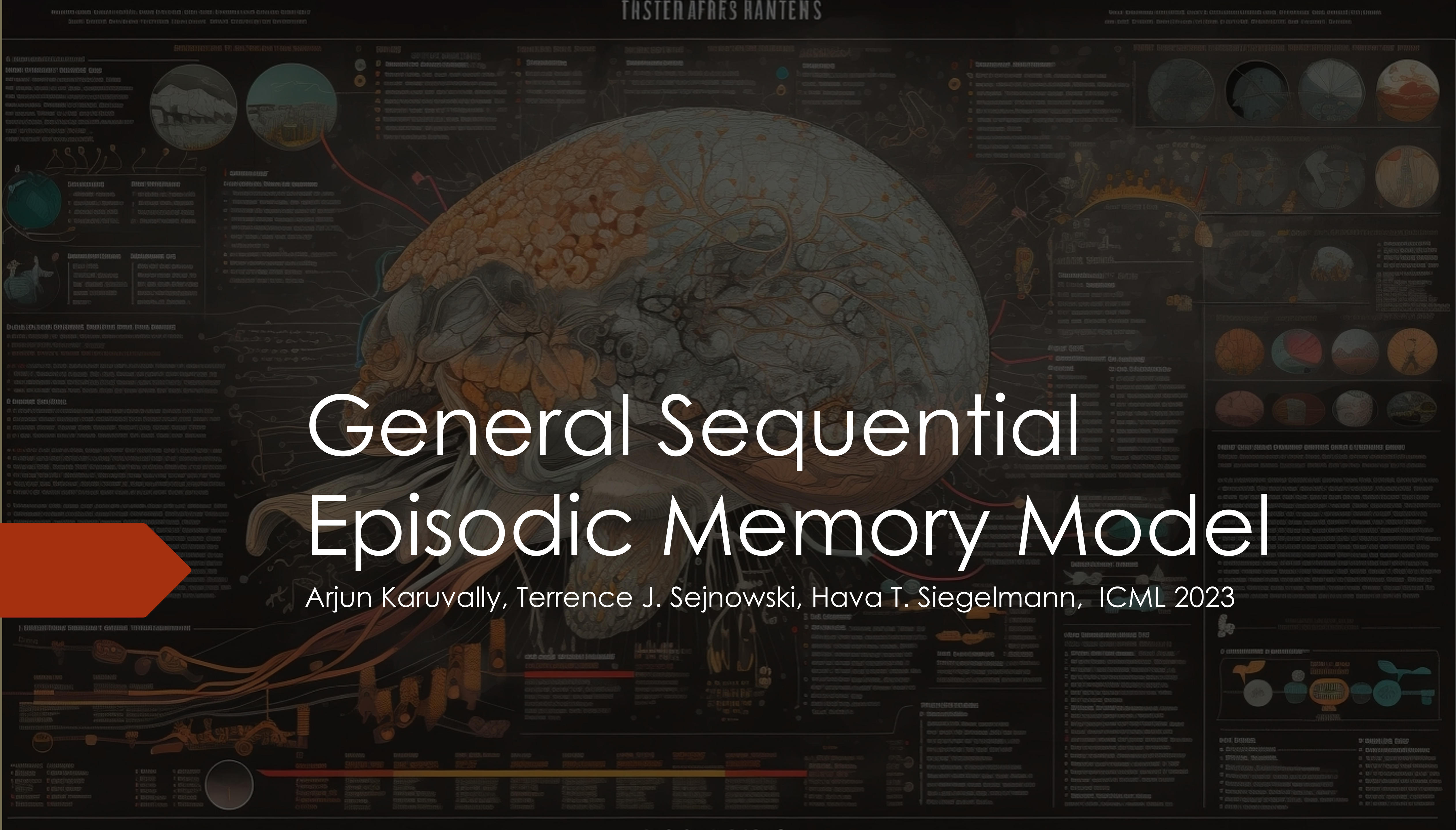


General Sequential Episodic Memory Model

Arjun Karuvally, Terrence J. Sejnowski, Hava T. Siegelmann, ICML 2023

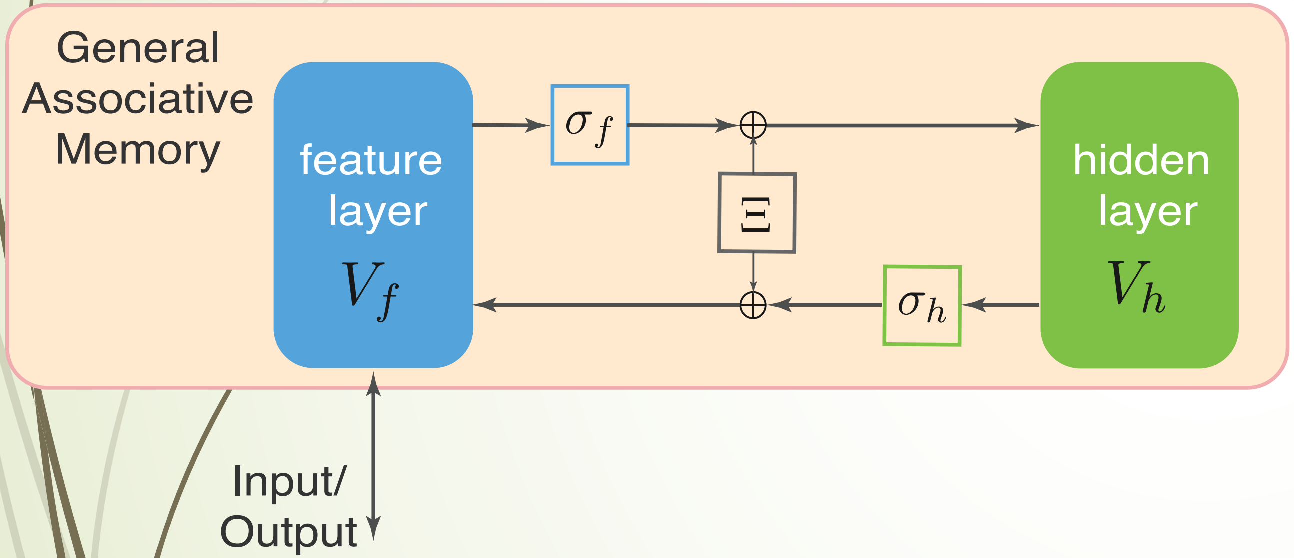
Outline

- General Associative Memory Model
- General Sequential Episodic Memory Model
  - Architecture
  - Energy function and dynamics
  - Sequence memory capacity

Outline

- **General Associative Memory Model**
- General Sequential Episodic Memory Model
  - Architecture
  - Energy function and dynamics
  - Sequence memory capacity

Generalized Associative Memory Model (GAMM)



Schematic diagram of network operations.

- *GAMM is a two-layer architecture of neurons with symmetric interactions.*
- *The symmetric interactions result in a dissipative dynamical system whose dynamical behavior can be explained using an energy function.*
- *Memories are the local minima states of the energy function.*
- *The contribution of each memory to the energy function depends on the interaction between the memories*
- *The hidden layer activation function defines how the stored memories (Ξ) interact resulting in models with different memory capacities.*
- *GAMM can store and retrieve singleton memories.*
- *Memory capacity relates the number of neurons in the feature layer to the number of memories that can be stored*

Singleton memory developments

classical Hopfield model (1982)

- Discrete states
- Discrete updates
- Linear memory interaction
- **Low capacity (0.14 N)**

Dense Associative Memory (2016)

- Discrete states
- Discrete updates
- Polynomial memory interaction
- **High capacity (N^{d-1})**

modern Hopfield networks (2017)

- Continuous states
- Discrete updates
- Exponential memory interaction
- **High capacity ($2^{N/2}$)**

general associative memory (2021)

- Continuous states
- Continuous updates
- Unified theory for associative memory explaining capacity by the hidden layer activation function

Sequence Memory developments

Kleinfeld (1986)

Multi-timescale networks (2019)

- Discrete states
- Linear memory interaction
- Discrete updates
- **Low capacity (0.14 N)**

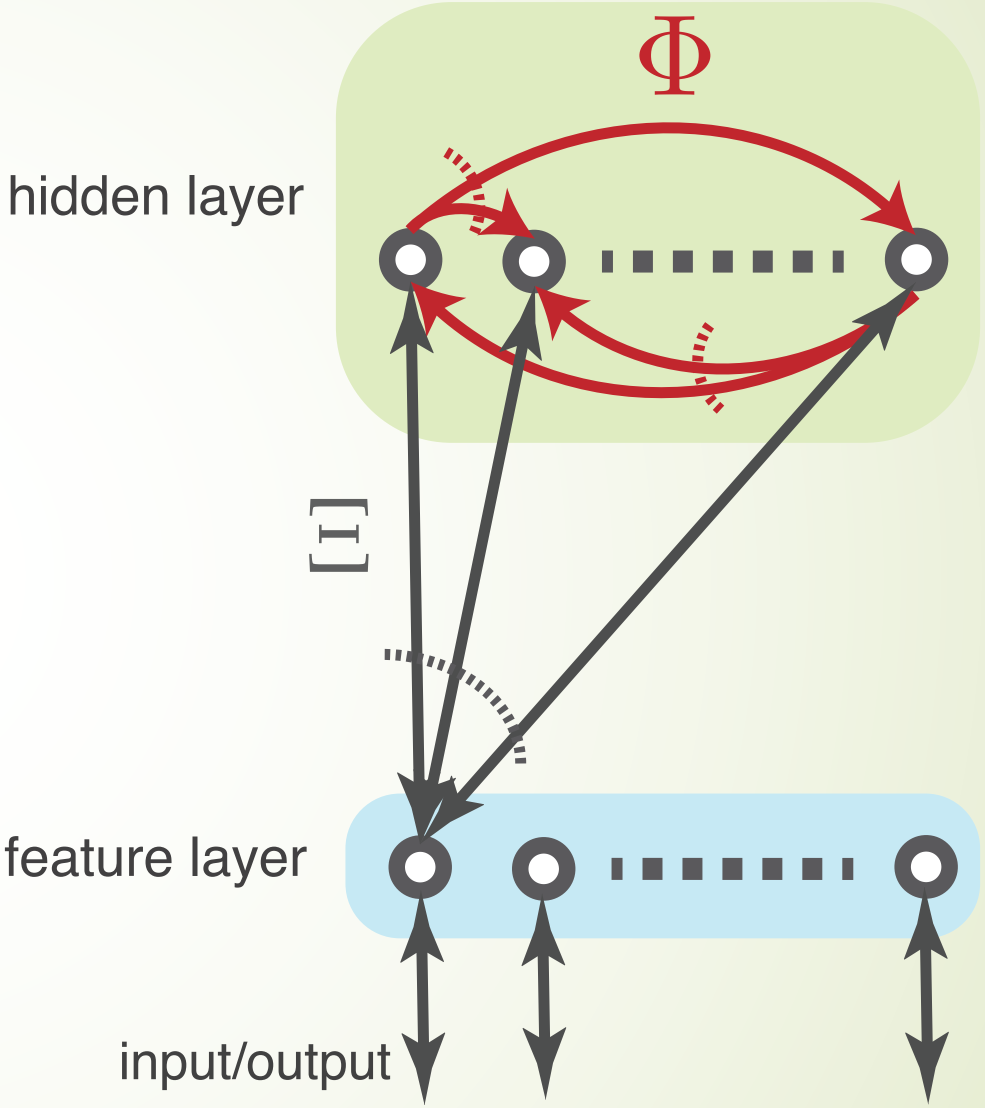
- Continuous states
- Continuous updates
- **Low capacity (0.14 N)**

Outline

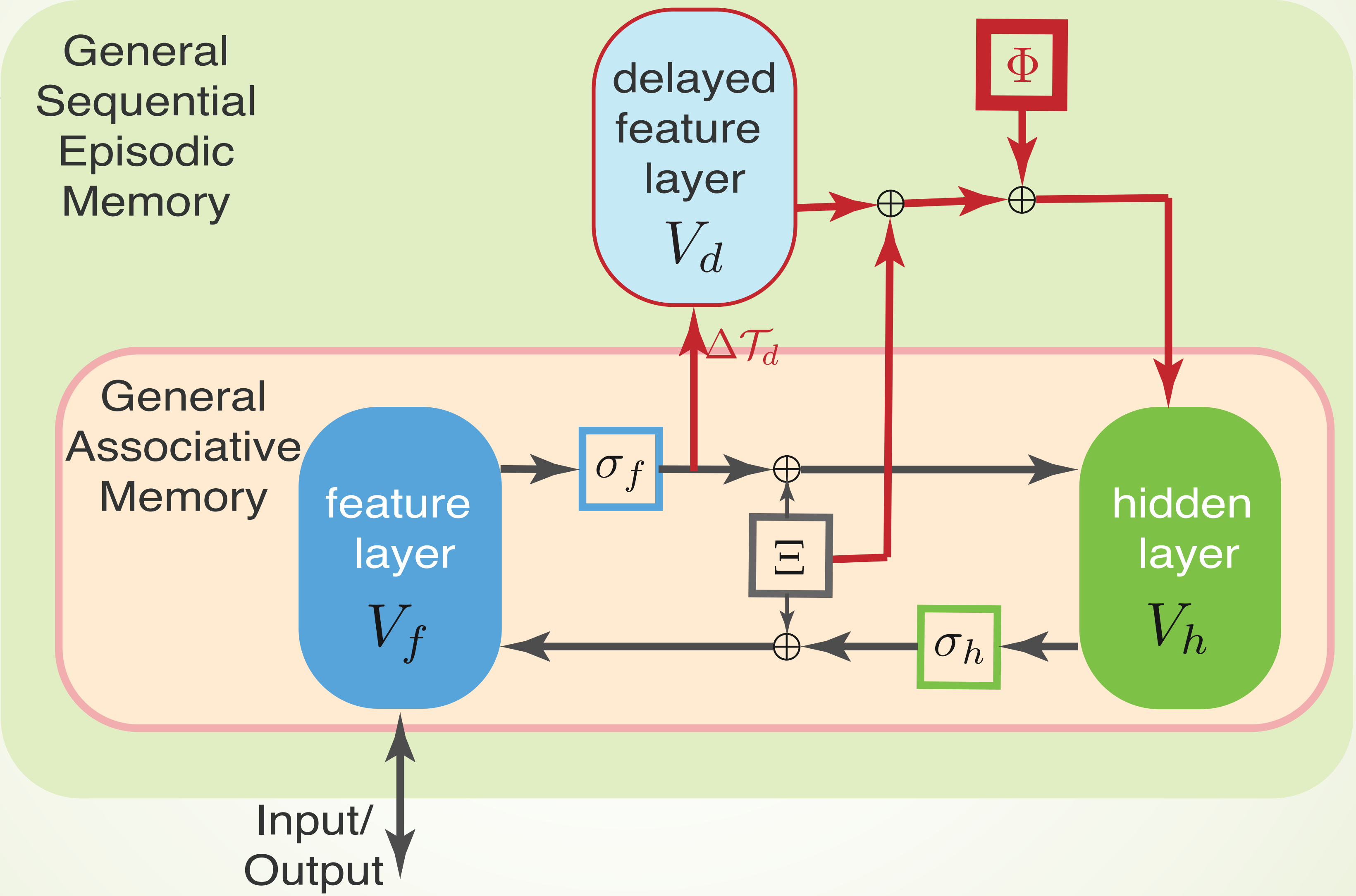
- General Associative Memory Model
- General Sequential Episodic Memory Model
  - **Architecture**
  - Energy function and dynamics
  - Sequence memory capacity

GSEMM neural architecture

- Sequential Episodic Memory (SEM) retrieval requires the ability to store and retrieve temporally related memories.
- Traditional energy-based models are restricted to single memory retrieval.
- General Sequential Episodic Memory Model (GSEMM) has a dynamic energy landscape capable of storing and retrieving of temporal memories in its dynamical evolution.
- The system consists of two layers of neurons organized analogous to General Associative Memory Model (GAMM) with additional delay-based intra-hidden layer connectivity.
- In contrast to GAMM, GSEMM exhibits a dynamic energy surface controlled by the delay signal resulting in a system with *instantaneous* fixed-point dynamics



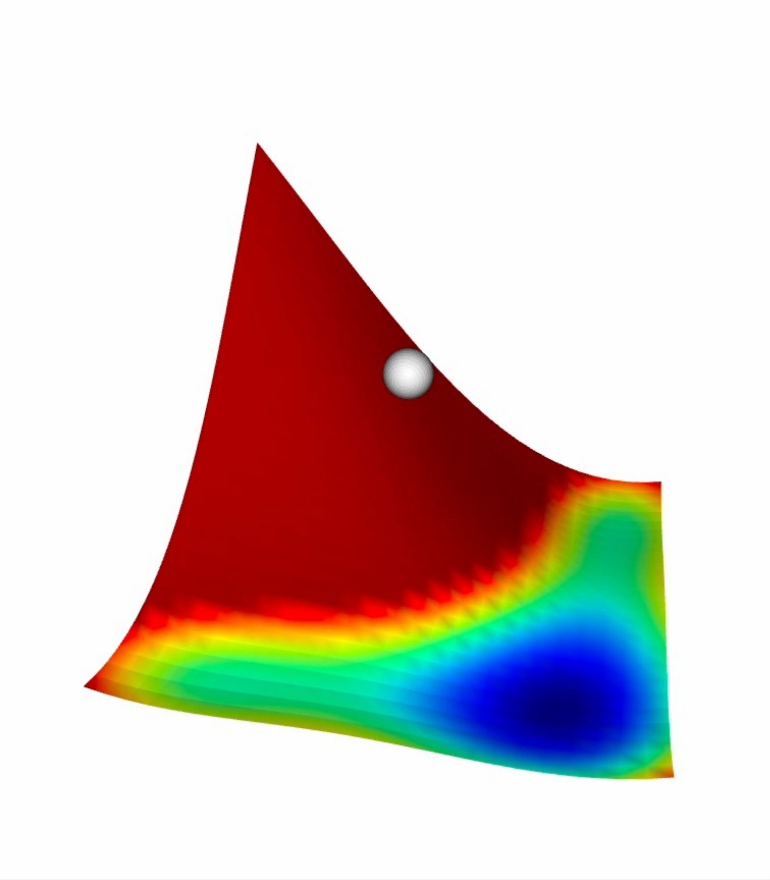
Schematic Representation of GSEMM operation

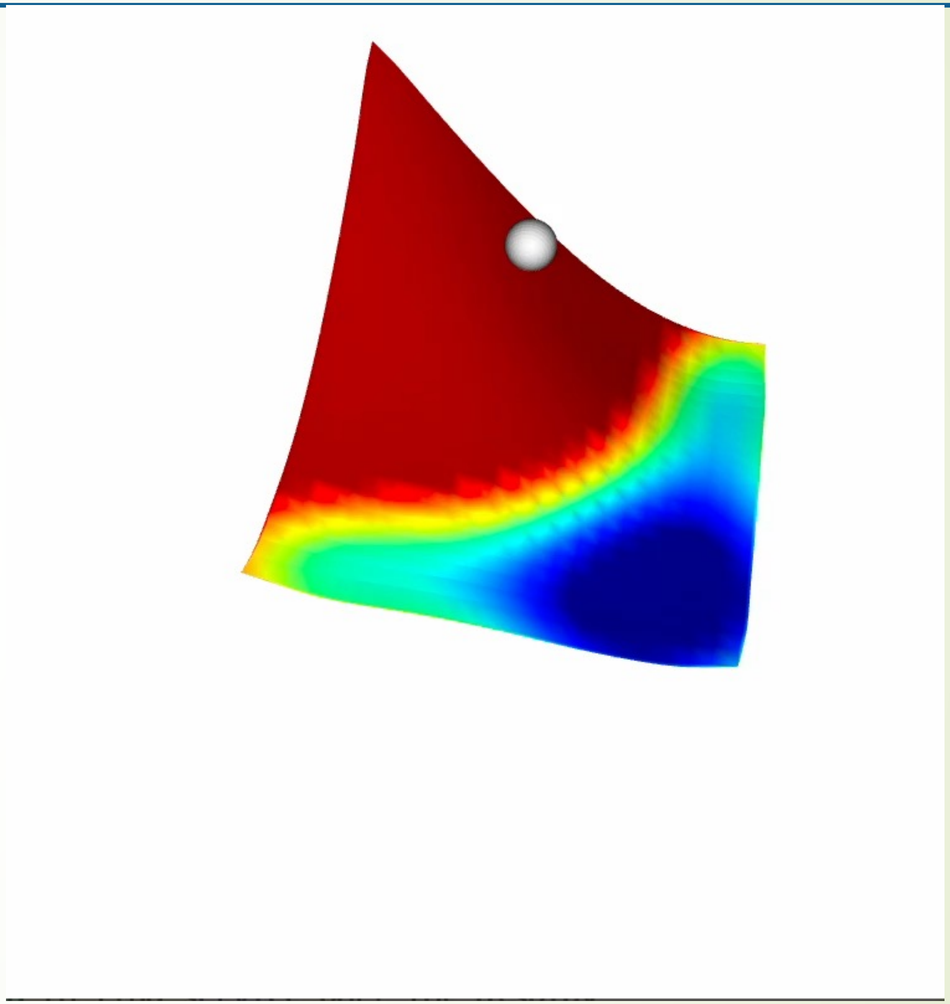


Outline

- General Associative Memory Model
- General Sequential Episodic Memory Model
  - Architecture
  - **Energy function and dynamics**
  - Sequence memory capacity

Energy Dynamics of GAMM vs GSEMM





- GAMM has a single attractor state the system converges to.
- When the delay is sufficiently high in GSEMM, it has a sequence of meta-stable attractors that the system visits in the dynamical evolution.

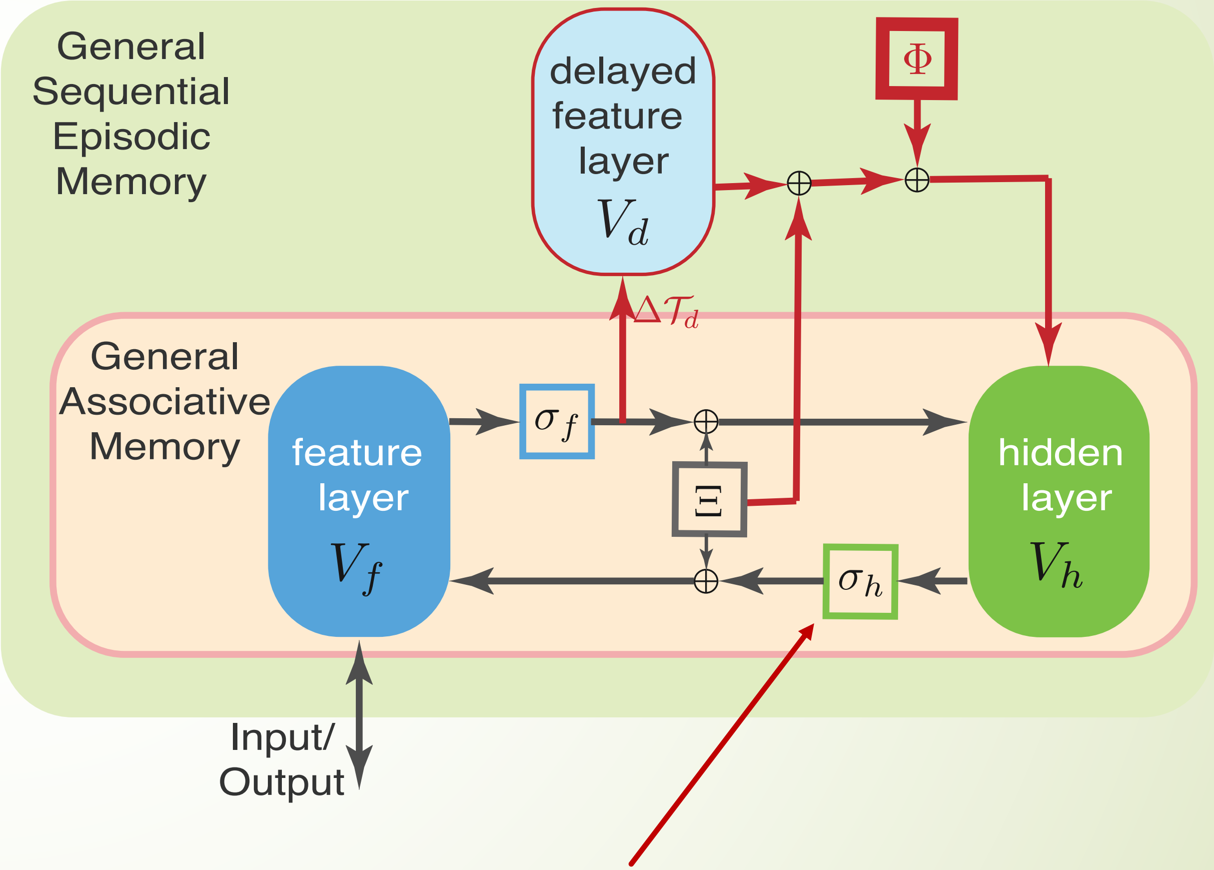
Practical Models from GSEMM

Linear GSEMM

- Analogous to current multi timescale sequence memory models.
- Derived by considering non-linear activation function for the feature layer and the identity activation function for the hidden layer
- In this setting, the stored memories interact linearly.

Dense GSEMM

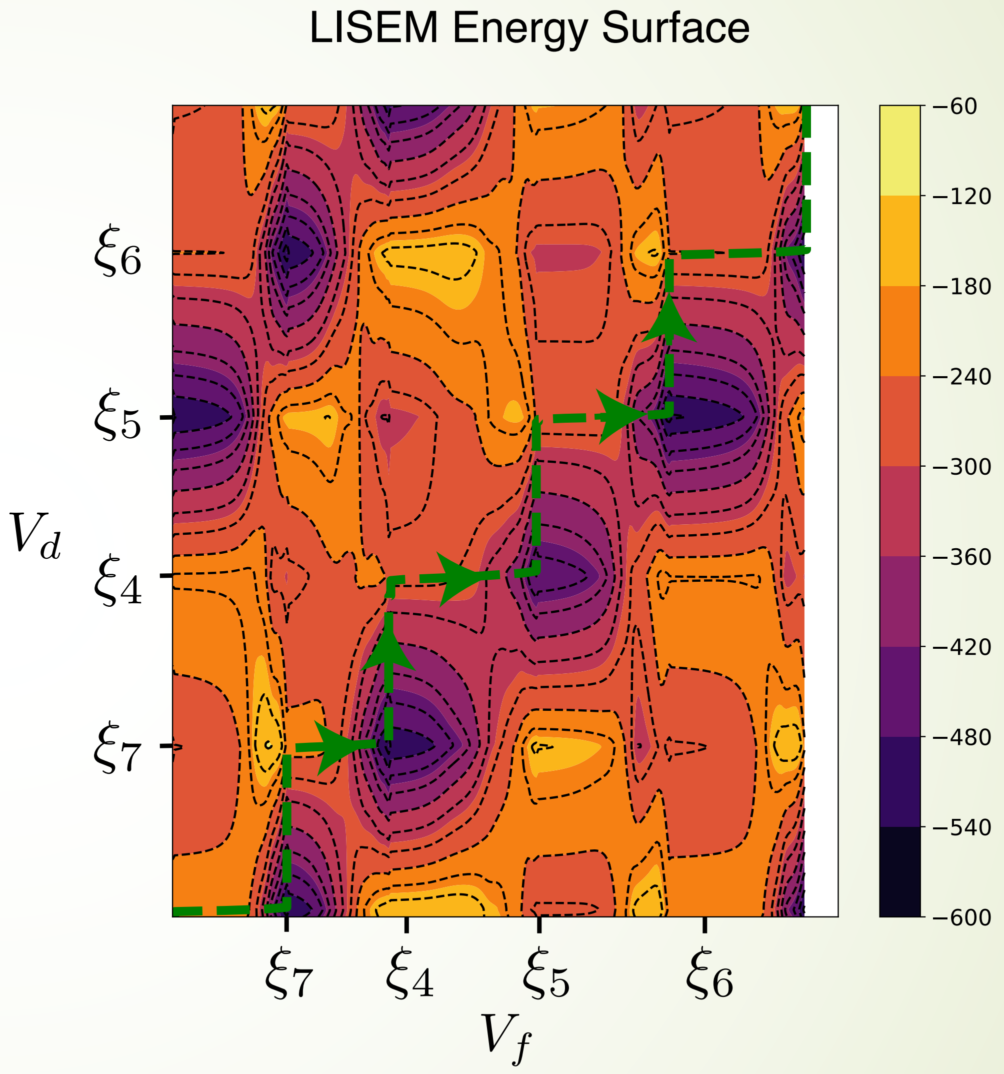
- Introduce polynomial non-linearity in the hidden layer activation.
- Stored memories interact non-linearly resulting in capacity improvements over LISEM.
- Note that LISEM is the degree 1 case of Dense GSEMM



Linear for Linear GSEMM
Non-linear for Dense GSEMM

Representative Example – Storage and Retrieval

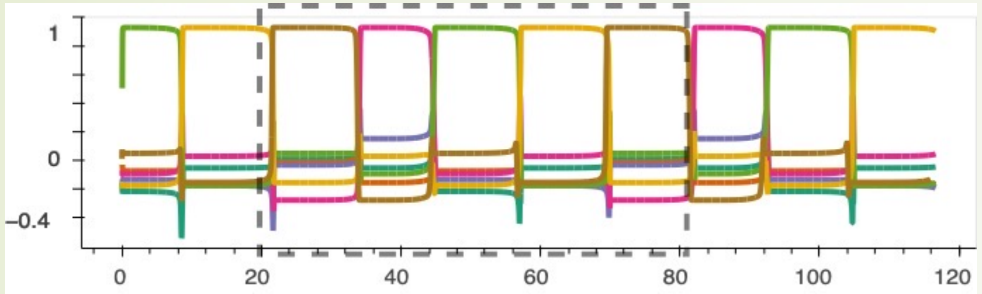
- Storage of sequence memory cycle - $\xi_4 \rightarrow \xi_5 \rightarrow \xi_6 \rightarrow \xi_7 \rightarrow \xi_4$ in the presence of 3 other memories
- Each memory ξ_i is a random binary vector (-1 or +1) of size 100 preloaded in the synapses (Ξ).
- The sequential relationship (the temporal connection between memories) is stored as an adjacency matrix in Φ
- The energy surface diagram shows how V_d controls the minima of the energy surface



Representative Example – Energy Dynamics

Output dynamics of LISEM

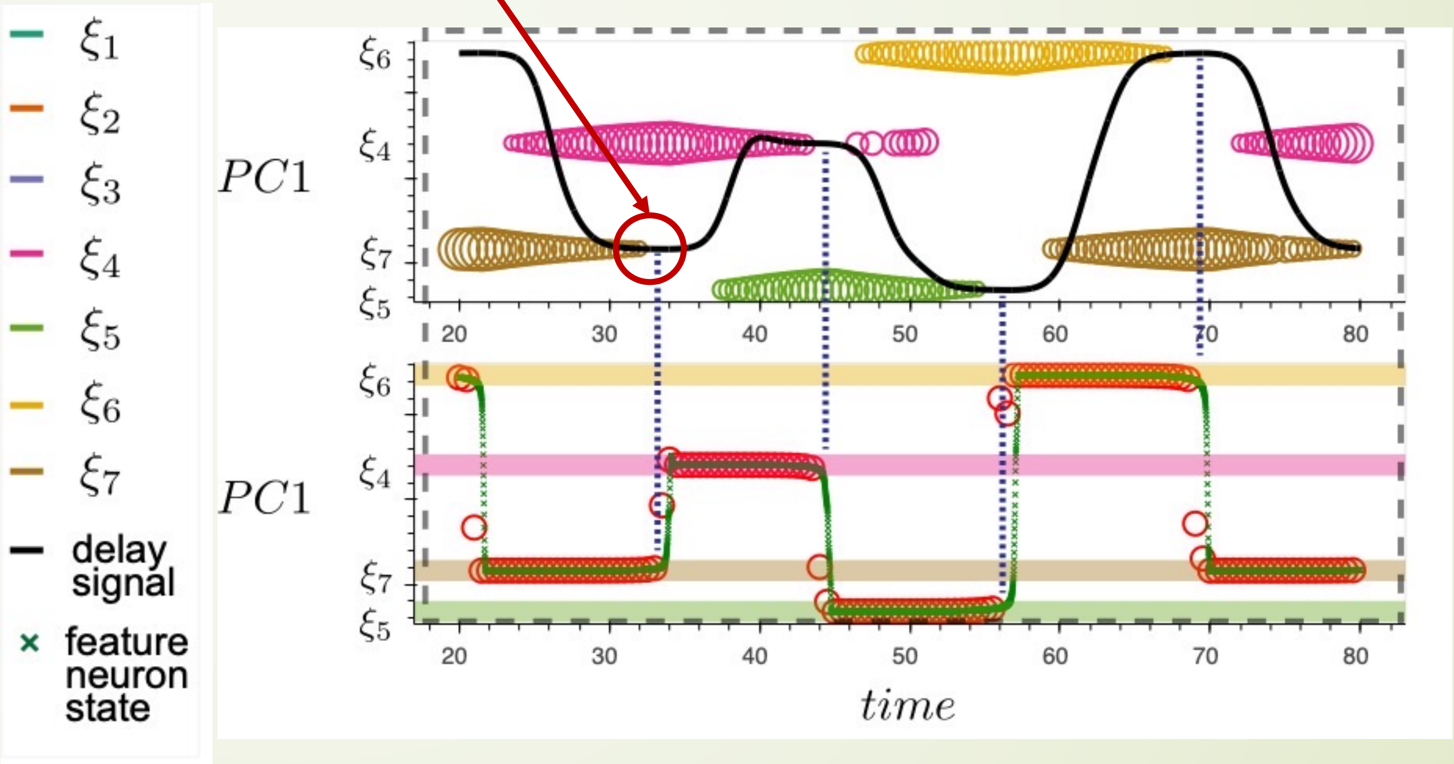
Memory overlap



- The first principle components of delayed neurons show how it influences the fixed points near all stored memories. Fixed points are depicted by colored circles.
- The size of the circle is inversely proportional to the memories' energy – higher the energy, smaller the circle.
- (top) The first principle component of the dynamical evolution of the nearest fixed point (red circle) of the energy surface and the current state of the fast sub-system (green cross).
- (bottom) The figure shows how the system is attracted to the nearest fixed point from the current state of the energy surface at each point in time.

Point where ξ_7 loses stability and the system is attracted to the basin of ξ_4

Attractor dynamics

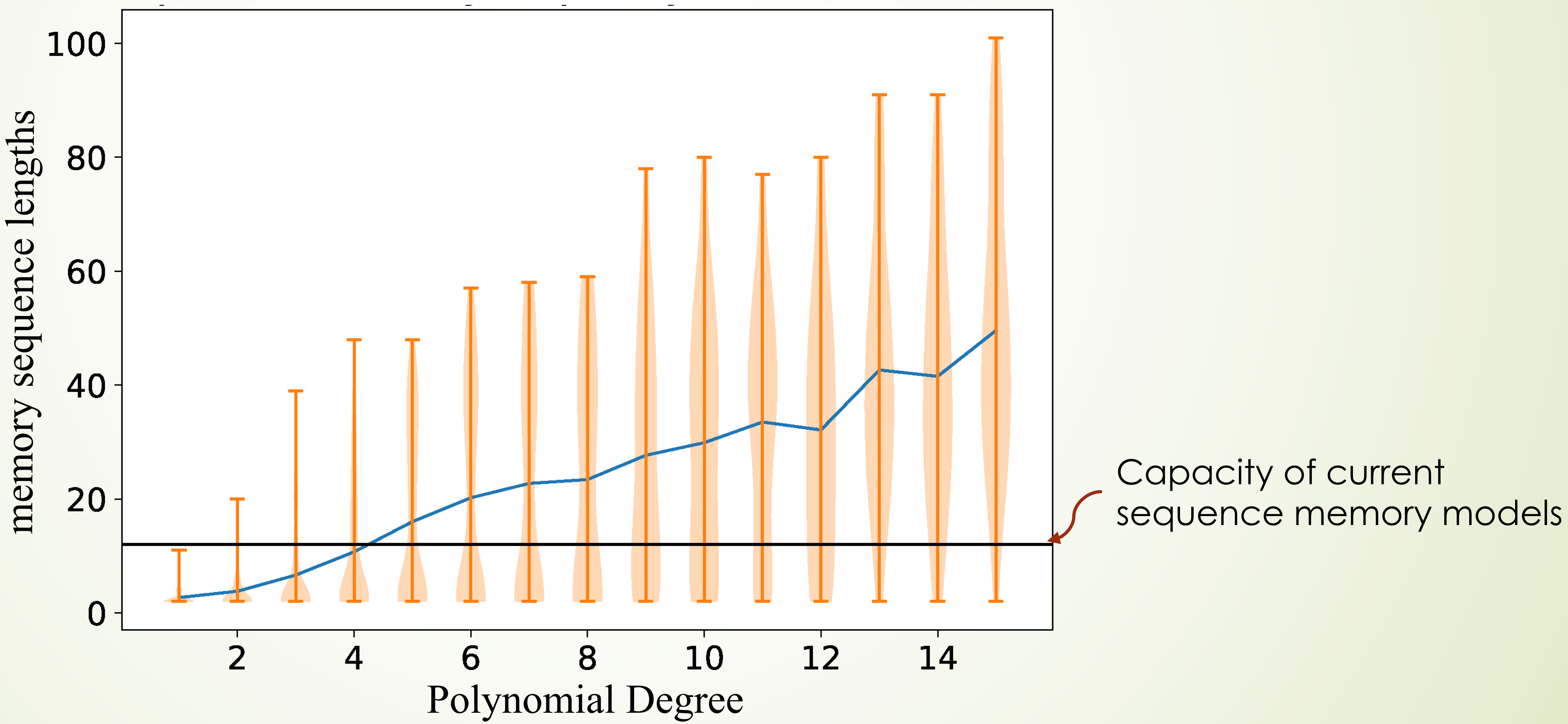


Outline

- General Associative Memory Model
- General Sequential Episodic Memory Model
  - Architecture
  - Energy function and dynamics
  - **Sequence memory capacity**

Exponential Capacity Improvement in Dense GSEMM

- Comparison of Linear GSEMM and Dense GSEMM based on their ability to store and retrieve cyclical memories with varying sequence lengths in a 100-neuron system.



Questions?

References

- Hopfield, John J.. “Neural networks and physical systems with emergent collective computational abilities.” Proceedings of the National Academy of Sciences of the United States of America 79 8 (1982): 2554-8 .
- Chen, H. H. et al. “High order correlation model for associative memory.” (1987).
- Krotov, Dmitry and John J. Hopfield. “Dense Associative Memory for Pattern Recognition.” *NIPS* (2016).
- Demircigil, Mete et al. “On a Model of Associative Memory with Huge Storage Capacity.” Journal of Statistical Physics 168 (2017): 288-299.
- Ramsauer, Hubert et al. “Hopfield Networks is All You Need.” *ArXiv* abs/2008.02217 (2020): n. pag.
- Krotov, Dmitry and John J. Hopfield. “Large Associative Memory Problem in Neurobiology and Machine Learning.” *ArXiv* abs/2008.06996 (2020): n. pag.
- Kleinfeld, David. “Sequential state generation by model neural networks.” *Proceedings of the National Academy of Sciences of the United States of America* 83 24 (1986): 9469-73 .
- Kurikawa, Tomoki et al. “Repeated sequential learning increases memory capacity via effective decorrelation in a recurrent neural network.” *ArXiv* abs/1906.11770 (2019): n. pag.