information retention gives rise to global learnability. Experiments demonstrate our theoretical insights and point toward the future design of scalable dynamical stochastic networks.

2. Theoretical Analysis

Throughout this section, we assume that the entire stochastic system of the network, which evolves in continuous time, is ergodic and is defined on a finite configuration space. A detailed overview of the intelligent field framework (Zhang and Koyejo, 2025) and complete derivations of the theoretical results are provided in the Appendix A&B, respectively.

2.1. The Intelligent Field Framework

We model the network as a space $\mathcal{X}$, with each point $x \in \mathcal{X}$ representing a node. The system dynamics are described by a field $\omega(t,x) \in \Omega_x$ evolving stochastically in spacetime, where Ω_x is a finite set, and the global configuration set is $\Omega = \prod_x \Omega_x$. The theory tracks the exact dynamics of the probability distribution over configurations, denoted by vectors $|\varphi(t)\rangle \in \mathcal{H}$. A pure state concentrated on configuration $\omega \in \Omega$ is written as $|\omega\rangle \in \mathcal{H}$. The evolution of $|\varphi(t)\rangle$ is generated by a linear operator $\mathbf{G} : \mathcal{H} \to \mathcal{H}$, i.e., its infinitesimal generator, where locality implies that $\mathbf{G}$ decomposes into a sum of local generators $\mathbf{G}(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}|\varphi(t)\rangle = \mathbf{G}|\varphi(t)\rangle, \qquad \mathbf{G} = \sum_x \mathbf{G}(x).$$

As expected from locality, the local generators commute, i.e., $[\mathbf{G}(x), \mathbf{G}(x')] = 0$, for non-neighboring x, x', and each node's behavior is fully determined by its local generator.

To model how the system receives objective signals such as penalties and rewards, a higher-dimensional space $\widetilde{\mathcal{H}} := \mathcal{H} \otimes \mathcal{H}$ is introduced, where it has basis vectors in the form of $|\omega'\omega\rangle$. The generator $\mathbf{G}$ is lifted to $\widetilde{\mathbf{G}} : \mathcal{H} \to \widetilde{\mathcal{H}}$. Each local node x is then associated with an objective operator $\mathbf{\Gamma}_x : \widetilde{\mathcal{H}} \to \mathbb{R}$, and is evolving to minimize its objective value: the time average of the objective signals it receives. The objective value $\bar{\gamma}(x)$ and its gradient with respect to a local generator $\mathbf{G}(x)$ are given by

$$\bar{\gamma}(x) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}t\ \mathbf{\Gamma}_x \widetilde{\mathbf{G}}|\varphi(t)\rangle, \qquad \frac{\partial \bar{\gamma}(x)}{\partial G(x)^{\omega'}_{\omega}} = \mathbf{\Gamma}_x(1 + \widetilde{\mathbf{G}}\mathbf{S}\mathbf{\Pi})\widetilde{\mathbf{A}}(x)^{\omega}_{\omega'}|\bar{\varphi}\rangle, \tag{1}$$

where $\mathbf{S} := \int_0^\infty \mathrm{d}t\ e^{t\mathbf{G}}$ integrates future influences; $\mathbf{\Pi}|\omega'\omega\rangle := |\omega\rangle - |\omega\rangle$ is a projection operator; $G(x)^{\omega'}_{\omega}$ denotes the transition rate at which the local configuration jumps from ω to ω'; $\widetilde{\mathbf{A}}(x)^{\omega}_{\omega'}$is a local action operator that enacts the transition $\omega \to \omega'$ locally at x; and $|\bar{\varphi}\rangle$ denotes the stationary distribution.

2.2. Learnability and Information Retention

We define learnability in the most straightforward way: as the magnitude of the local gradient (1). If the gradient is close to zero, it becomes difficult for local nodes to determine how to adjust their behavior. As demonstrated later in the experiments, obtaining large gradient magnitudes is considerably more challenging than obtaining small ones.

Definition 1 (Learnability) *The learnability at a local point x is defined as*

$$\Lambda(x) := \left\| \frac{\partial \bar{\gamma}}{\partial \mathbf{G}(x)} \right\|,$$

where the norm is arbitrary and can be specified as needed.

If we inspect the gradient (1) closely, without knowing the objective $\mathbf{\Gamma}$, the gradient norm is determined by how fast the system mixes, as captured by the following quantities.

Definition 2 (KL Mixing Distance) *As $e^{t\mathbf{G}}|\omega\rangle$ represents the distribution of the system evolving from a pure state $|\omega\rangle$ for time t, we characterize its KL divergence from the stationary distribution $|\bar{\varphi}\rangle$ by: $D_\omega(t) := D_{KL}(e^{t\mathbf{G}}|\omega\rangle \| |\bar{\varphi}\rangle)$ and $D(t) := \sup_{\omega\in\Omega} D_{KL}(e^{t\mathbf{G}}|\omega\rangle \| |\bar{\varphi}\rangle)$.*

Theorem 3 (Learnability is Bounded by KL Mixing Distance) *The learnability, measured in ℓ_∞-norm, denoted as Λ_∞, is bounded by the KL mixing distance. Formally, for any point $x \in \mathcal{X}$: $\Lambda_\infty(x) \leq 2\sqrt{2}\ \|\mathbf{\Gamma}\widetilde{\mathbf{G}}\|_\infty \cdot \int_0^\infty \mathrm{d}t\ \sqrt{D(t)}$.*

The above quantity only accounts for the worst case. A more accurate characterization may be obtained by considering the average case. In this case, it is standard to show that the KL mixing distance at time t is essentially the amount of information retention.

Definition 4 (Information Retention) *Sample $\omega_0 \sim |\bar{\varphi}\rangle$, evolve it for time t we have a random variable ω_t, where its distribution follows $e^{t\mathbf{G}}|\omega\rangle$. We denote the global information retention to be*

$$\Theta(t) := I(\omega_0; \omega_t)/H(|\bar{\varphi}\rangle),$$

where $I(\cdot;\cdot)$ denotes mutual information. The Shannon entropy $H(|\bar{\varphi}\rangle)$ of the stationary distribution $|\bar{\varphi}\rangle$ serves as a normalization factor such that $\Theta(t) \in [0, 1]$ (detailed in Proposition 7).

Proposition 5 (Information Retention $\iff$ KL Mixing Distance) *The following identity holds. $\Theta(t) \cdot H(|\bar{\varphi}\rangle) = \mathbb{E}_{\omega\sim|\bar{\varphi}\rangle}[D_\omega(t)]$.*

Therefore, we can see that information retention is closely related to learnability. We can marginalize the information retention $\Theta(t)$ to obtain a locally observable quantity.

Definition 6 (Local Information Retention) *Sample $\omega_0 \sim |\bar{\varphi}\rangle$, evolve ω for time t we have a random variable ω_t. Given local space $\Omega_x \subset \Omega$, we observe local random variables $(\omega_0)_{|\Omega_x}, (\omega_t)_{|\Omega_x} \in \Omega_x$. We denote the local information retention over Ω_x to be $\Theta_{|\Omega_x}(t) := I((\omega_0)_{|\Omega_x}; (\omega_t)_{|\Omega_x})/H(|\bar{\varphi}\rangle_{|\Omega_x})$, where $H(|\bar{\varphi}\rangle_{|\Omega_x})$ is the Shannon entropy of the marginalized stationary distribution $|\bar{\varphi}\rangle_{|\Omega_x}$.*

Proposition 7 *The local information retention $\Theta_{|\Omega_x}(t)$, for any non-empty $\Omega_x \subseteq \Omega$, has the following properties: (i) $\Theta_{|\Omega_x}(0) = 1$; (ii) $\lim_{t\to\infty} \Theta_{|\Omega_x}(t) = 0$; (iii) $\Theta_{|\Omega_x}(t) \in [0, 1]$, but may not be monotonic in t.*

Based on the theoretical insights, numerical simulations (Appendix C) show that the network exhibits a sharp transition into a high-learnability regime as local information retention increases from low to high.

3. Discussion and Future Work

The preliminary results highlight imminent future steps: conducting a deeper theoretical analysis of learnability, incorporating information retention as a design constraint, and implementing dynamical stochastic neural networks to address practical AI tasks.

References

Matthew Cook et al. Universality in elementary cellular automata. *Complex systems*, 15 (1):1–40, 2004.

Ken-ichi Funahashi and Yuichi Nakamura. Approximation of dynamical systems by continuous time recurrent neural networks. *Neural networks*, 6(6):801–806, 1993.

John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.

Chi Jin, Qinghua Liu, Yuanhao Wang, and Tiancheng Yu. V-learning—a simple, efficient, decentralized algorithm for multiagent reinforcement learning. *Mathematics of Operations Research*, 49(4):2295–2322, nov 2024. doi: 10.1287/moor.2021.0317. URL https://pubsonline.informs.org/doi/10.1287/moor.2021.0317.

Wolfgang Maass. Networks of spiking neurons: the third generation of neural network models. *Neural networks*, 10(9):1659–1671, 1997.

Valerio Mante, David Sussillo, Krishna V Shenoy, and William T Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, 503(7474): 78–84, 2013.

Frans A Oliehoek, Christopher Amato, et al. *A concise introduction to decentralized POMDPs*, volume 1. Springer, 2016.

Shayegan Omidshafiei, Jason Pazis, Christopher Amato, Jonathan P How, and John Vian. Deep decentralized multi-task multi-agent reinforcement learning under partial observability. In *International Conference on Machine Learning*, pages 2681–2690. PMLR, 2017.

Paul Rendell. Turing universality of the game of life. In *Collision-based computing*, pages 513–539. Springer, 2002.

Hava T Siegelmann and Eduardo D Sontag. On the computational power of neural nets. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 440–449, 1992.

Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8:229–256, 1992.

Kaiqing Zhang, Zhuoran Yang, Han Liu, Tong Zhang, and Tamer Basar. Fully decentralized multi-agent reinforcement learning with networked agents. In *International conference on machine learning*, pages 5872–5881. PMLR, 2018.

Yibo Jacky Zhang and Sanmi Koyejo. A framework for objective-driven dynamical stochastic fields. *arXiv preprint arXiv:2504.16115*, 2025.

Appendix A. An Overview of Intelligent Fields

Intelligent field refers to the framework of objective-driven dynamical stochastic fields (Zhang and Koyejo, 2025), which can be used to model dynamical stochastic networks. This field-theoretical framework is precise in describing the dynamics using a compact mathematical language.

We begin by considering the network as a space $\mathcal{X}$, where each point $x \in \mathcal{X}$ corresponds to a node in the network. The dynamics of the network system are thus modeled by a field $\omega(t, x) \in \Omega_x$ evolving in spacetime. The set of global configurations is denoted as $\Omega = \prod_x \Omega_x$. An illustrative figure, as provided originally in (Zhang and Koyejo, 2025), is shown in Figure 2.

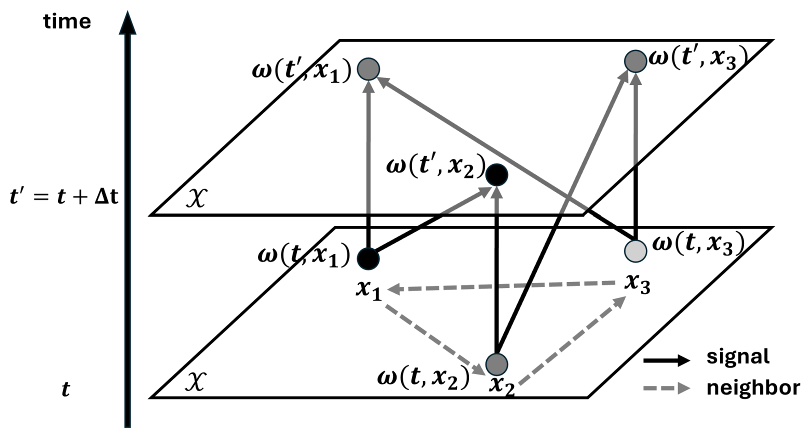


Figure 2: (Illustration of the intelligent field theoretical framework) A spacetime diagram illustrating the evolution of local configurations $\omega(t, x)$ over time for three entities x_1, x_2, x_3 in a discrete space $\mathcal{X}$. Time progresses vertically from t to $t' = t + \Delta t$, where Δt is an infinitesimal time step. Each horizontal layer corresponds to the system at a specific time. Dashed gray arrows represent directed neighboring relationships, indicating directions of signal propagation (e.g., x_1 receives signals from x_3, but not from x_2). Solid arrows represent communication and objective signals, which only propagate forward in time and are limited to immediate neighbors as defined by the dashed links. This shows that the updated local configuration $\omega(t', x)$ depends only on the previous configurations of the entity and its neighbors at time t. The local objective value is defined as the long-term average of the received objective signals, and each entity evolves to minimize its own local objective.

The theory characterizes the dynamics of the probability distribution of the field configurations, and thus no approximation is made. The probability distribution concentrated in one configuration $\omega \in \Omega$ is represented by a basis vector $|\omega\rangle \in \mathcal{H}$ where $\mathcal{H}$ is the Hilbert space spanned by all basis vectors. The inner product of two basis vectors is denoted by $\langle \omega' \mid \omega \rangle = \delta_\omega^{\omega'}$, where $\delta_\omega^{\omega'}$ is the Kronecker delta function. A generic probability distribution over system configuration is represented by a state vector $|\varphi\rangle \in \mathcal{H}$ with normalization $\sum_{\omega \in \Omega} \langle \omega \mid \varphi \rangle = 1$ and $\langle \omega \mid \varphi \rangle \geq 0$, where $\langle \omega \mid \varphi \rangle$ is the probability of observing configuration ω given the distribution $|\varphi\rangle$. The evolution of the system's state vector $|\varphi(t)\rangle$ is governed by its infinitesimal generator $\mathbf{G} : \mathcal{H} \to \mathcal{H}$, and locality implies that the global generator can be decomposed into a sum of local generators $\mathbf{G}(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}|\varphi(t)\rangle = \mathbf{G}|\varphi(t)\rangle, \qquad \mathbf{G} = \sum_x \mathbf{G}(x). \tag{2}$$

As expected from locality, the local generators commute, i.e., $[\mathbf{G}(x), \mathbf{G}(x')] = 0$, if x and x' are not neighbors.

Drawing an analogy to quantum field theory, the generators in this framework correspond to the system's Hamiltonian, and a path integral formalism naturally arises. The Lagrangian $L(w_t, w_t^+)$ of the entire system is expressed as a sum over local Lagrangians $L(w_{t,x}, w_{t,x}^+)$. The transition probability $\langle\omega'|e^{\mathbf{G}T}|\omega\rangle$, representing the evolution of the system from configuration ω to ω' over time T, can be formulated via a path integral using this Lagrangian:

$$L(w_t, w_t^+) = \sum_x L(w_{t,x}, w_{t,x}^+), \qquad \langle\omega'|e^{\mathbf{G}T}|\omega\rangle = \int \mathcal{D}\omega \ \exp\left\{-\int_0^T \mathrm{d}t \ L(w_t, w_t^+)\right\}. \tag{3}$$

As we take the limit $H \to 0$, where the entropy H quantifies the system's stochasticity (as rigorously defined in Proposition 2.22 of (Zhang and Koyejo, 2025)), the path that minimizes the Lagrangian integral becomes the dominant contribution to the path integral. This parallels how taking $\hbar \to 0$ would reveal the classic limit in Feynman's path integral of a quantum system.

To model how the system receives objective signals such as penalties and rewards, a higher-dimensional space $\widetilde{\mathcal{H}} := \mathcal{H} \otimes \mathcal{H}$ is introduced, where it has basis vectors in the form of $|\omega'\omega\rangle$, and the generator $\mathbf{G}$ is lifted to $\widetilde{\mathbf{G}} : \mathcal{H} \to \widetilde{\mathcal{H}}$. Each local node x is then associated with an objective operator $\mathbf{\Gamma}_x : \widetilde{\mathcal{H}} \to \mathbb{R}$, and evolves to minimize its objective value: the time average of the objective signals it receives over an infinite time horizon. The objective value $\bar{\gamma}(x)$ and its gradient with respect to a local generator $\mathbf{G}(x)$ are given by

$$\bar{\gamma}(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}t \ \mathbf{\Gamma}_x \widetilde{\mathbf{G}}|\varphi(t)\rangle, \qquad \frac{\partial\bar{\gamma}(x)}{\partial G(x)_\omega^{\omega'}} = \mathbf{\Gamma}_x(1 + \widetilde{\mathbf{G}}\mathbf{S}\mathbf{\Pi})\widetilde{\mathbf{A}}(x)_{\omega'}^{\omega}|\bar{\varphi}\rangle, \tag{4}$$

where $\mathbf{S} := \int_0^\infty \mathrm{d}t \ e^{t\mathbf{G}}$ is a linear operator that integrates future influences; $\mathbf{\Pi}|\omega'\omega\rangle := |\omega\rangle - |\omega\rangle$ is a projection operator; $G(x)_\omega^{\omega'}$ denotes the transition rate at which the local configuration jumps from ω to ω'; $\widetilde{\mathbf{A}}(x)_{\omega'}^{\omega}$ is a local action operator that enacts the transition $\omega \to \omega'$ locally at x; and $|\bar{\varphi}\rangle$ denotes the stationary distribution.

Credit assignment of the network is done via the propagation of objective signals. It is shown that any objective propagation can be represented as a linear operator, which is termed the objective propagator. In particular, an interesting class of such propagators is

$$\mathbf{P}[\mathbf{Q}] := 1 + \widetilde{\mathbf{G}}\mathbf{S}\mathbf{Q}\mathbf{\Pi}, \tag{5}$$

where $\mathbf{Q}$ can be any linear operator that satisfies a normalization constraint. In this way, each node can locally compute a closed-form gradient expression that guides how it should adapt its behavior.

In summary, Zhang and Koyejo (2025) provide a theoretical framework that describes both the dynamics and credit assignment in intelligent fields through objective propagation.

Appendix B. Complete Derivation of the Theoretical Results

Theorem 8 (Theorem 3 Restated) *The learnability, measured in ℓ_∞-norm, denoted as Λ_∞, is bounded by the KL mixing distance. Formally, for any point $x \in \mathcal{X}$:*

$$\Lambda_\infty(x) \leq 2\sqrt{2}\ \|\mathbf{\Gamma}\widetilde{\mathbf{G}}\|_\infty \cdot \int_0^\infty \mathrm{d}t\ \sqrt{D(t)}. \tag{6}$$

Proof Consider a two-entity view of the system: it is composed of two entities, x and the rest $y := \mathcal{X}\backslash x$. We call x the acting entity and the rest the environmental entity.

In the two-entity view, the following notations become handy and clearer. The entity x possesses a configuration $\mu \in \mathcal{M}$ and $\alpha \in \mathcal{A}$, where only α is observable to the environmental entity y. Symmetrically, entity y possesses a configuration $\nu \in \mathcal{N}$ and $\beta \in \mathcal{B}$, where only β is observable to entity x. Therefore, the system configuration is described by:

$$\omega = (\alpha, \beta, \mu, \nu) \in \Omega, \qquad \text{where } \ \Omega = \mathcal{A} \times \mathcal{B} \times \mathcal{M} \times \mathcal{N}. \tag{7}$$

It is proven (Zhang and Koyejo, 2025) that locality implies the following decomposition of the generator $\mathbf{G}$.

$$\mathbf{G} = \mathbf{M} + \mathbf{N}, \tag{8}$$

where $\mathbf{G}$ is the generator of the whole system, $\mathbf{M} = \mathbf{G}(x)$ is the generator of the acting entity, and $\mathbf{N} = \sum_{x' \in \mathcal{X}\backslash x} \mathbf{G}(x')$ is the generator of the environmental entity. Note that $\mathbf{M}$ completely characterizes the behavior of the acting entity.

The gradient of the objective $\bar{\gamma}$ given by the environment with respect to the agent's generator can thus be formulated as

$$\frac{\partial \bar{\gamma}}{\partial M_{\alpha\beta\mu}^{\alpha'\mu'}} = \mathbf{\Gamma}\widetilde{\mathbf{G}}\mathbf{S}\mathbf{\Pi}\widetilde{\mathbf{A}}_{\alpha'\mu'}^{\alpha\beta\mu}|\bar{\varphi}\rangle. \tag{9}$$

Note that $|\bar{\varphi}\rangle$ is the stationary state, representing the stationary distribution of the whole system. Let us denote $\bar{p}(\cdot)$ as the stationary distribution over $\omega = (\alpha, \beta, \mu, \nu) \in \Omega$. Then, the above gradient formula can be rewritten as

$$\frac{\partial \bar{\gamma}}{\partial M_{\alpha\beta\mu}^{\alpha'\mu'}} = \int_0^\infty \mathrm{d}t\ \mathbb{E}_{\nu\sim\bar{p}(\cdot|\alpha\beta\mu)} \left[\mathbf{\Gamma}\widetilde{\mathbf{G}}e^{t\mathbf{G}}(|\alpha'\beta\mu'\nu\rangle - |\alpha\beta\mu\nu\rangle)\right] \cdot \bar{p}(\alpha\beta\mu). \tag{10}$$

We start from the gradient formula (10):

$$\Lambda_\infty(x) = \max_{\alpha\beta\mu\alpha'\mu'} \left| \frac{\partial \bar{\gamma}}{\partial M_{\alpha\beta\mu}^{\alpha'\mu'}} \right| \tag{11}$$

$$= \max_{\alpha\beta\mu\alpha'\mu'} \left| \int_0^\infty \mathrm{d}t \; \mathbb{E}_{\nu\sim\bar{p}(\cdot|\alpha\beta\mu)} \left[\mathbf{\Gamma}\widetilde{\mathbf{G}} e^{t\mathbf{G}} (|\alpha'\beta\mu'\nu\rangle - |\alpha\beta\mu\nu\rangle) \right] \cdot \bar{p}(\alpha\beta\mu) \right| \tag{12}$$

$$\le \max_{\alpha\beta\mu\alpha'\mu'} \left| \int_0^\infty \mathrm{d}t \; \mathbb{E}_{\nu\sim\bar{p}(\cdot|\alpha\beta\mu)} \left[\mathbf{\Gamma}\widetilde{\mathbf{G}} e^{t\mathbf{G}} (|\alpha'\beta\mu'\nu\rangle - |\alpha\beta\mu\nu\rangle) \right] \right| \tag{13}$$

$$\le \max_{\alpha\beta\mu\alpha'\mu'} \mathbb{E}_{\nu\sim\bar{p}(\cdot|\alpha\beta\mu)} \left[\int_0^\infty \mathrm{d}t \; \left| \mathbf{\Gamma}\widetilde{\mathbf{G}} e^{t\mathbf{G}} (|\alpha'\beta\mu'\nu\rangle - |\alpha\beta\mu\nu\rangle) \right| \right] \tag{14}$$

$$\le \max_{\alpha\beta\mu\alpha'\mu'} \mathbb{E}_{\nu\sim\bar{p}(\cdot|\alpha\beta\mu)} \left[\|\mathbf{\Gamma}\widetilde{\mathbf{G}}\|_\infty \cdot \int_0^\infty \mathrm{d}t \; \underbrace{\left\| e^{t\mathbf{G}} (|\alpha'\beta\mu'\nu\rangle - |\alpha\beta\mu\nu\rangle) \right\|_1}_{(A)} \right]. \tag{15}$$

In the last step, the norm $\|\cdot\|_1$ is the L^1 norm which applies to Hilbert space $\mathcal{H}$ with the canonical orthonormal basis $\{|\omega\rangle\}_{\omega\in\Omega}$. Accordingly, the norm of the linear operator $\mathbf{\Gamma}\widetilde{\mathbf{G}}$ is the operator norm with respect to the L^1 norm, which is equivalent to the L^∞ norm on $\mathcal{H}$.

In the next step, we deal with the term (A) as above. The trick is that we insert the stationary distribution $|\bar{\varphi}\rangle \in \mathcal{H}$, where $|\bar{\varphi}\rangle = e^{t\mathbf{G}}|\bar{\varphi}\rangle$ is a fixed point, and then apply Pinsker's inequality.

$$(A) = \left\| e^{t\mathbf{G}} (|\alpha'\beta\mu'\nu\rangle - |\bar{\varphi}\rangle + |\bar{\varphi}\rangle - |\alpha\beta\mu\nu\rangle) \right\|_1 \le \left\| e^{t\mathbf{G}} (|\alpha'\beta\mu'\nu\rangle - |\bar{\varphi}\rangle) \|_1 + \| e^{t\mathbf{G}} (|\bar{\varphi}\rangle - |\alpha\beta\mu\nu\rangle) \right\|_1 \tag{16}$$

$$= \left\| e^{t\mathbf{G}} |\alpha'\beta\mu'\nu\rangle - |\bar{\varphi}\rangle \|_1 + \| e^{t\mathbf{G}} |\alpha\beta\mu\nu\rangle - |\bar{\varphi}\rangle \right\|_1 \tag{17}$$

$$\le \sqrt{2D_{KL}(e^{t\mathbf{G}}|\alpha'\beta\mu'\nu\rangle \| |\bar{\varphi}\rangle)} + \sqrt{2D_{KL}(e^{t\mathbf{G}}|\alpha\beta\mu\nu\rangle \| |\bar{\varphi}\rangle)} \qquad \text{(Pinsker's inequality)}$$

$$\le 2\sqrt{2D(t)}. \qquad \text{(Definition 2)}$$

Combining the above, we arrive at

$$\Lambda_\infty(x) \le \max_{\alpha\beta\mu\alpha'\mu'} \mathbb{E}_{\nu\sim\bar{p}(\cdot|\alpha\beta\mu)} \left[\|\mathbf{\Gamma}\widetilde{\mathbf{G}}\|_\infty \cdot \int_0^\infty \mathrm{d}t \; 2\sqrt{2D(t)} \right] \tag{18}$$

$$= \|\mathbf{\Gamma}\widetilde{\mathbf{G}}\|_\infty \cdot \int_0^\infty \mathrm{d}t \; 2\sqrt{2D(t)}. \tag{19}$$

Rearranging the above inequality completes the proof. ■

Proposition 9 (Proposition 5 Restated) *The following identity holds.*

$$\Theta(t) \cdot H(|\bar{\varphi}\rangle) = \mathbb{E}_{\omega\sim|\bar{\varphi}\rangle}[D_\omega(t)]. \tag{20}$$

Proof Denoting the transition probability $p_t(\omega' \mid \omega) := \langle \omega' | e^{t\mathbf{G}} | \omega \rangle$, we start from the definition of the KL divergence:

$$\mathbb{E}_{\omega\sim\bar{p}}[D_\omega(t)] = \sum_{\omega} \bar{p}(\omega) \sum_{\omega'} p_t(\omega' \mid \omega) \log \frac{p_t(\omega' \mid \omega)}{\bar{p}(\omega')} \tag{21}$$

$$= \sum_{\omega} \bar{p}(\omega) \sum_{\omega'} p_t(\omega' \mid \omega) \log \frac{\bar{p}(\omega) p_t(\omega' \mid \omega)}{\bar{p}(\omega)\bar{p}(\omega')}. \tag{22}$$

Next, to connect it with mutual information, let us examine more closely at the distribution of (ω_0, ω_t). Note that we denote $\bar{p}$ as the probability mass function of the stationary distribution, which is essentially the same as the stationary state $|\bar{\varphi}\rangle \in \mathcal{H}$. Given that $\omega_0 \sim \bar{p}$ is from the stationary distribution, the marginal distribution of ω_t is also the stationary distribution (formally, $|\bar{\varphi}\rangle = e^{t\mathbf{G}}|\bar{\varphi}\rangle$ is the fixed point). Therefore, we have $\Pr(\omega_t = \omega') = \bar{p}(\omega')$. Given this observation, we can continue from above:

$$(22) = \sum_{\omega} \Pr(\omega_0 = \omega) \sum_{\omega'} \Pr(\omega_t = \omega' \mid \omega_0 = \omega) \log \frac{\Pr(\omega_0 = \omega)\Pr(\omega_t = \omega' \mid \omega_0 = \omega)}{\Pr(\omega_0 = \omega)\Pr(\omega_t = \omega')} \tag{23}$$

$$= \sum_{\omega,\omega'} \Pr(\omega_0 = \omega, \omega_t = \omega') \log \frac{\Pr(\omega_0 = \omega, \omega_t = \omega')}{\Pr(\omega_0 = \omega)\Pr(\omega_t = \omega')} \tag{24}$$

$$= I(\omega_0; \omega_t) = \Theta(t) \cdot H(\bar{p}). \tag{25}$$

■

Proposition 10 (Proposition 7 Restated) *The local information retention $\Theta_{|\Omega_x}(t)$, for any non-empty $\Omega_x \subseteq \Omega$, has the following properties: (i) $\Theta_{|\Omega_x}(0) = 1$; (ii) $\lim_{t\to\infty} \Theta_{|\Omega_x}(t) = 0$; (iii) $\Theta_{|\Omega_x}(t) \in [0, 1]$, but may not be monotonic in t.*

Proof First, by the definition of mutual information,

$$I((\omega_0)_{|\Omega_x}; (\omega_0)_{|\Omega_x}) = H((\omega_0)_{|\Omega_x}) = H(\bar{p}_{|\Omega_x}). \tag{26}$$

Note that we denote $\bar{p}$ as the probability mass function of the stationary distribution, which is essentially the same as the stationary state $|\bar{\varphi}\rangle \in \mathcal{H}$.

This implies

$$\Theta_{|\Omega_x}(0) = \frac{I((\omega_0)_{|\Omega_x}; (\omega_0)_{|\Omega_x})}{H(\bar{p}_{|\Omega_x})} = 1. \tag{27}$$

Then, given the ergodicity, there is a unique stationary distribution $|\bar{\varphi}\rangle$, and its marginal is also unique. Therefore, the random variable $(\omega_\infty)_{|\Omega_x})$ follows the distribution $\bar{p}_{|\Omega_x}$, which does not have information about ω_0. Therefore, $I((\omega_0)_{|\Omega_x}; (\omega_\infty)_{|\Omega_x}) = 0$.

Lastly, it is easy to see that

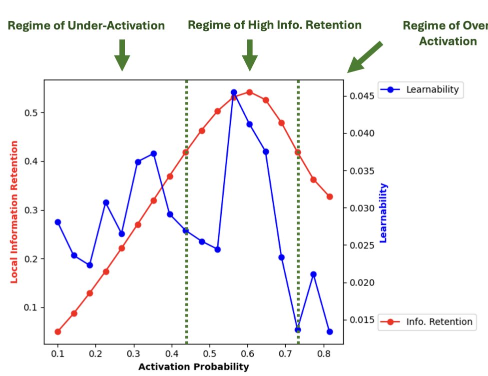
$$0 \le I((\omega_0)_{|\Omega_x}; (\omega_t)_{|\Omega_x}) = H((\omega_0)_{|\Omega_x}) - H((\omega_0)_{|\Omega_x} \mid (\omega_t)_{|\Omega_x}) \le H((\omega_0)_{|\Omega_x}) = H(\bar{p}_{|\Omega_x}). \tag{28}$$

Therefore, $\Theta_{|\Omega_x}(t) \in [0, 1]$. Noting that the marginal process ω_t may not be Markovian, we can see that the mutual information $I((\omega_0)_{|\Omega_x}; (\omega_t)_{|\Omega_x})$ is not necessarily non-increasing.

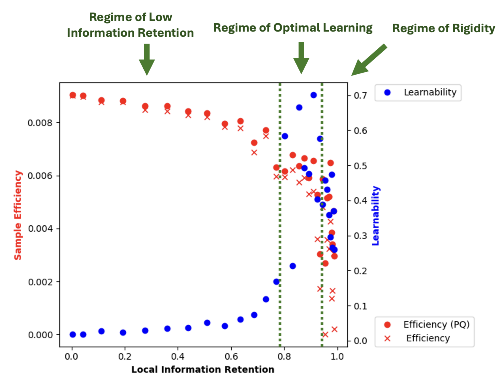
■

Appendix C. Numerical Simulation

In this synthetic experiment, we ask a small binary dynamical stochastic network to learn the XOR operation. The network consists of 13 nodes, each connected to two other nodes, with recurrent connections included. When an input is given to two nodes, the network runs for 30 steps, after which the environment reads the output from a node to check whether it matches the result of the XOR operation. The results are presented and discussed in Figure 3, showing how local information retention is closely related to global learnability.



(a) As the activation probability across all nodes is gradually increased, the network transitions from a under-activated regime to an over-activated regime. Both local information retention and learnability peak between these two phases. However, in this scenario, both learnability and information retention are significantly lower than those observed in (b). This suggests that simply controlling the activation probability and operating at the edge of under/over activation does not guarantee great learnability.



(b) As local information retention is controlled and increased, the network transitions from the regime of low information retention to a regime of rigidity where it becomes almost deterministic. During this transition, learnability sees a sudden improvement. However, sample efficiency undergoes a sudden decline. It's important to note that sample efficiency improves when the network utilizes the $\mathbf{P}[\mathbf{Q}]$ objective propagation, which decomposes and propagates objective signals more effectively.

Figure 3: A binary dynamical stochastic network consisting of 13 nodes is tasked with learning the XOR operation. Each node is connected to two other nodes, with recurrent connections present. When given an input to two nodes, the network is run for 30 steps, after which the environment reads the output from a node to check if it matches the result of the XOR operation.