UNDERSTANDING MEMORY IN NEURAL NETWORKS THROUGH FISHER INFORMATION DIFFUSION

Haodong Qin, Tatyana Sharpee
University of California San Diego, Salk Institute of Biological Studies

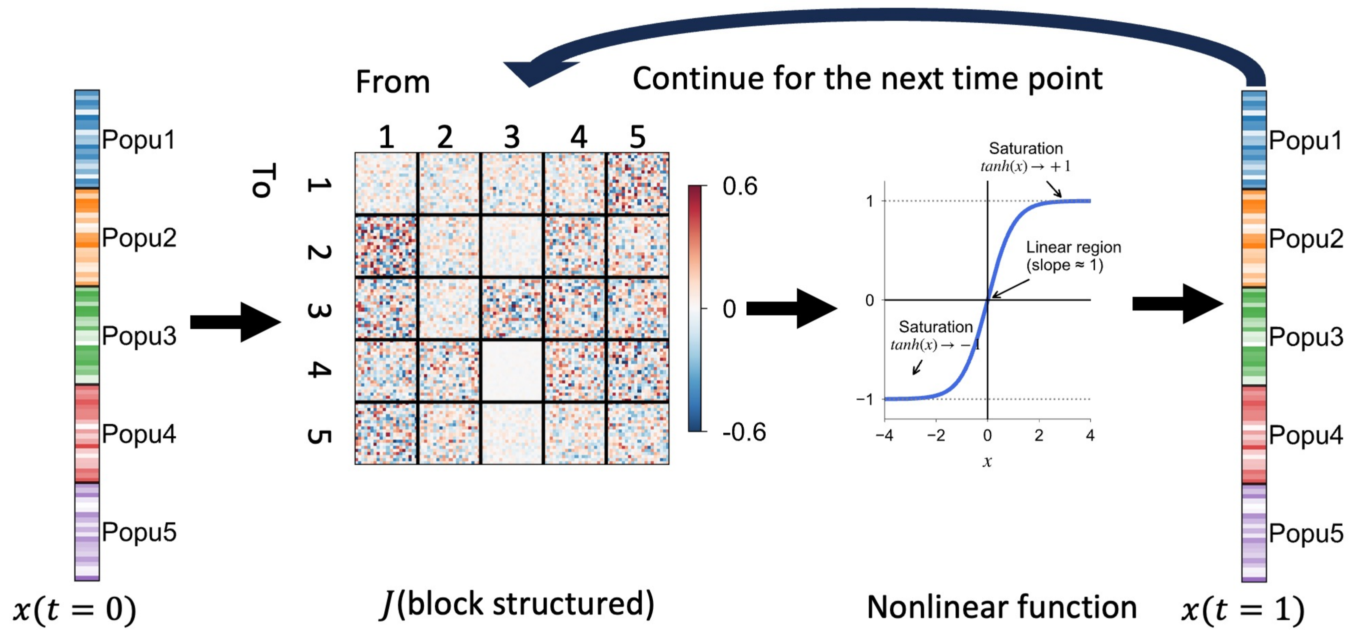
Abstract

- **Dynamic Memory and Information Flow**
Stimulus representations evolve along stable manifolds, preserving input geometry rather than converging to static attractors.
- **General Framework for Recurrent Networks**
Theoretical framework showing how information can be maintained dynamically while **preserving the geometry of inputs** and supports flexible input and output configuration.
- **Information diffusion**
Derive a **Fisher information diffusion operator** that quantifies how information propagates between interacting populations.
- Principled Design Rules
Simple, theory-driven **initialization schemes** ensure stability at the **edge of chaos**.
- Dynamic memory: Sequential memory
Improved **sequential memory** and **training efficiency** in recurrent networks.

Introduction

- **Working memory requires dynamics**
Edge-of-chaos criteria characterize network stability, but they do not directly measure memory—how distinguishable past inputs remain as representations evolve, nor how information is tracked, propagated, across populations over time.
- **Need for a Multi-Population Framework**
Theoretical models often assume **homogeneous connectivity**, failing to capture **structured, multi-area circuits** in brains and AI architectures.

RNN dynamics

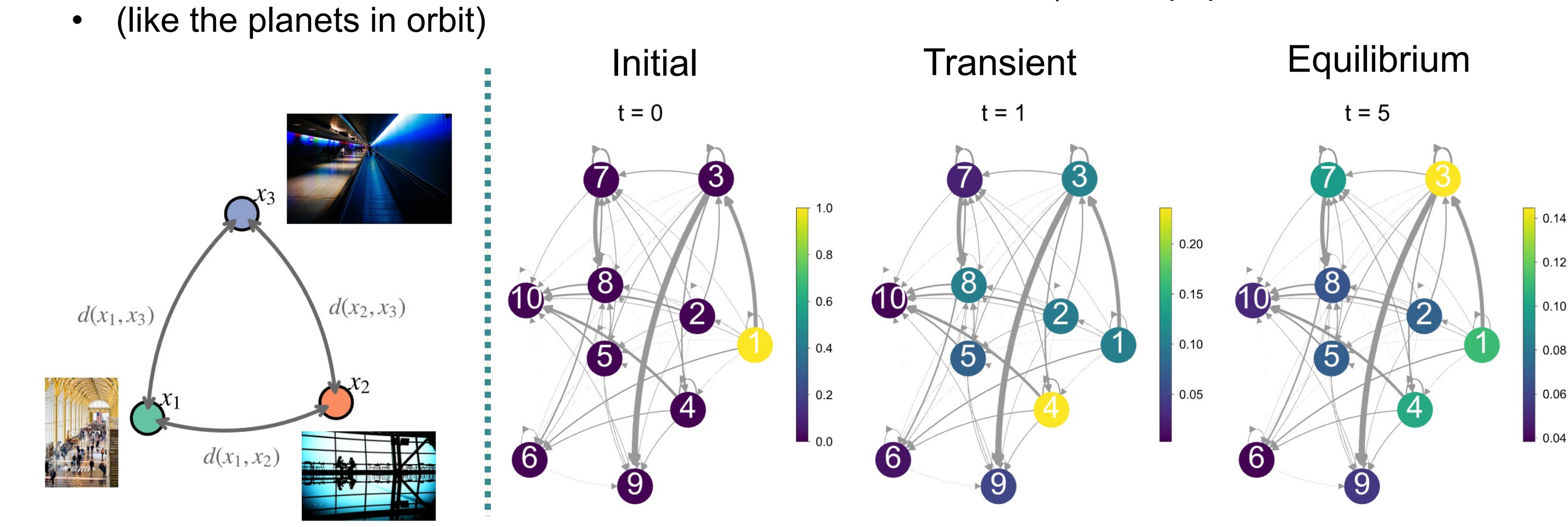


Findings

RNN: $h_i(t) = \sum_{j=1}^{N} J_{ij} S_j(t) + \eta_i(t), \qquad S_j(t) = \phi[w_j x(t) + h_j(t-1)]$

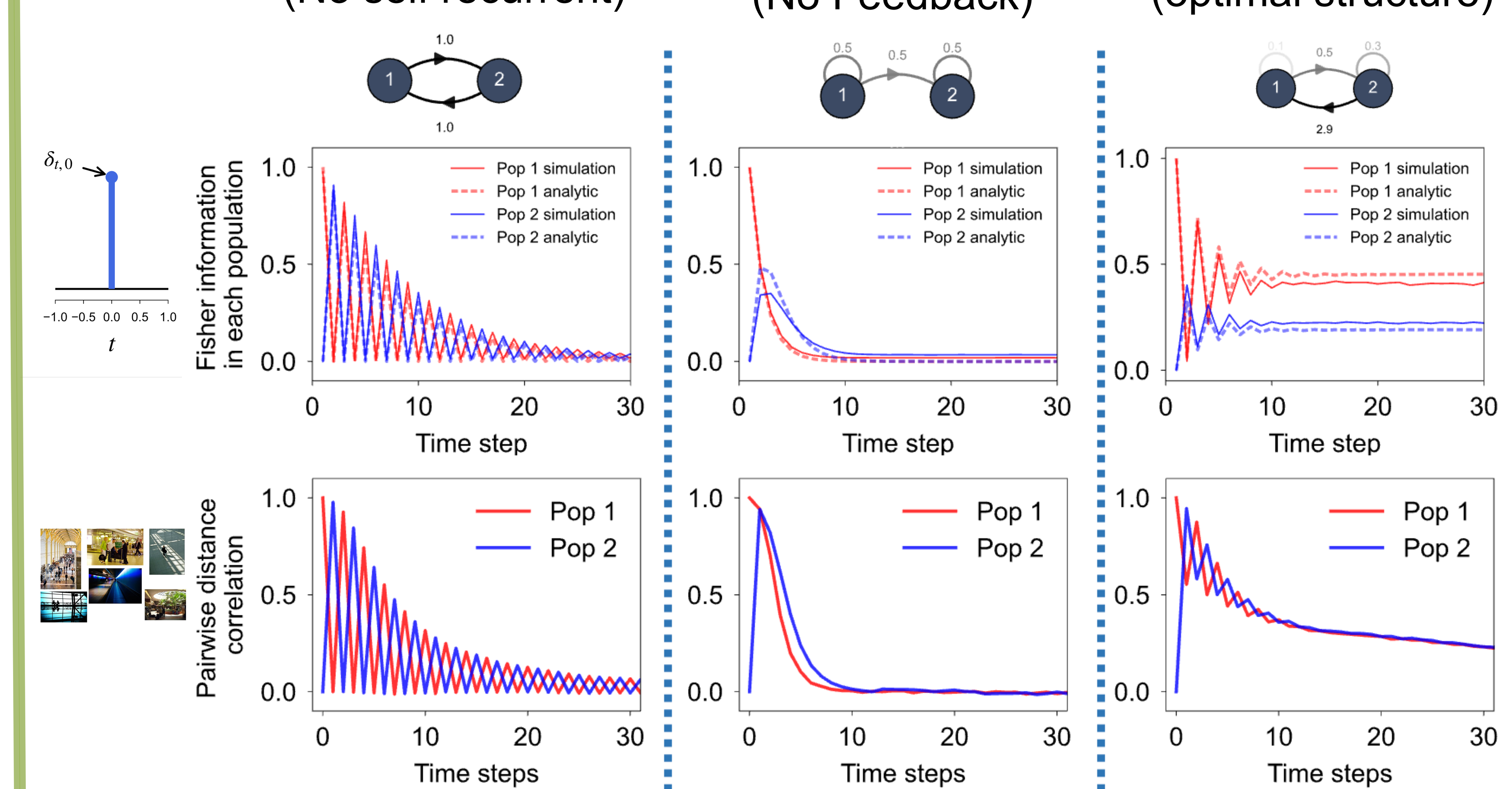
- Representation of the input changes with time
- Memory as preservation of pairwise Distance
  - (like the planets in orbit)

Diffusion of Fisher information (not activities) in subpopulations



Validation of Analytic Solution Against Simulation

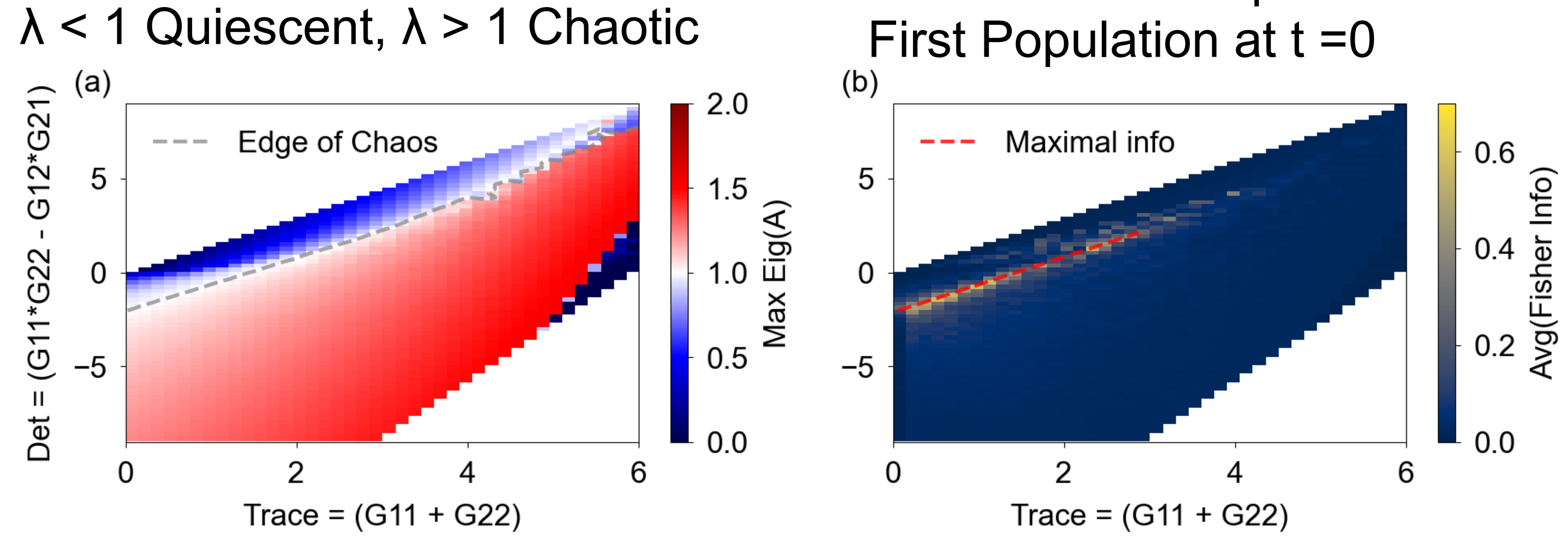
Oscillation (No self recurrent) | Fast decay (No Feedback) | Saturation (optimal structure)



Phase diagram for two populations

Eigenvalue Spectrum:
λ = 1 Dynamic Stability,
λ < 1 Quiescent, λ > 1 Chaotic

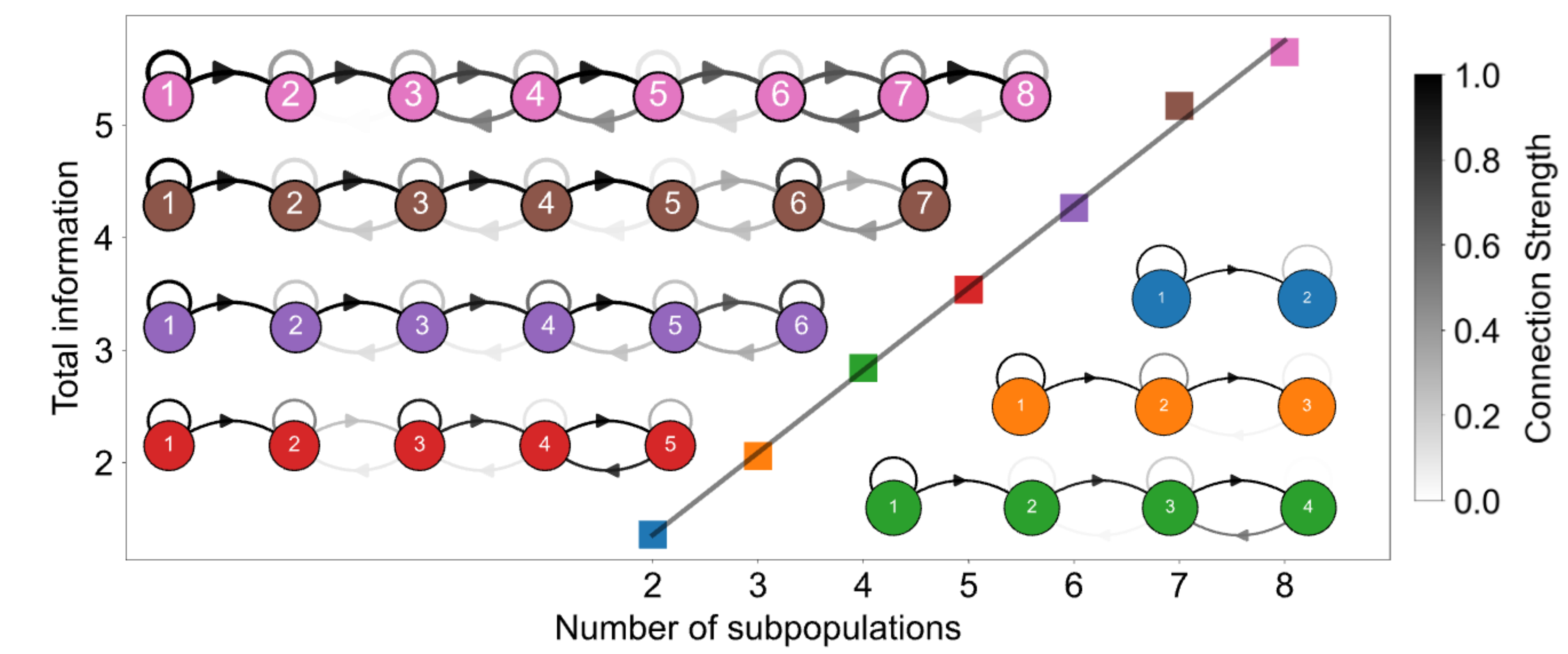
Cumulative Fisher Information for Input to the First Population at t =0



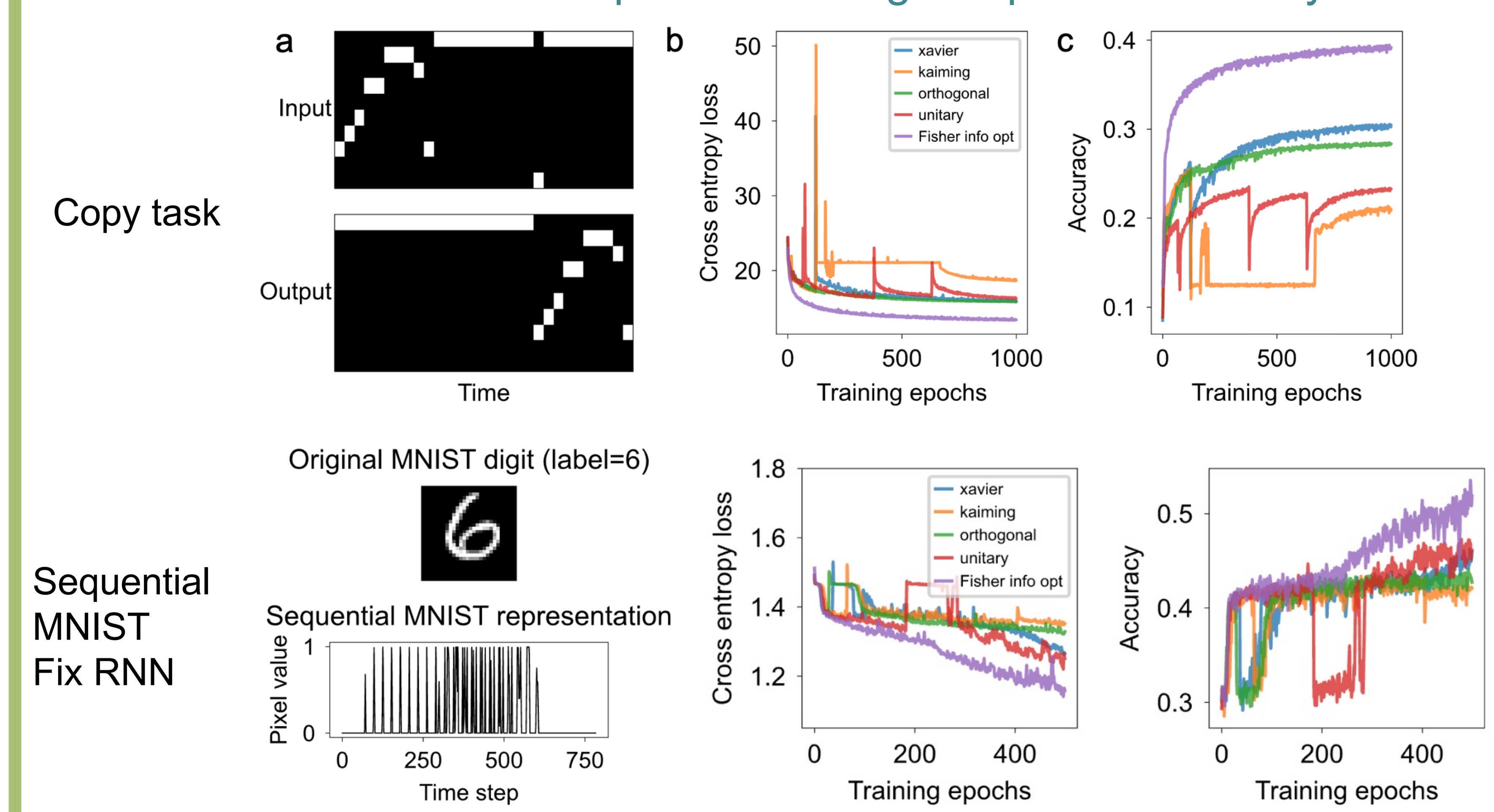
Optimal structure for information storage: Criticality + Alignment

Applications

Optimal structure for Fisher information



Quick initialization rule improves training: Sequential memory task



Conclusion

- **Dynamic representation** of stimuli through evolving neural activity, not static attractors
- **Information = preservation** of input geometry, beyond number of fixed patterns
- **General framework** with multi-population and flexible input – output coupling
- **Physics inspired model:** information flow as diffusion
- **Optimal structure for Fisher Memory:** criticality + alignment
- **Scalable computation:** complexity ∝ depth, not network size
- **Fisher-optimized initialization accelerates training** and enhances sequential memory in recurrent networks.

Thanks also to Arjun Karuvally for valuable discussions and helpful suggestions.