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

NEURAL INFORMATION PROCESSING SYSTEMS

Yibo Jacky Zhang  Sanmi Koyejo
Department of Computer Science, Stanford University

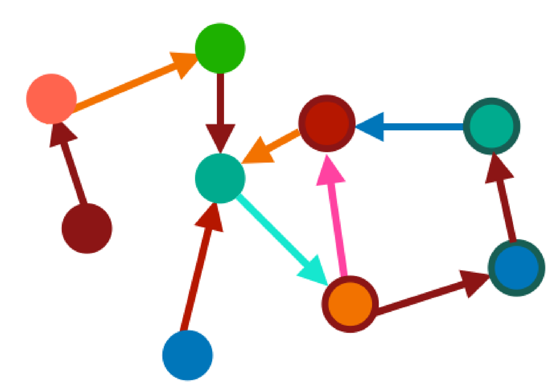
## Introduction

### Dynamical Stochastic Networks

- A network of stochastically evolving nodes
- Each neuron has its own state
- Each neuron interacts with its neighbors
- It may receive input or generate output signals
- Local evolution and local learning rule

### Examples

- **Bio-plausible (artificial) neural networks**
- Generalized cellular automata
- Multi-agent systems



### Research Question

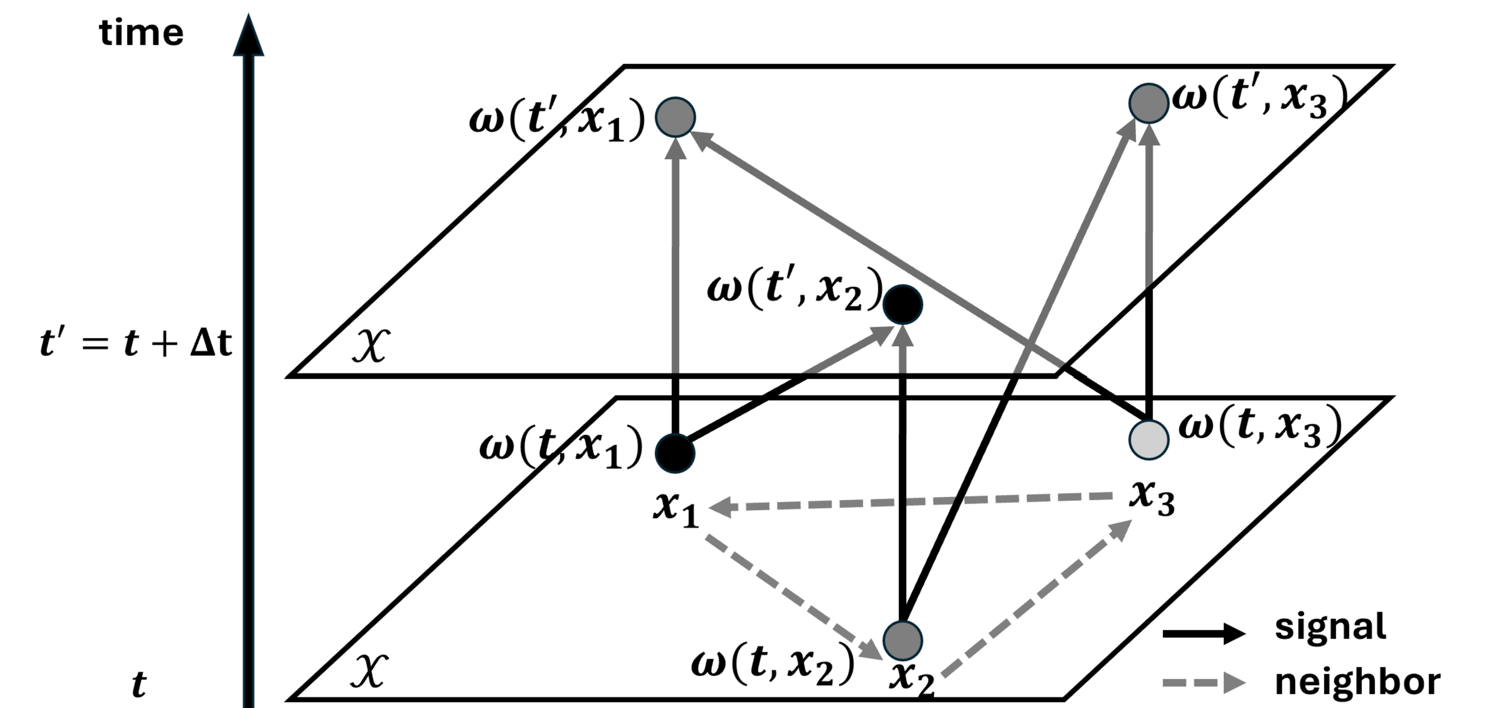
**How can individual neurons, with a <u>local learning rule</u>, <u>efficiently</u> achieve a given <u>global objective</u>?**

### Challenges to Answer the Question

- **<u>Challenge 1: Theoretical Modeling.</u>**
  - ➢ To understand and analyze the learning in the stochastic dynamical system
- **<u>Challenge 2: Credit Assignment.</u>**
  - ➢ We consider credit assignment conveyed through the local propagation of scalar <u>objective signals</u> (e.g., reward or penalty).
  - ➢ <u>Objective propagation</u>: each neuron sends objective signals to its neighbors; each neuron adapts its behavior based on what it receives
- **<u>Challenge 3: Regimes of Learnability</u>**
  - ➢ Inefficient learning If neurons fire chaotically
  - ➢ Inefficient learning if neurons fire too sparsely
  - ➢ <u>What is the fundamental property that is necessary to efficient learning?</u>

## Intelligent Field Theory

We use the theory of *objective-driven dynamical stochastic fields*, also referred to as *intelligent fields* [1].




- The evolving network is modeled by a field spanned in spacetime; each neuron correspond to a label $x$
- Field configuration $\omega$; local configuration $\omega(t, x)$
- Evolving probability distribution of the field configurations, i.e., "superposition" $|\varphi\rangle \in \mathcal{H}$
- Evolution governed by infinitesimal generators $G$:

$$\frac{d}{dt}|\varphi\rangle = G|\varphi\rangle$$

- The $G = \sum_x G(x)$ is the sum of local generators
- Local generators commute $[G(x), G(x')] = 0$ for non-neighboring neurons $x$ and $x'$
- There also is a path integral formalism
- The objective propagation is modeled by an operator called <u>objective propagator</u>:

$$P[Q] = 1 + \tilde{G}SQ\Pi$$

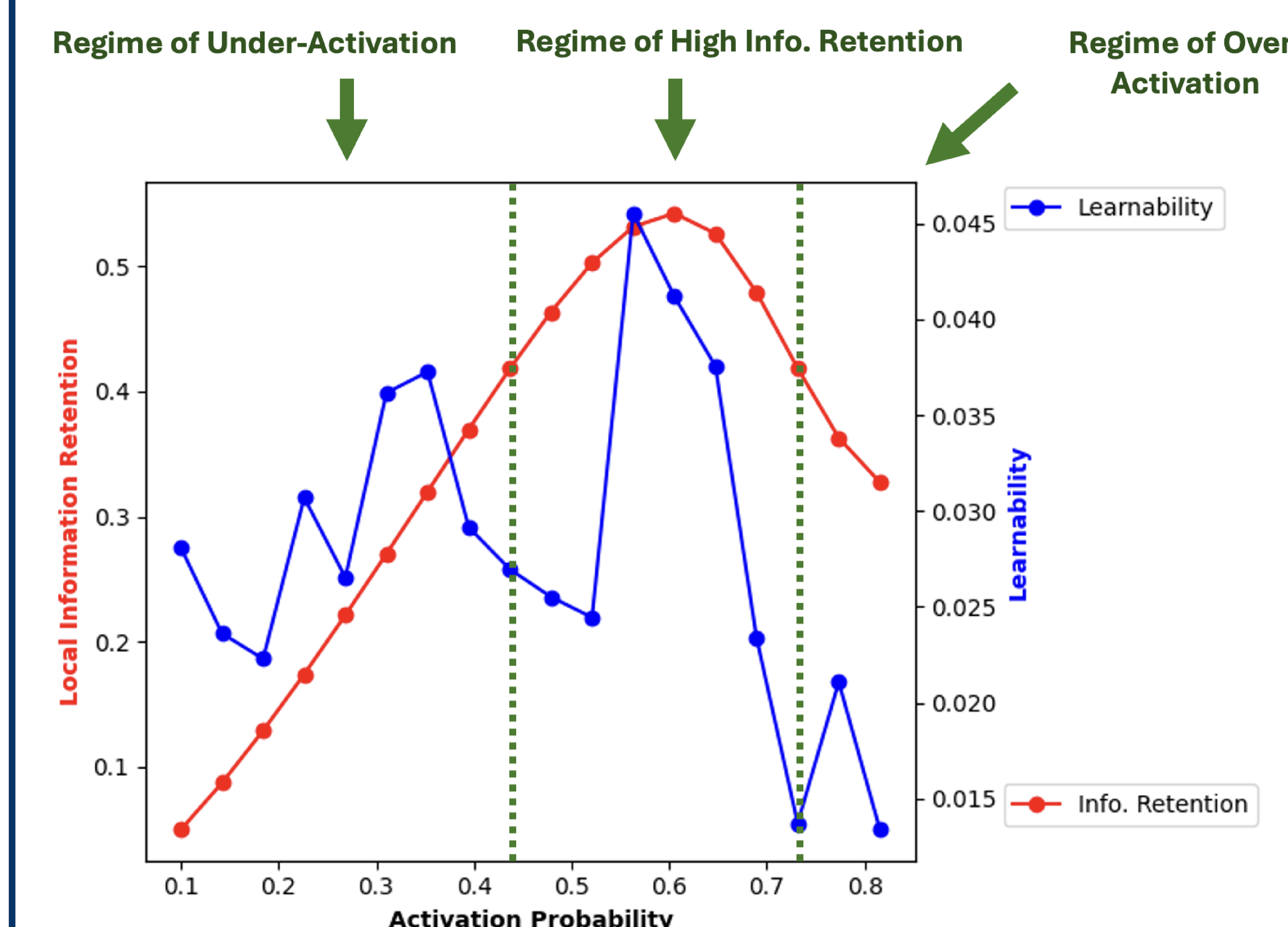
## Learnability

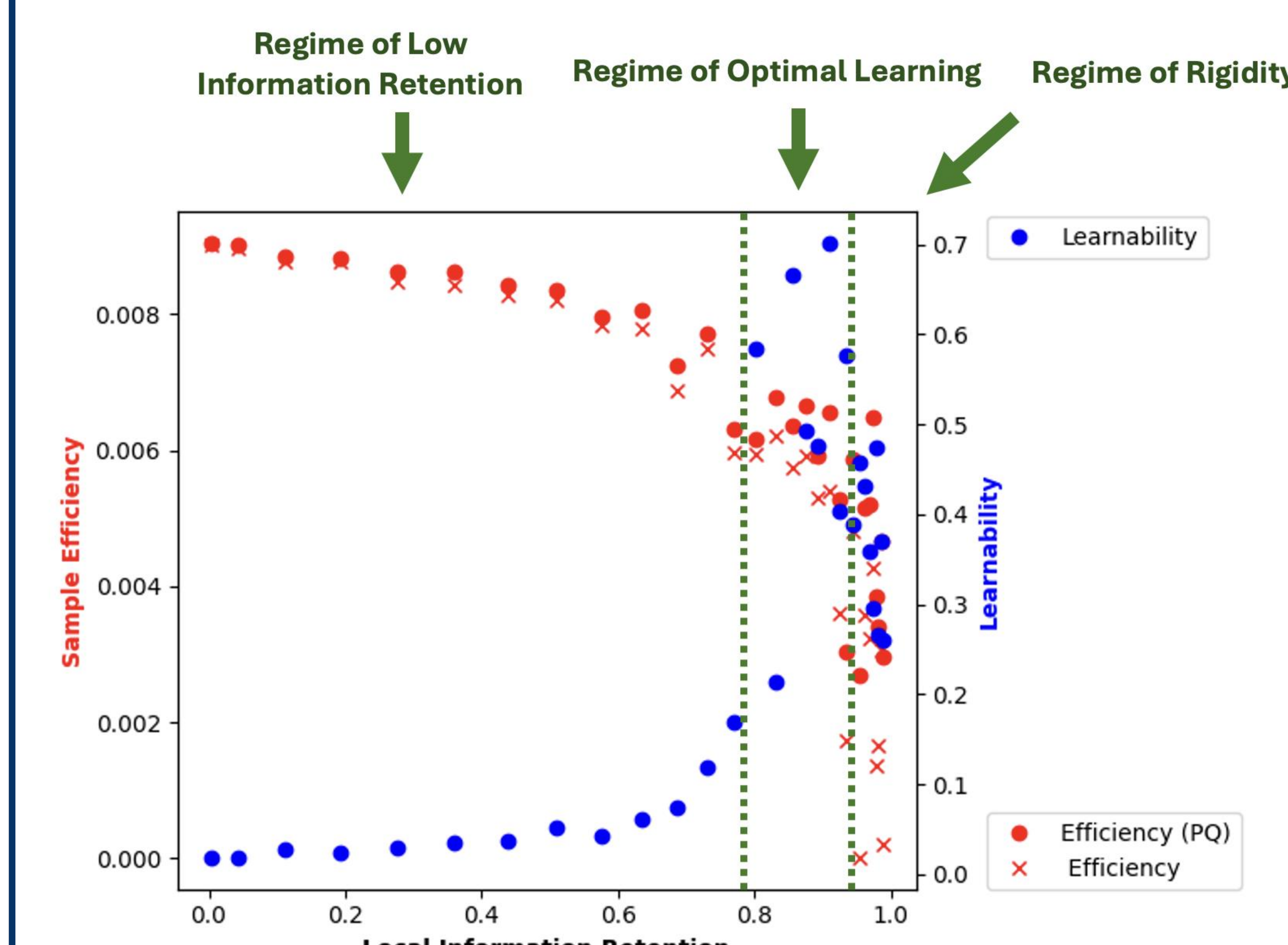
- **<u>Learnability:</u> The magnitude of the local gradient**
  - ➢ the ability of which each neuron can adjust its behavior towards its objective
- **Our theory suggests: <u>Information retention is necessary to efficient learning</u>**

- **<u>Information Retention</u>: the mutual information between the current and future observations**
  - ➢ How fast the system forgets what happened in the past

## Numerical Simulations

- Task: learn the XOR operation.
- The network consists of 13 binary neurons
- Each connected to two other neurons
- When an input is given to two neurons, 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.




[1] Zhang and Koyejo. A Framework for Objective-Driven Dynamical Stochastic Fields, 2025.