Optimal Oscillator Memory Networks

Connor Bybee

Advisor: Fritz Sommer

REDWOOD CENTER for Theoretical Neuroscience

Berkeley UNIVERSITY OF CALIFORNIA

intel

Motivation

Attractor networks are important models of memory in neuroscience & ML

- Memory networks models for hippocampus and other brain areas
- Error-correction in Vector Symbolic Architectures (VSA)
- “Modern Hopfield Networks” in Transformer networks

Many traditional attractor models are inefficient

- How to design efficient associative memories for neuromorphic hardware or coupled oscillators?

Classical "Hopfield" Associative Memory

An associative memory stores a set of patterns for robust recall

Binary Patterns

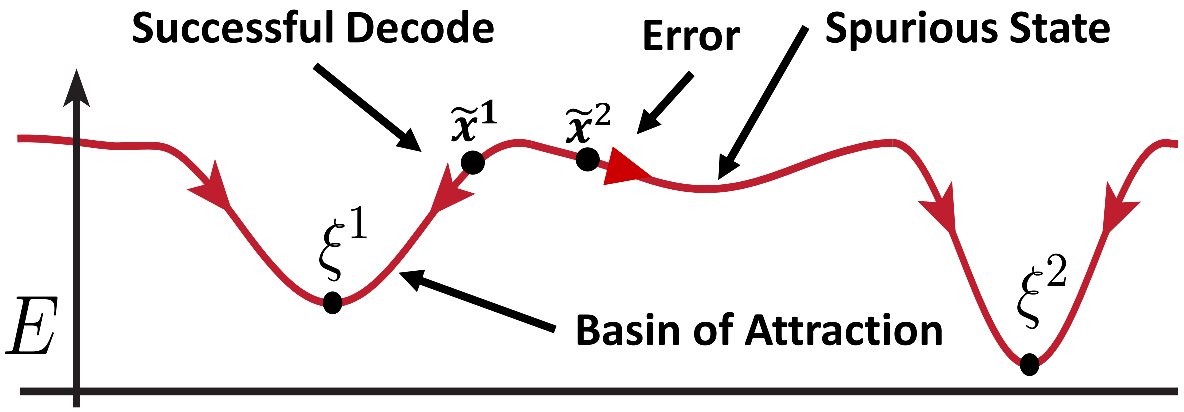
$\boldsymbol{\xi}^{\boldsymbol{k}} \in \{-1, 1\}^{N}$

Outer Product Learning Rule

$$\boldsymbol{W} = \frac{1}{M}\sum_{k=1}^{M} \boldsymbol{\xi}^{\boldsymbol{k}}\boldsymbol{\xi}^{\boldsymbol{k}T}$$

Energy Function

$E = -\boldsymbol{x}^{T}\boldsymbol{W}\boldsymbol{x}$



System Dynamics

$\Delta \boldsymbol{x} = f(-\nabla_{\boldsymbol{x}} E(\boldsymbol{x}))$

$\boldsymbol{x}(t+1) = f(\boldsymbol{W}\boldsymbol{x}(t))$

Error-Correction (Two Components)

- Collective State Computation
- State Quantization $f(x) = \text{sign}(x)$

Hopfield, John J. "Neural networks and physical systems with emergent collective computational abilities." *Proceedings of the national academy of sciences* 79.8 (1982): 2554-2558.

Efficiency of Associative Memories

Parameters

M	Patterns
N	Units/Neurons
$S = \lvert\{W_{ij}\}\rvert$	Synapses/Parameters
I_p	Pattern Information (Bits/Pattern)

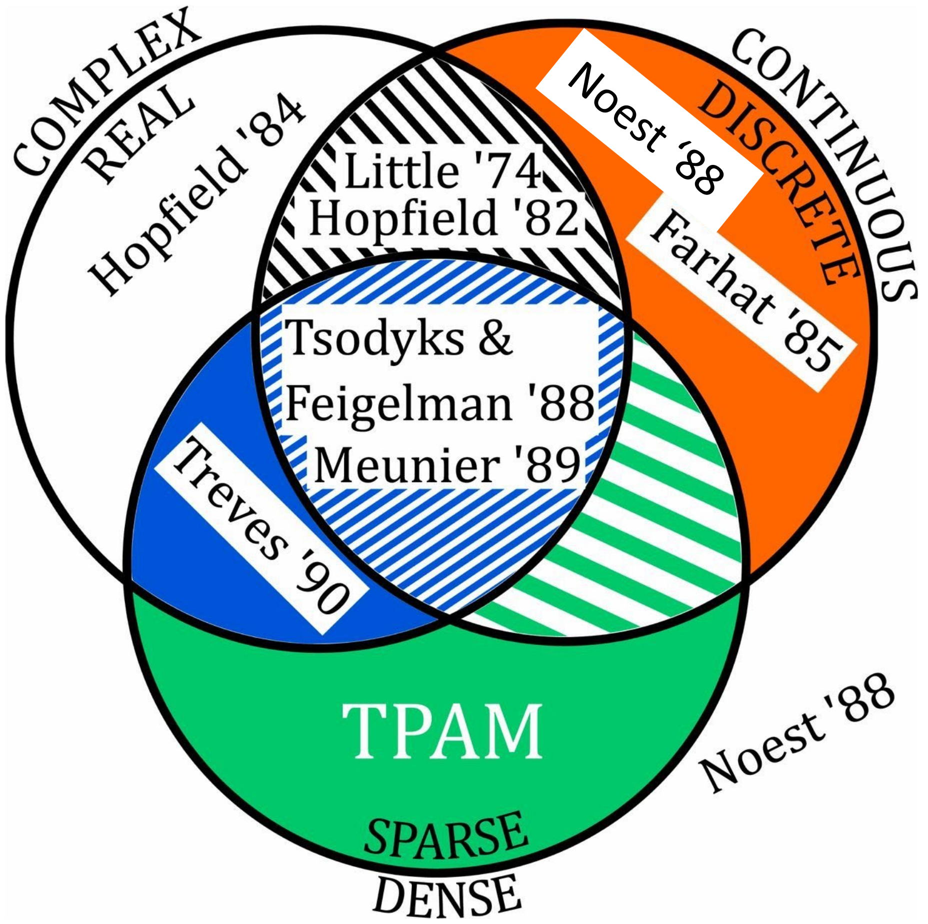
Pattern Capacity

$$\frac{M}{N} \quad \left[\frac{\text{Patterns}}{\text{Unit}}\right]$$

Information Capacity

$$\frac{I_p M}{S} \quad \left[\frac{\text{Bits}}{\text{Synapse}}\right]$$

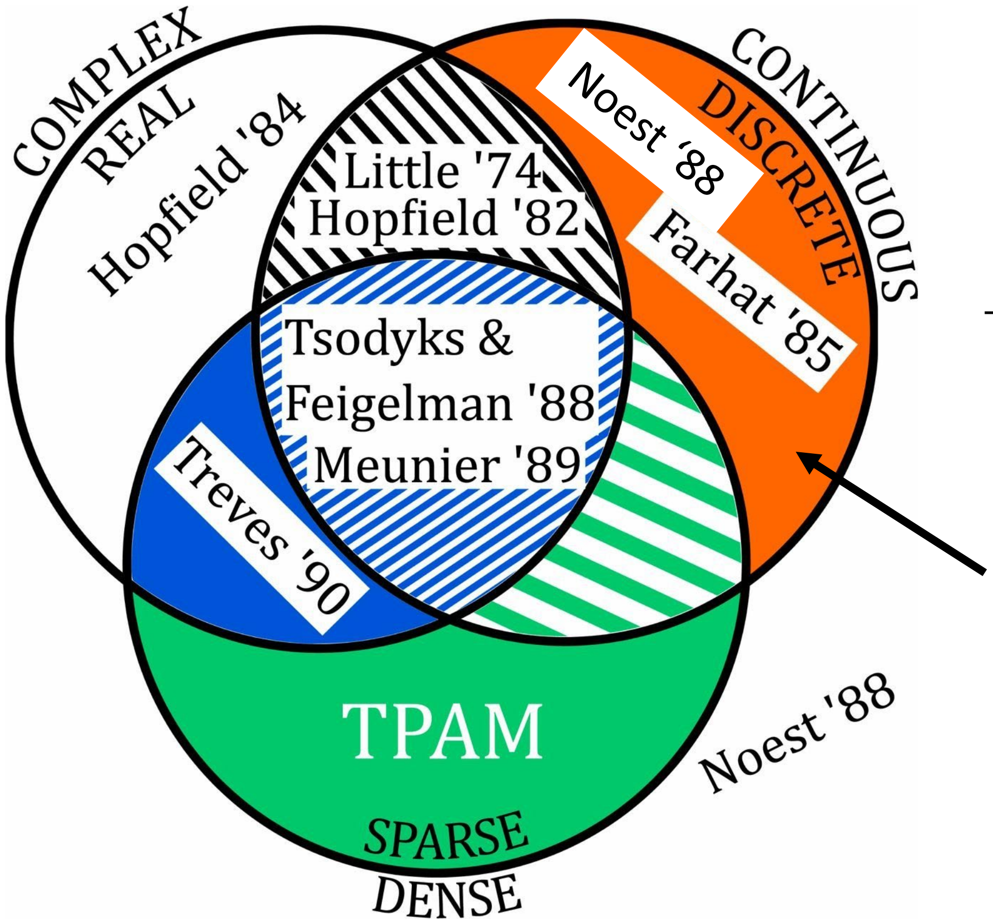
Existing Associative Memory Models



	Pattern Information	Pattern & Information Capacity
Binary & Dense & Discrete	High	Low
Sparse	Low	High
Complex & Dense & Continuous	High	Low

Frady, E. Paxon, and Friedrich T. Sommer. "Robust computation with rhythmic spike patterns." *Proceedings of the National Academy of Sciences* 116.36 (2019): 18050-18059.

Existing Associative Memory Models



	Pattern Complexity	Pattern & Information Capacity
Binary & Dense & Discrete	High	Low
Sparse	Low	High
Complex & Dense & Continuous	High	Low
Complex & Dense & Discrete	High	?

Frady, E. Paxon, and Friedrich T. Sommer. "Robust computation with rhythmic spike patterns." *Proceedings of the National Academy of Sciences* 116.36 (2019): 18050-18059.

Complex-Valued Phasor Associative Memory

Continuous Patterns
$\boldsymbol{\xi}^{\boldsymbol{k}} \in \mathbb{C}^N$

Complex Outer Product Learning Rule

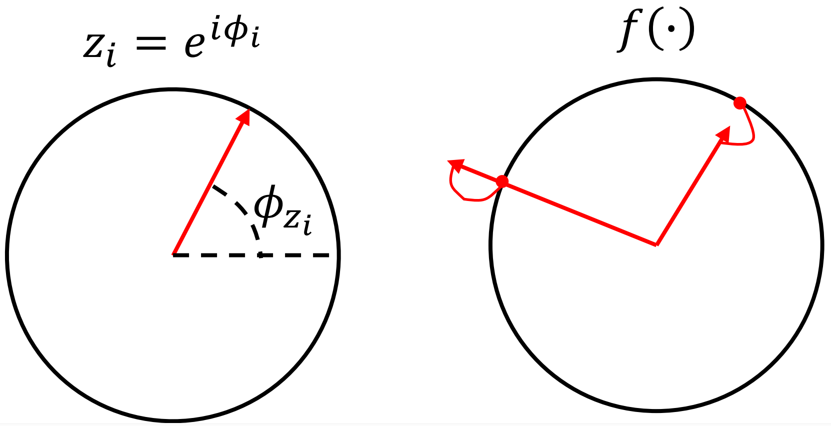
$$\boldsymbol{W} = \frac{1}{M}\sum_{k=1}^{M} \boldsymbol{\xi}^{\boldsymbol{k}}\left(\boldsymbol{\xi}^{k}\right)^{*\mathrm{T}}$$

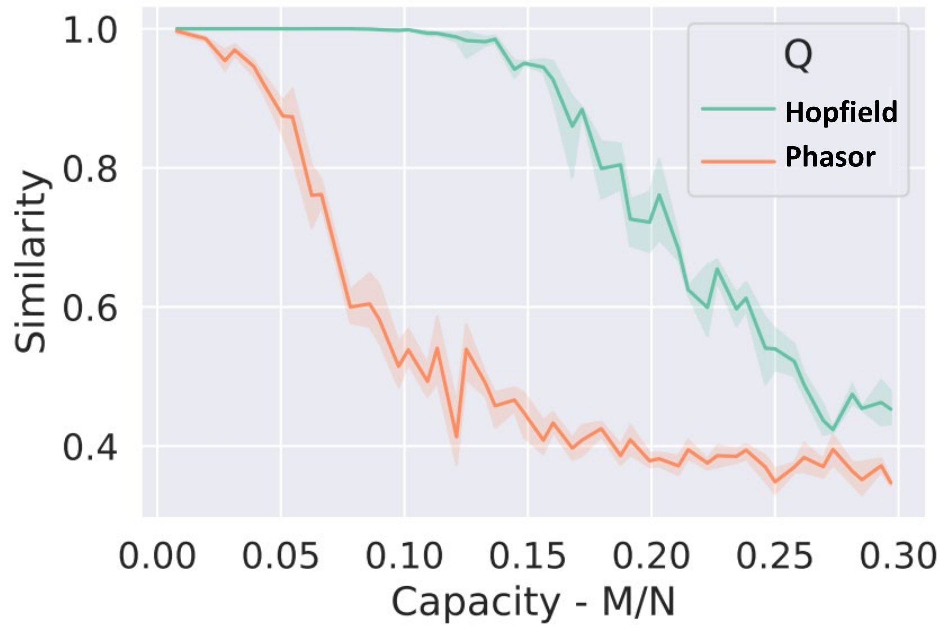
Energy
$\boldsymbol{E} = -(\boldsymbol{z})^{*\mathrm{T}}\boldsymbol{W}\boldsymbol{z}$

Dynamics

$$\boldsymbol{z}(t+1) = f\big(\boldsymbol{W}\boldsymbol{z}(t)\big)$$

$$f(z_i) = \frac{z_i}{|z_i|}$$





Noest, Andre. "Phasor neural networks." *Neural information processing systems*. 1987.

Q-State Phasor Neural Networks

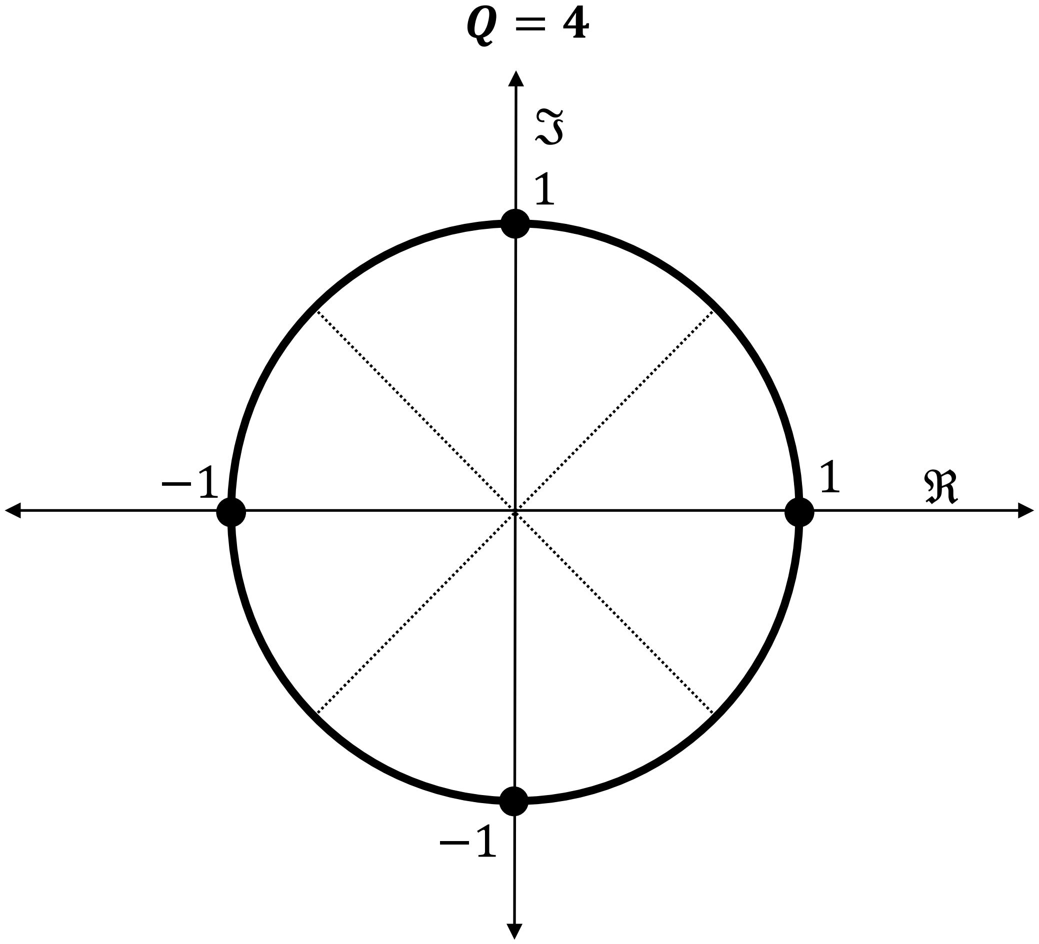
Patterns

$$\boldsymbol{\xi}^k = \exp\left(i\frac{2\pi \boldsymbol{q}^k}{Q}\right)$$

Dynamics

$$\boldsymbol{z}(t+1) = f\big(\boldsymbol{W}\boldsymbol{z}(t)\big)$$

$$f(u_i) = \exp\left(i\frac{2\pi}{Q}\operatorname{argmin}_q\left|\phi_i^{\mathrm{u}} - \frac{2\pi q}{Q}\right|\right)$$

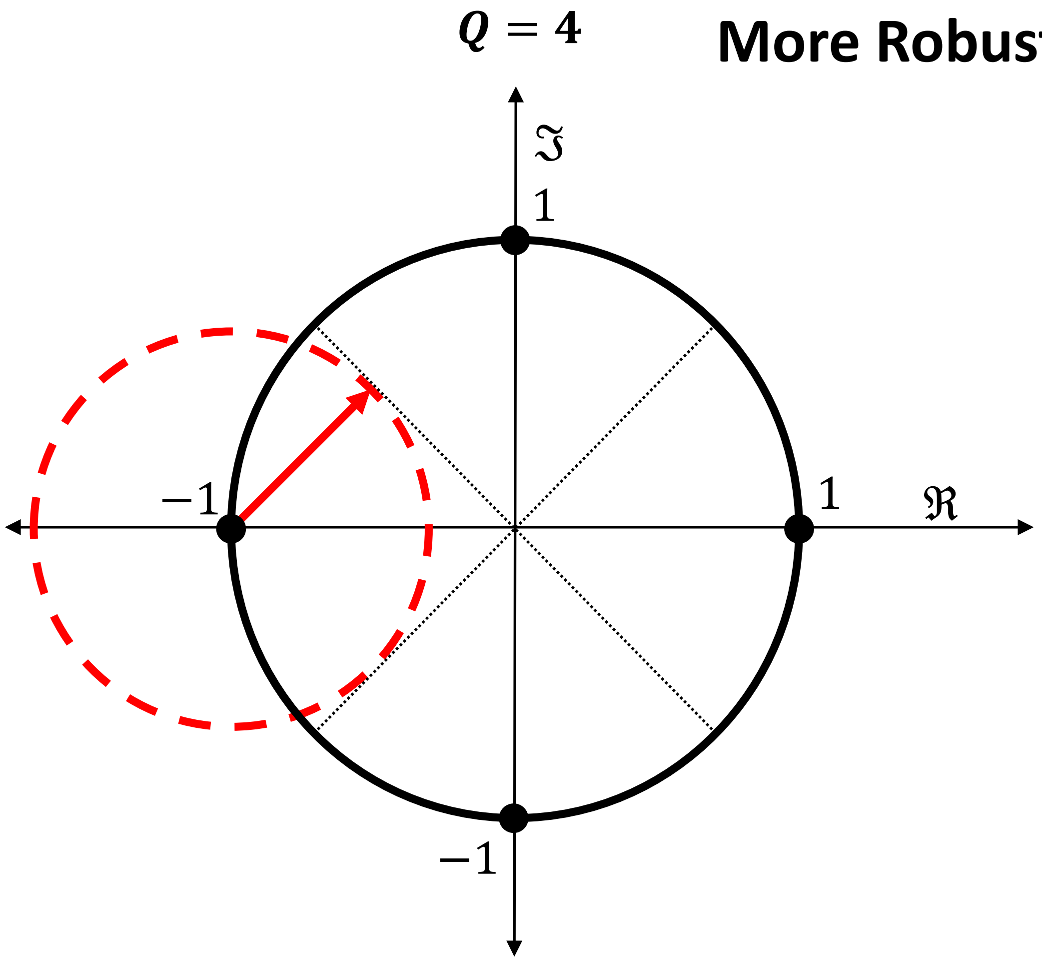


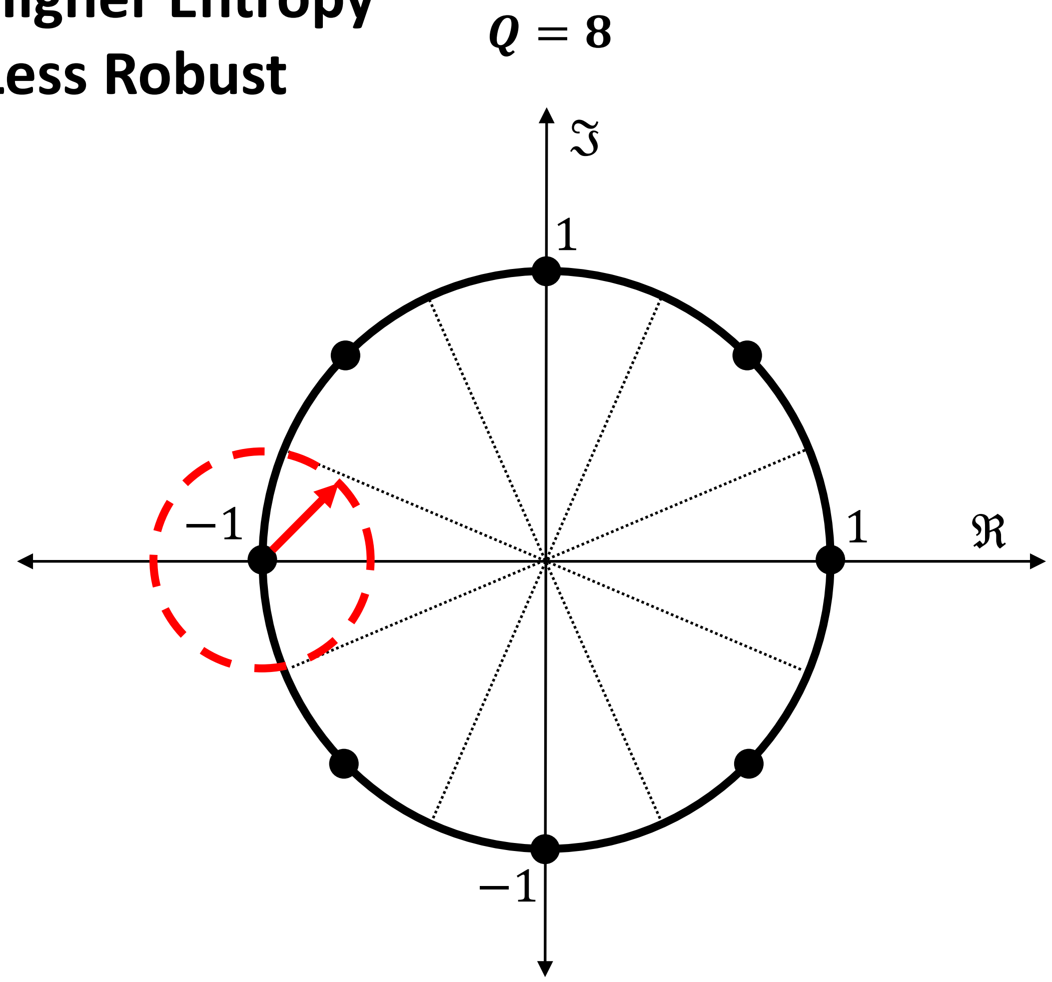
Noest, A. J. (1988). Discrete-state phasor neural networks. *Physical Review A*, *38*(4), 2196.

Trade-off Pattern Complexity vs. Error-Correction

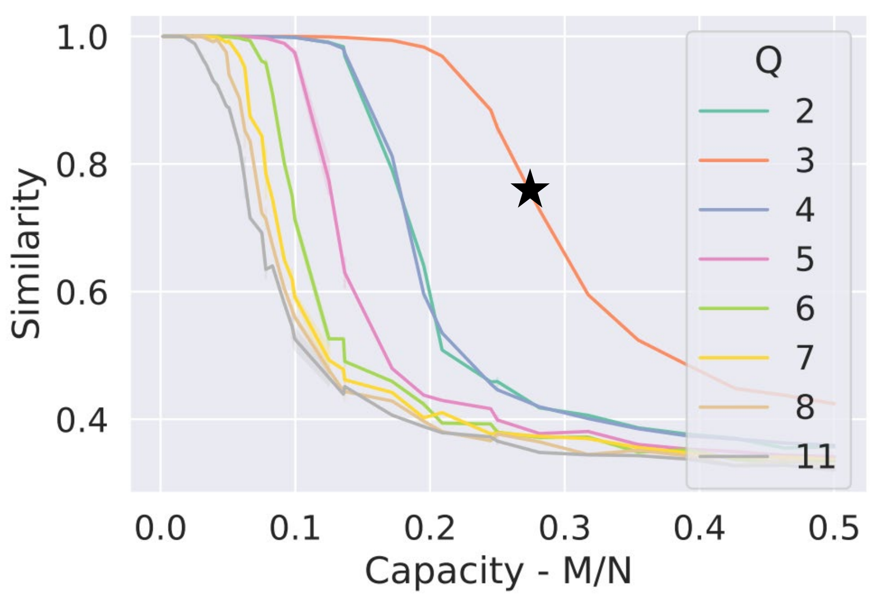
Lower Entropy
More Robust

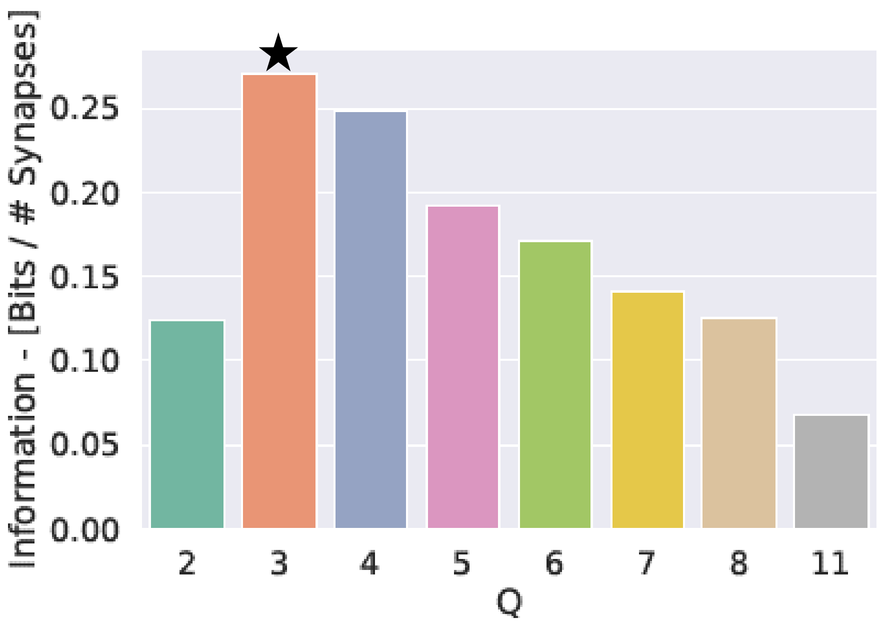
Higher Entropy
Less Robust





Capacity Results for Q-State Phasor Networks





★ Maximum capacity at Q=3 as predicted by mean-field theory - Cook, J. (1989)

Oscillator Networks

Mapping Associative Memories to Hardware

Synchronization in Weakly-Coupled Oscillators

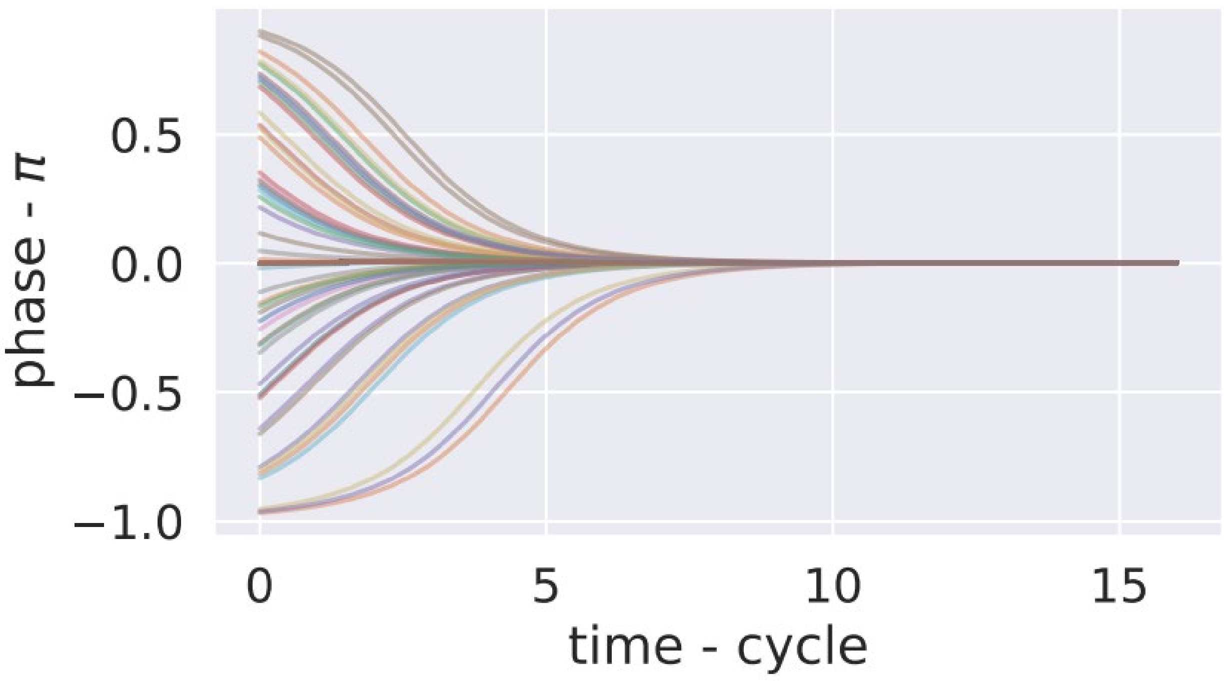
Continuous Phasors

$\boldsymbol{z}(t+1) = f(\boldsymbol{W}\boldsymbol{z}(t)); \; W_{ij} = 1$

$f(z_i) = \frac{z_i}{|z_i|}$

Kuramoto Model

$\dot{\phi}_i = \epsilon \sum_j \sin(\phi_j - \phi_i)$



Kuramoto Phasor Associative Memory

Phasor associative memories map to Kuramoto oscillator networks

Without state quantization fixed-points of dynamical system **ARE NOT*** stored patterns

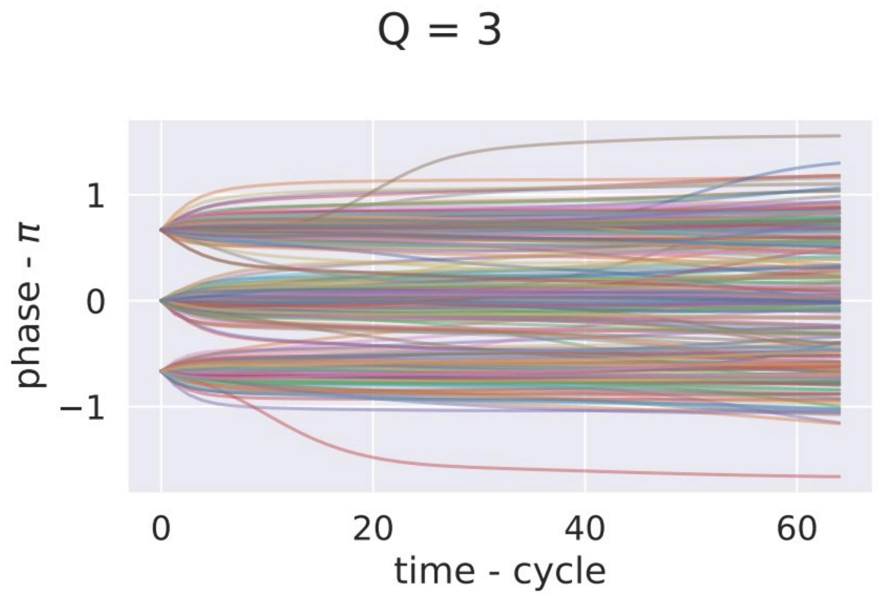
Dynamics

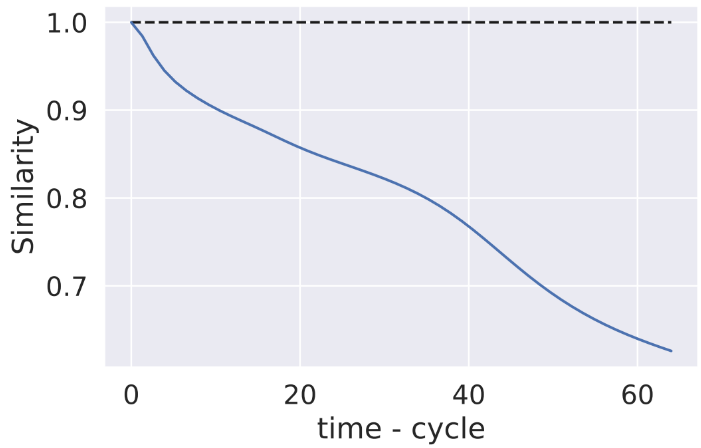
$$\dot{\phi}_i = \epsilon \sum_j R_{ij} \sin(\phi_j + \Phi_{ij} - \phi_i)$$

Parameters

$R_{ij} = |W_{ij}|$ Coupling Strength

$\Phi_{ij} = \arg(W_{ij})$ Coupling Phase Shift





*For M > 2

Phase Quantization

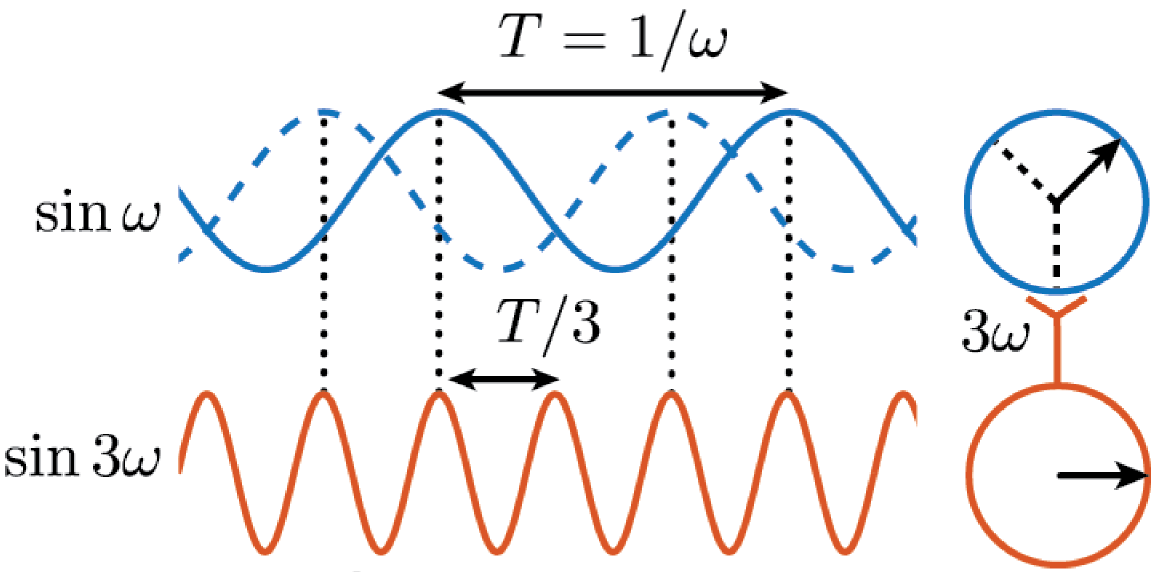
Phase Quantization

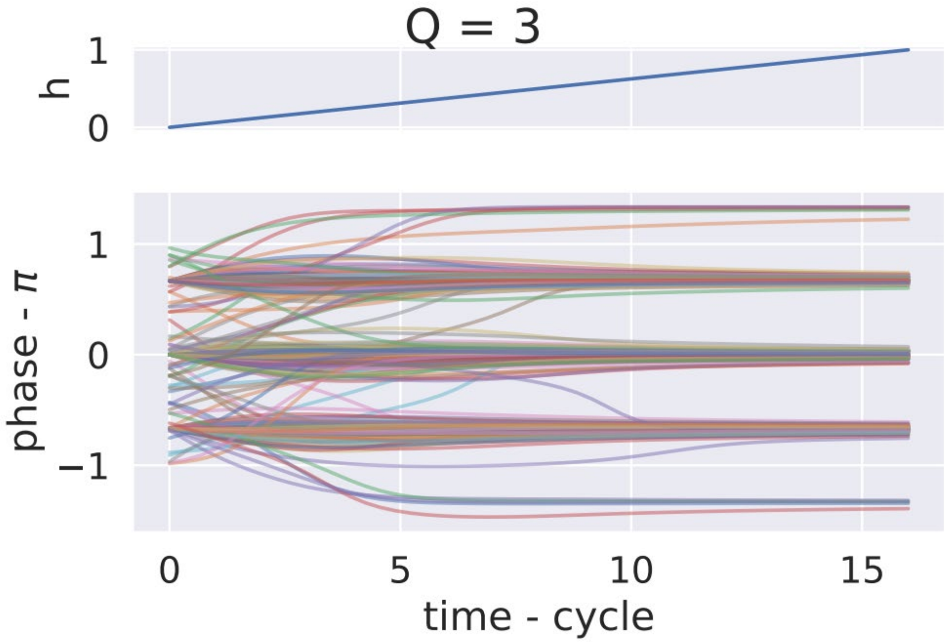
$$f(u_i) = \exp\left(i\frac{2\pi}{Q}\operatorname{argmin}_q\left|\phi_i^{\mathrm{u}} - \frac{2\pi q}{Q}\right|\right)$$

Harmonic Injection Locking (HIL)

$$\dot{\phi}_i = -\epsilon\frac{\partial E(\boldsymbol{\phi})}{\partial \phi_i} - h\sin(Q\phi_i - \phi_j)$$

$$Q = \frac{\omega_{inj}}{\omega_{mem}}$$ Harmonic Ratio



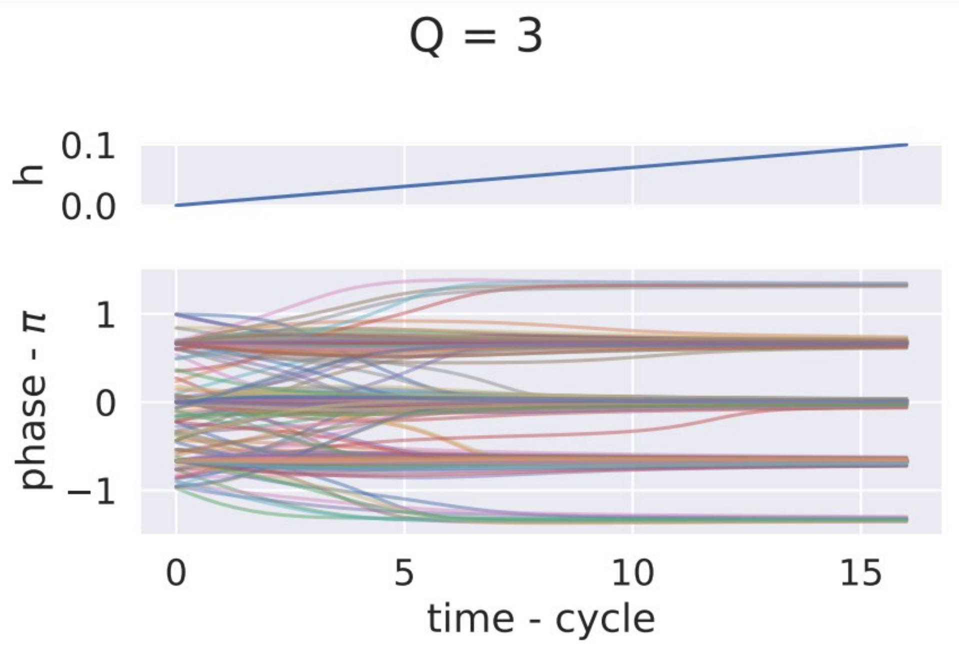


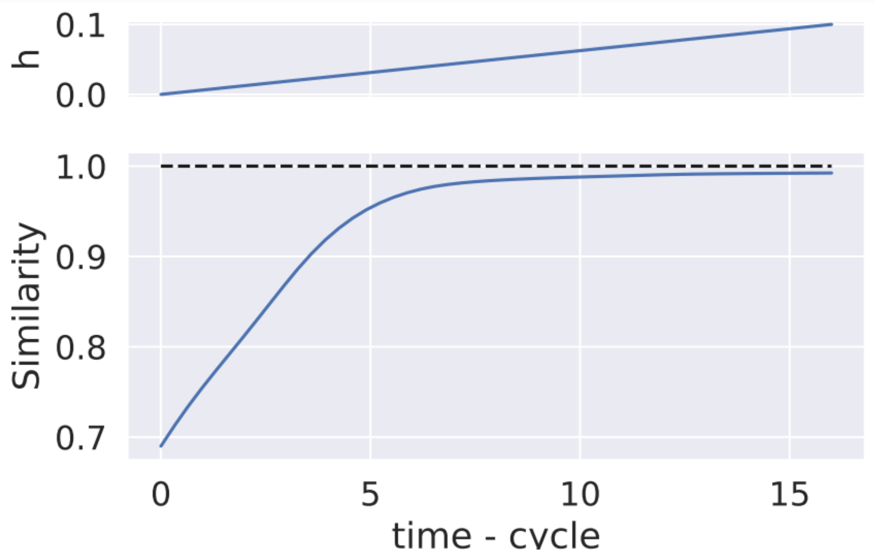
*Presented in Nishikawa et at. '92 for bipolar patterns

Result: Q-State Oscillator Network

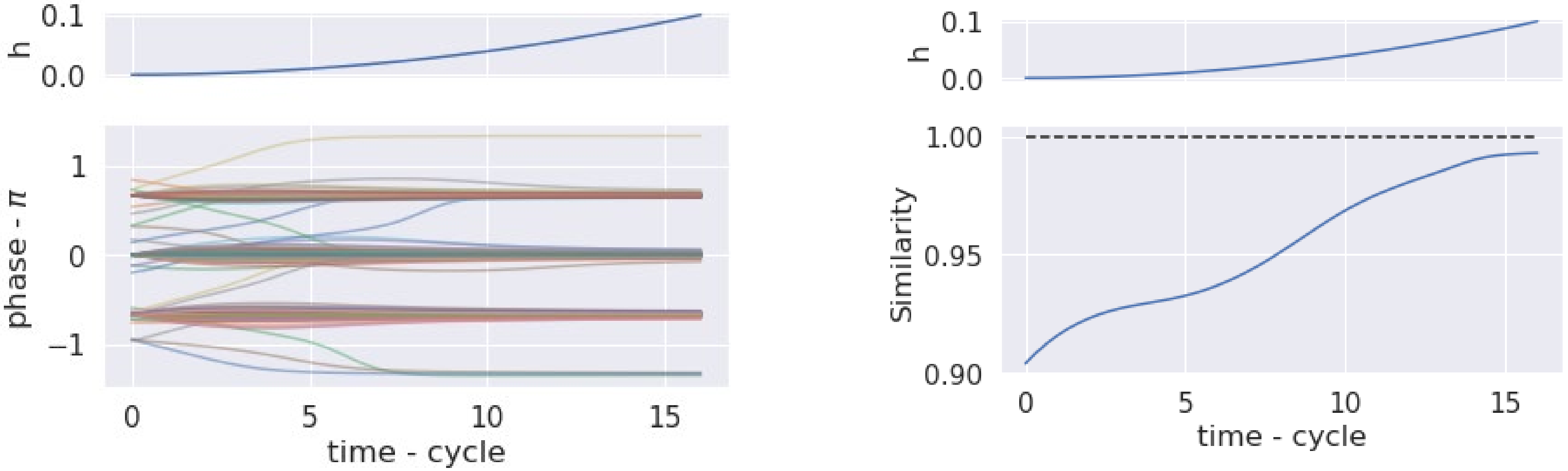
Complete System Dynamics

$$\dot{\phi}_i = \epsilon \sum_j R_{ij} \sin(\phi_j + \Phi_{ij} - \phi_i) - h \sin(Q\phi_i - \phi_j)$$



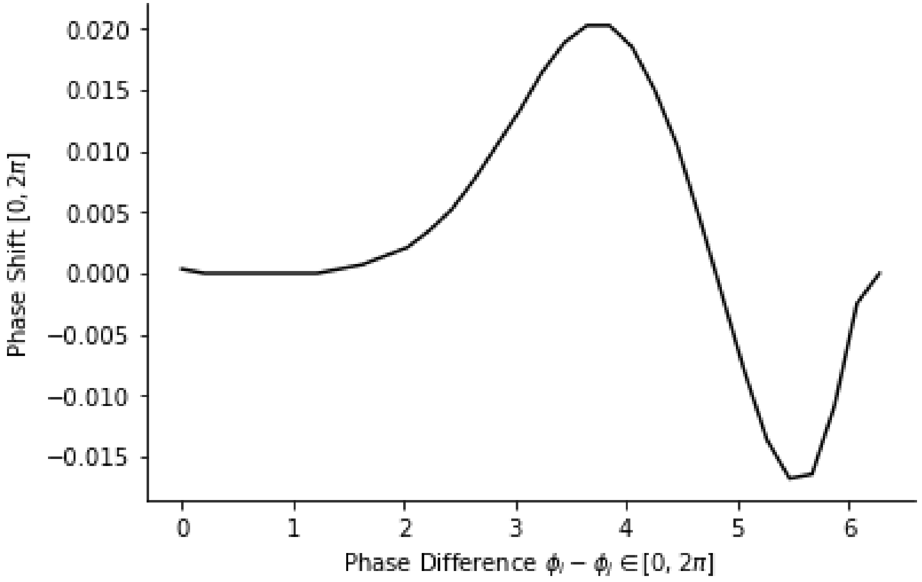


Capacity of Q-state Oscillator Models

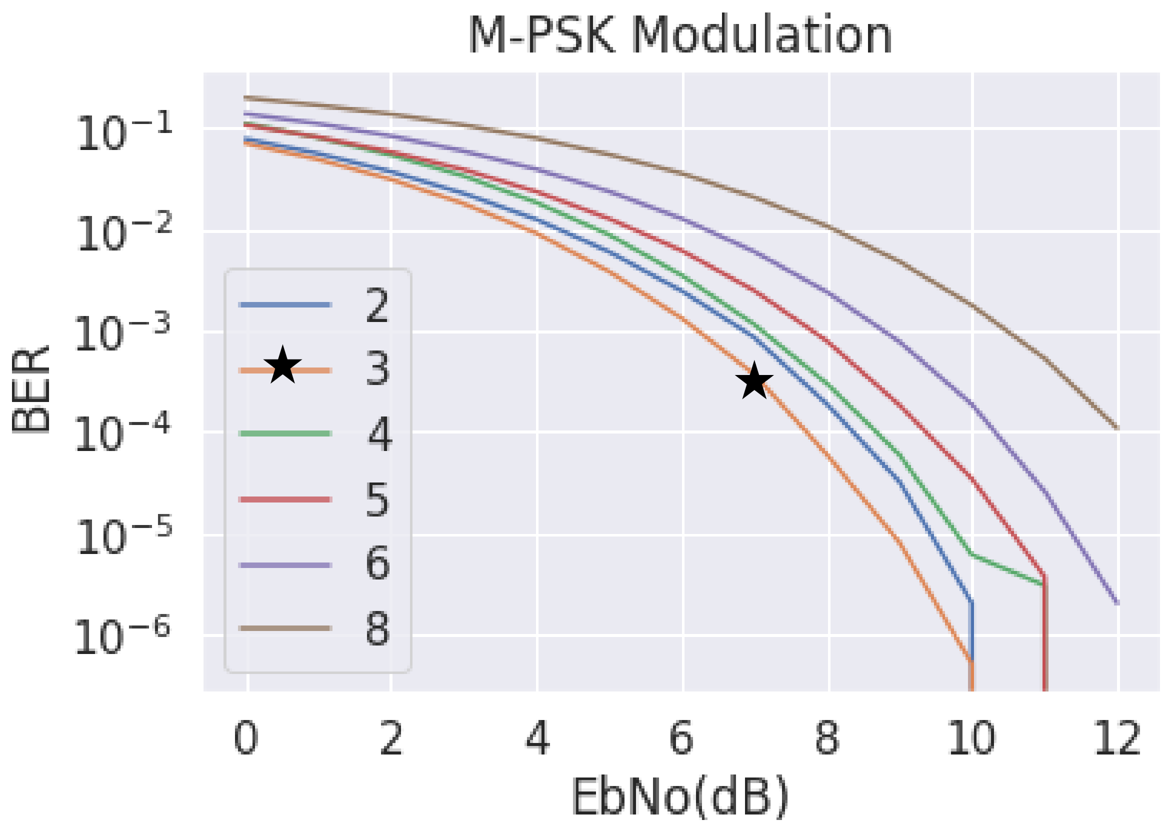


General Phase Coupling

$$\dot{\phi}_i = \sum_j C_{ij} g(\phi_j - \phi_i)$$



M-ary Phase Shift Keying (M-PSK)



★ Again, $Q = 3$ is optimal!!

Summary

Dense Hopfield associative memories in the literature have low capacity

Q-State Phasor Associative Memories achieve high capacity

Implementation of Q-state Phasor Associative Memories in couple oscillators with harmonic injection

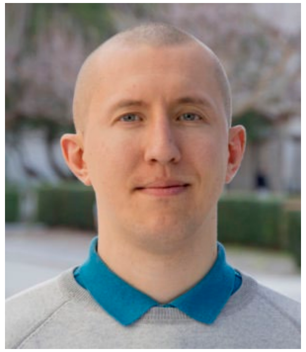
$Q = 3$ is best!

Acknowledgments

REDWOOD CENTER
for Theoretical Neuroscience


Fritz Sommer


Alex Belsten


Denis Kleyko


Paxon Frady


Bruno Olshausen


Dmitri Nikonov (Intel Labs)

Funding
NSF Graduate Research Fellowship Program
Intel Neuromorphic Computing Lab (NCL) Research Grant (INRC)

NSF GRFP intel