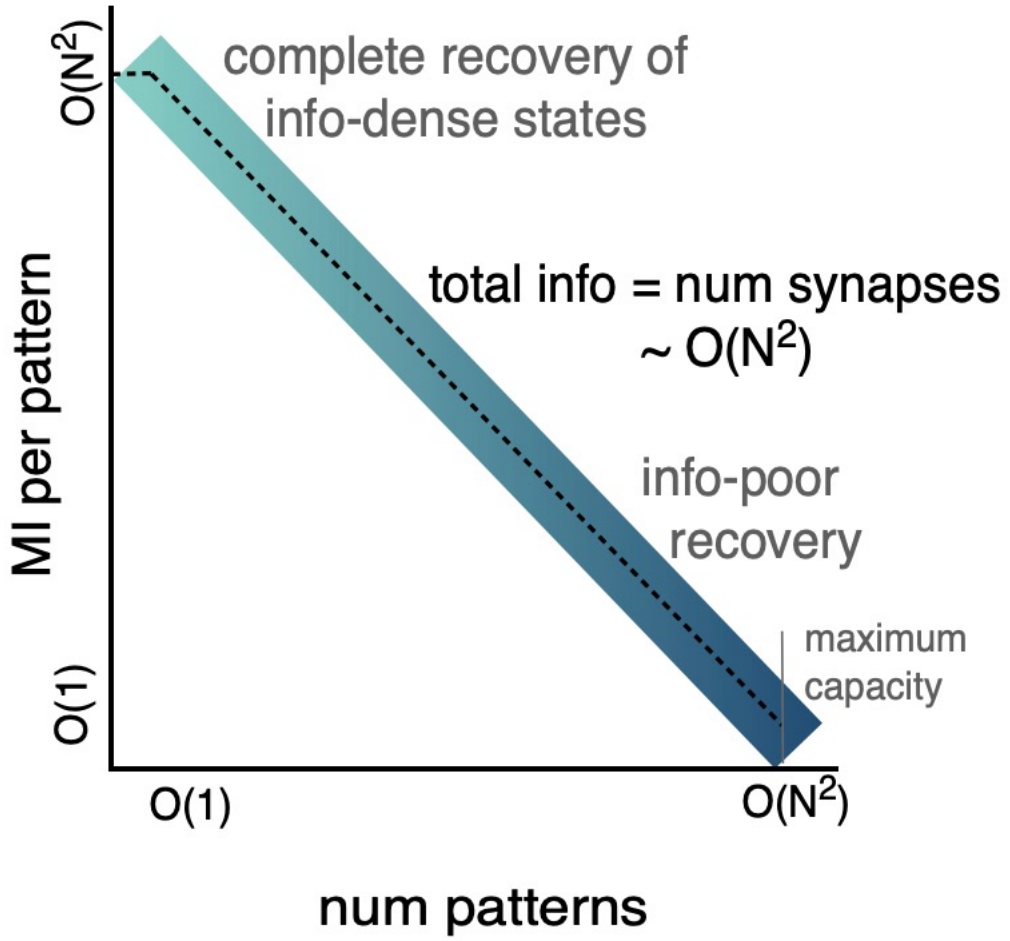
Content Addressable Memory Without Catastrophic Forgetting by Heteroassociation with a Fixed Scaffold

Sugandha Sharma, Sarthak Chandra*, Ila Fiete*

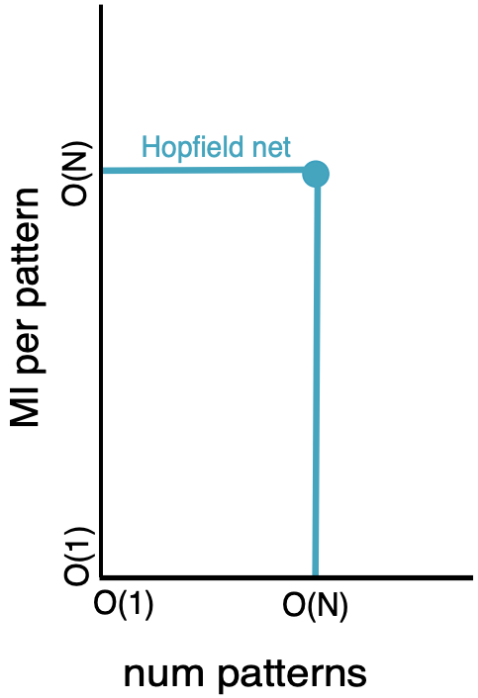
Sugandha Sharma
July 19, 2022

Any memory model should exhibit a Continuum



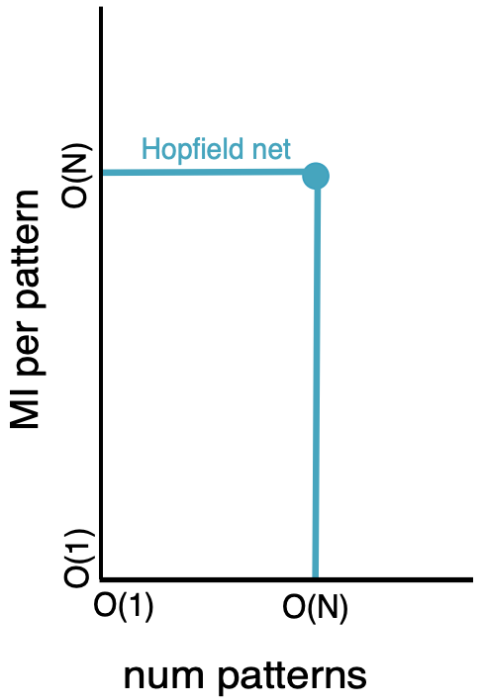
- Memories should only be gradually forgotten
- Key is to recover all stored memory states but with a lower information (accuracy) per memory

Existing CAM models lack a memory continuum

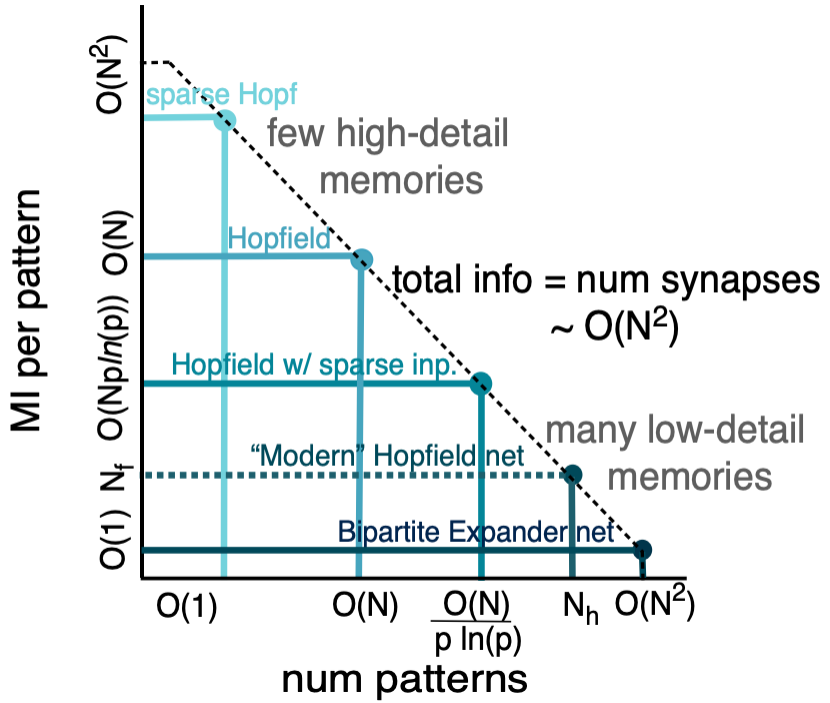


Hopfield network

Existing CAM models lack a memory continuum

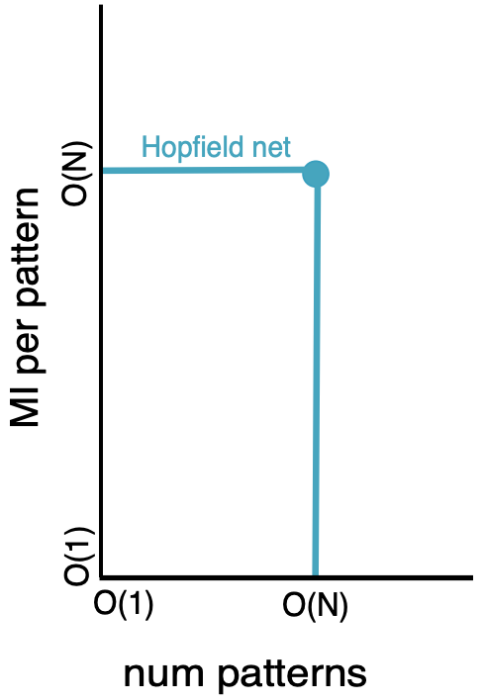


Hopfield network

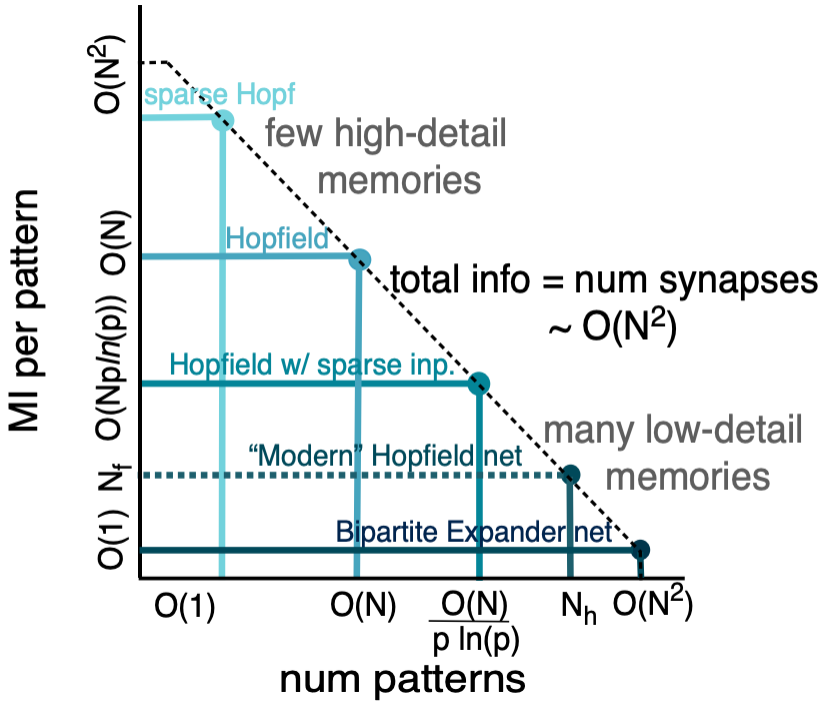


Existing models

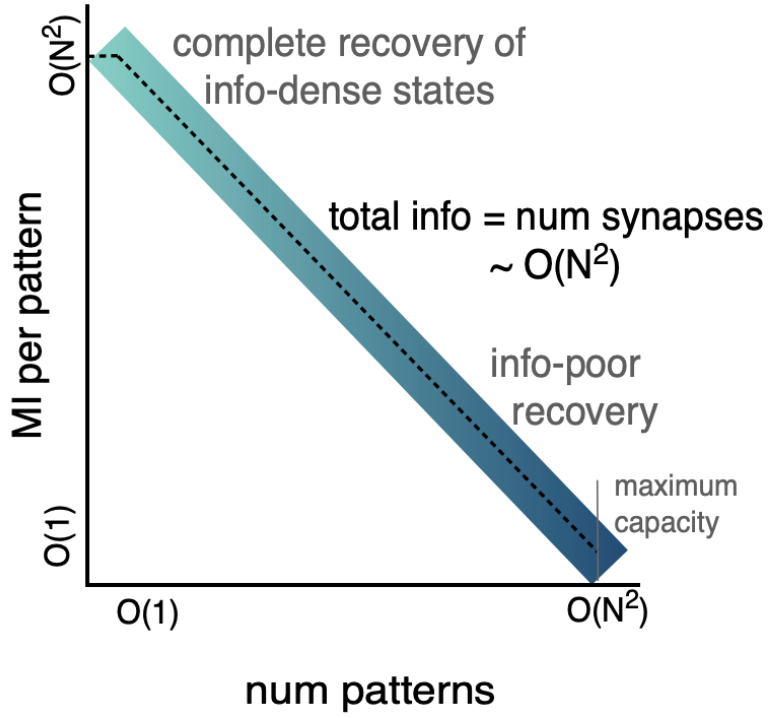
Existing CAM models lack a memory continuum



Hopfield network

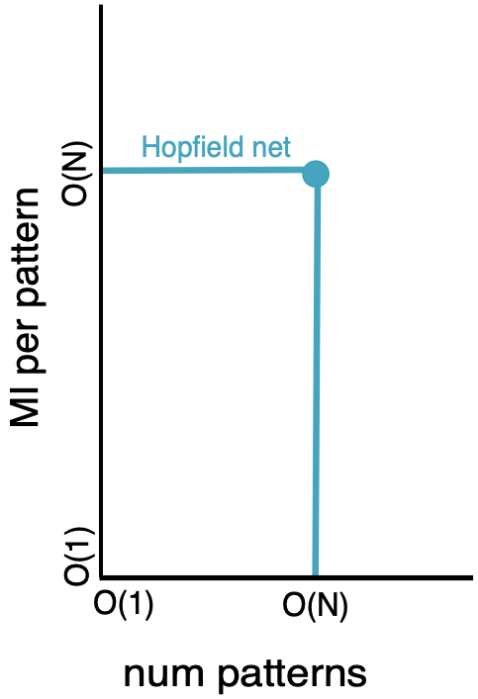


Existing models

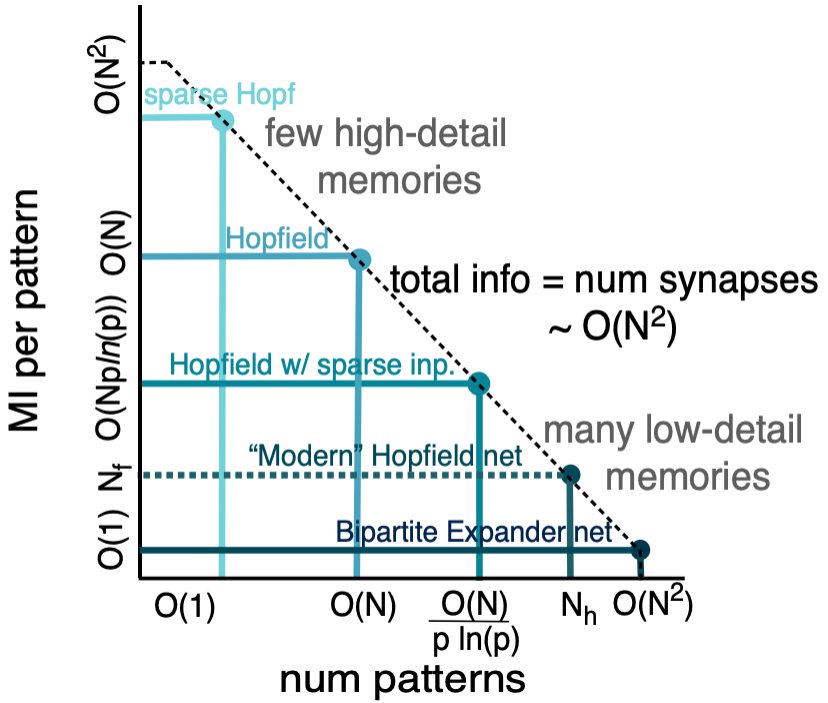


Desired CAM continuum

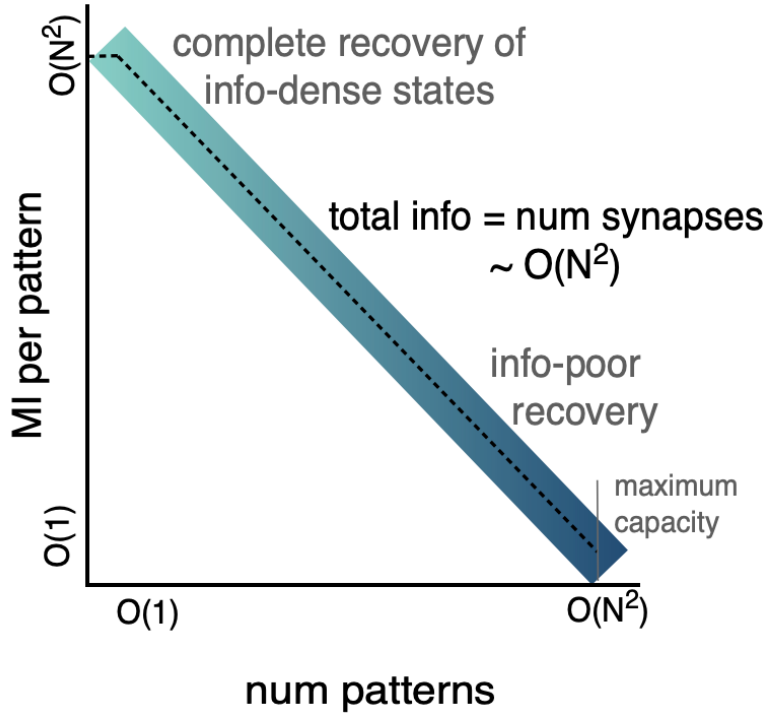
Existing CAM models lack a memory continuum



Hopfield network



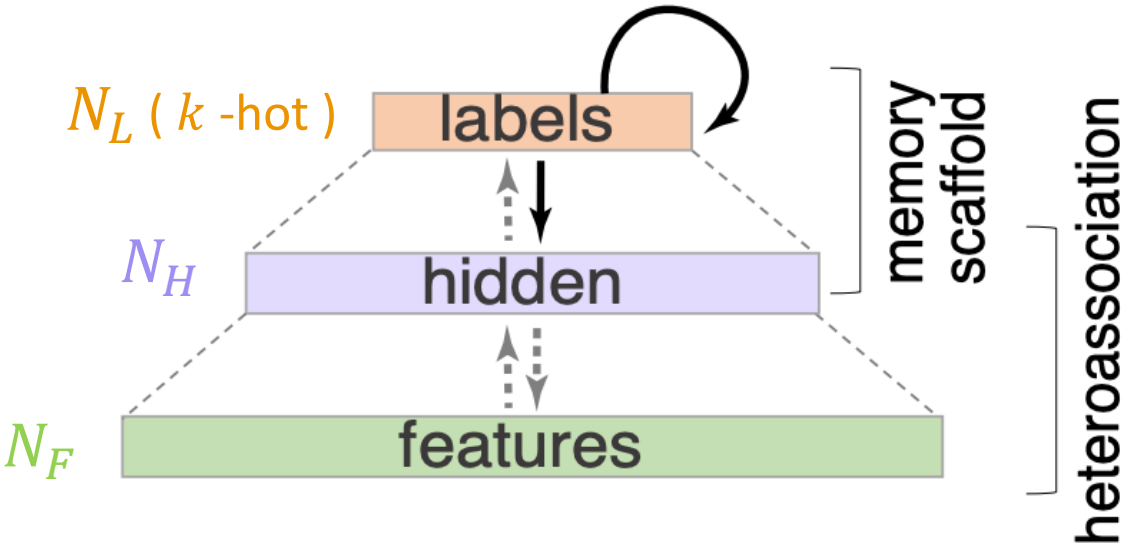
Existing models

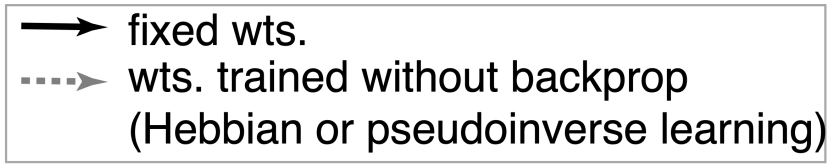


Desired CAM continuum achieved through our model!

Memory **S**caffold with **H**eteroassociation (MESH)

Separation of input features and persistence of memory

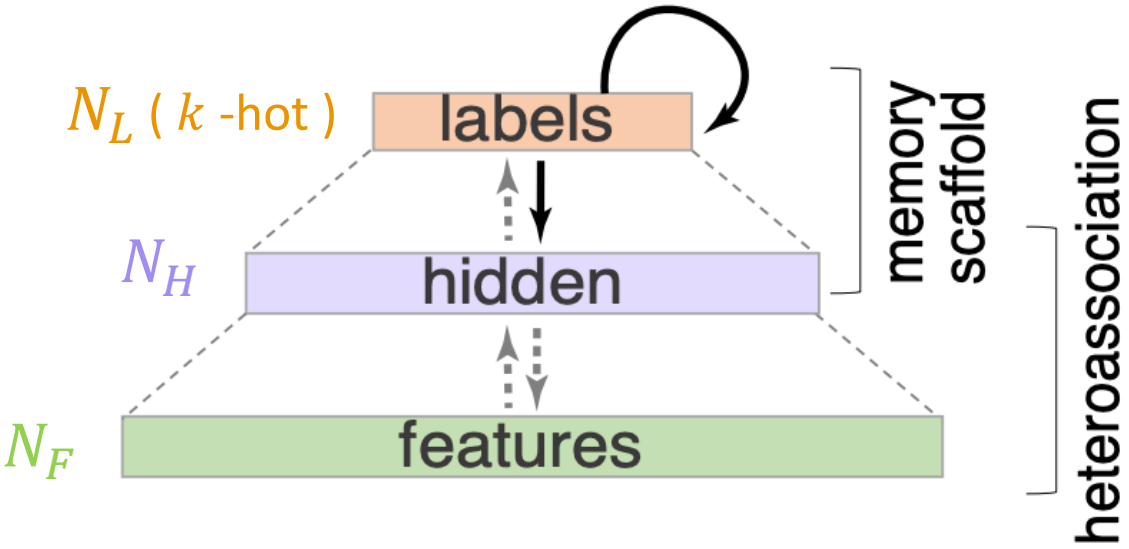


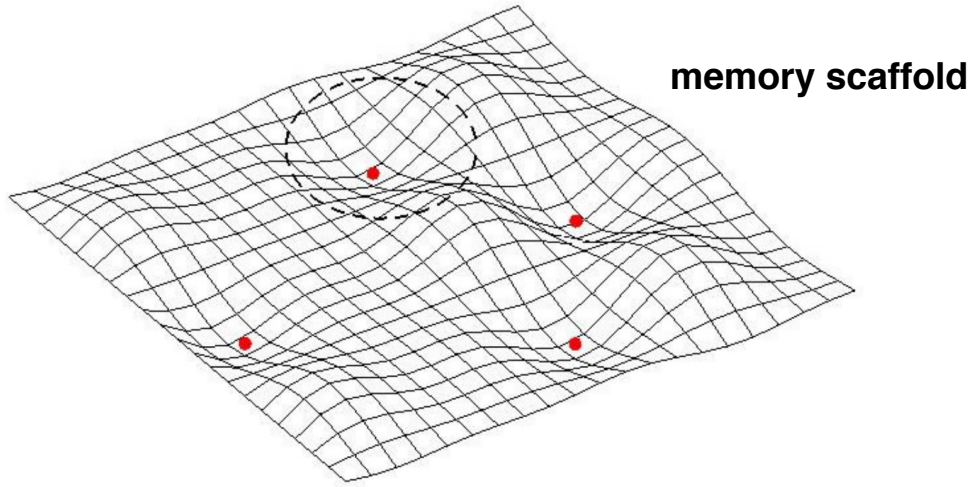


$N_H \sim O(1)$ $N_F \sim O(N_L^k)$

Memory **S**caffold with **H**eteroassociation (MESH)

Separation of input features and persistence of memory



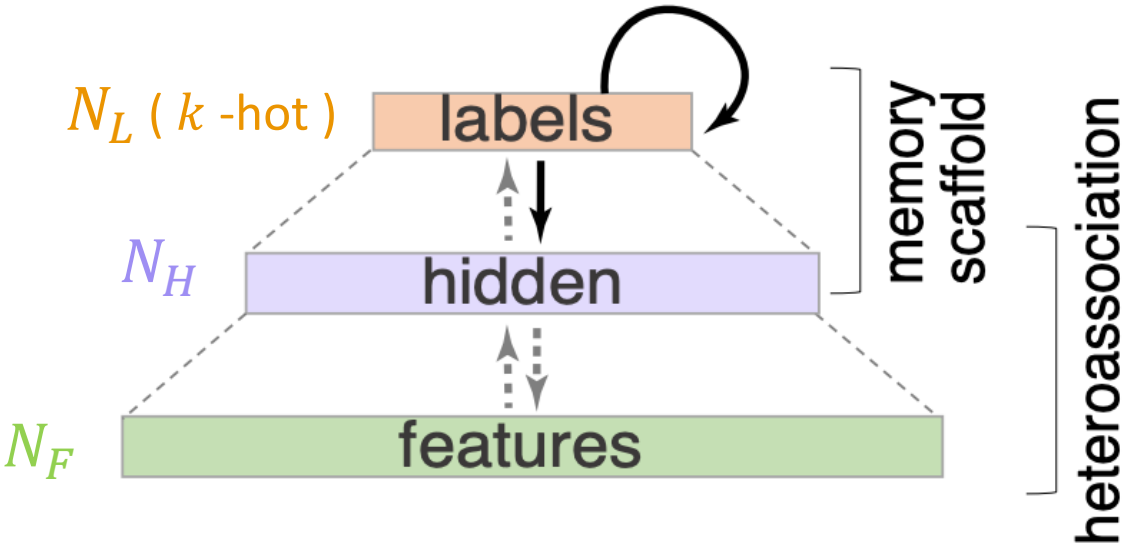


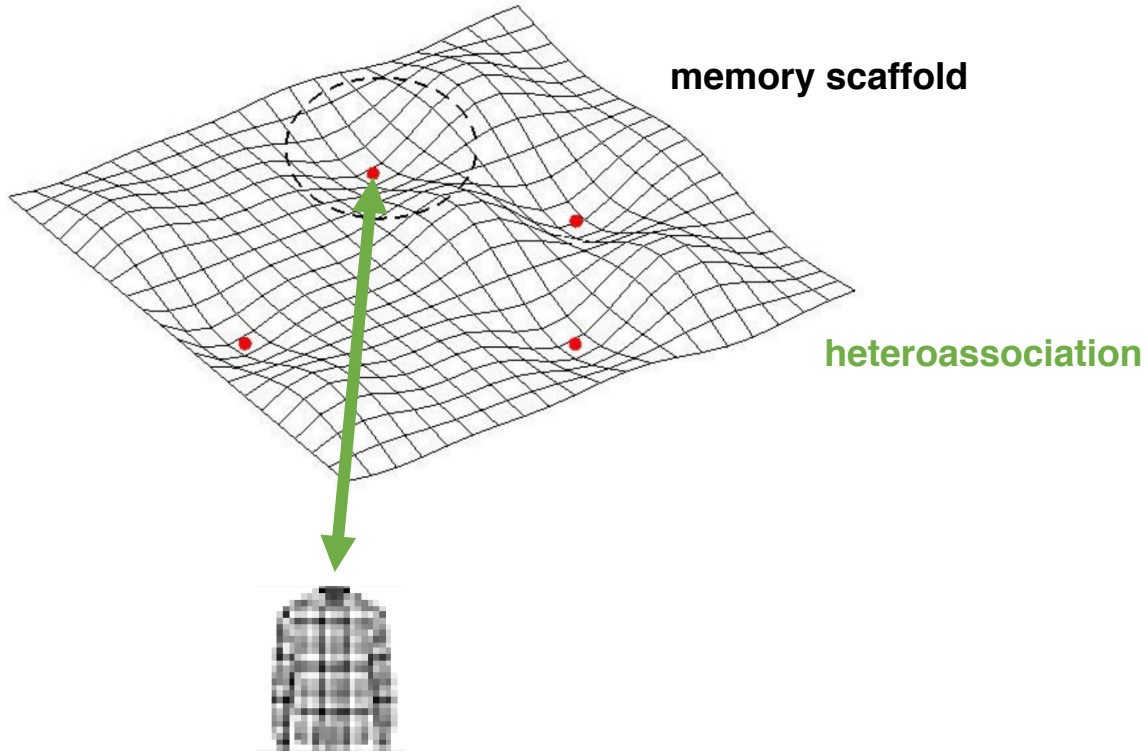
$N_H \sim O(1)$ $N_F \sim O(N_L^k)$

Total number of states: $\binom{N_L}{k} \sim \exp(dN_L)$

Memory **S**caffold with **H**eteroassociation (MESH)

Separation of input features and persistence of memory



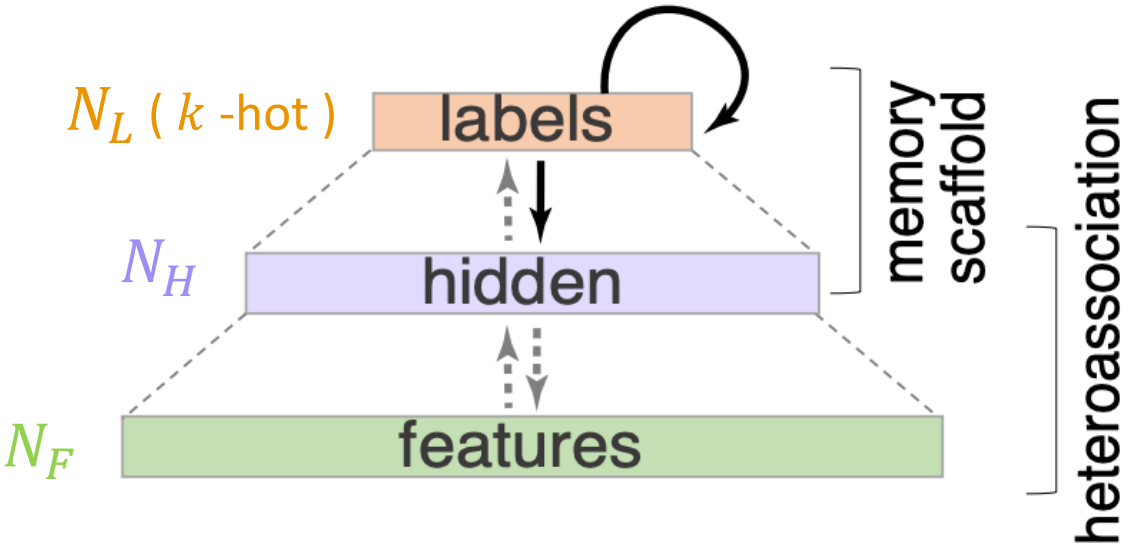


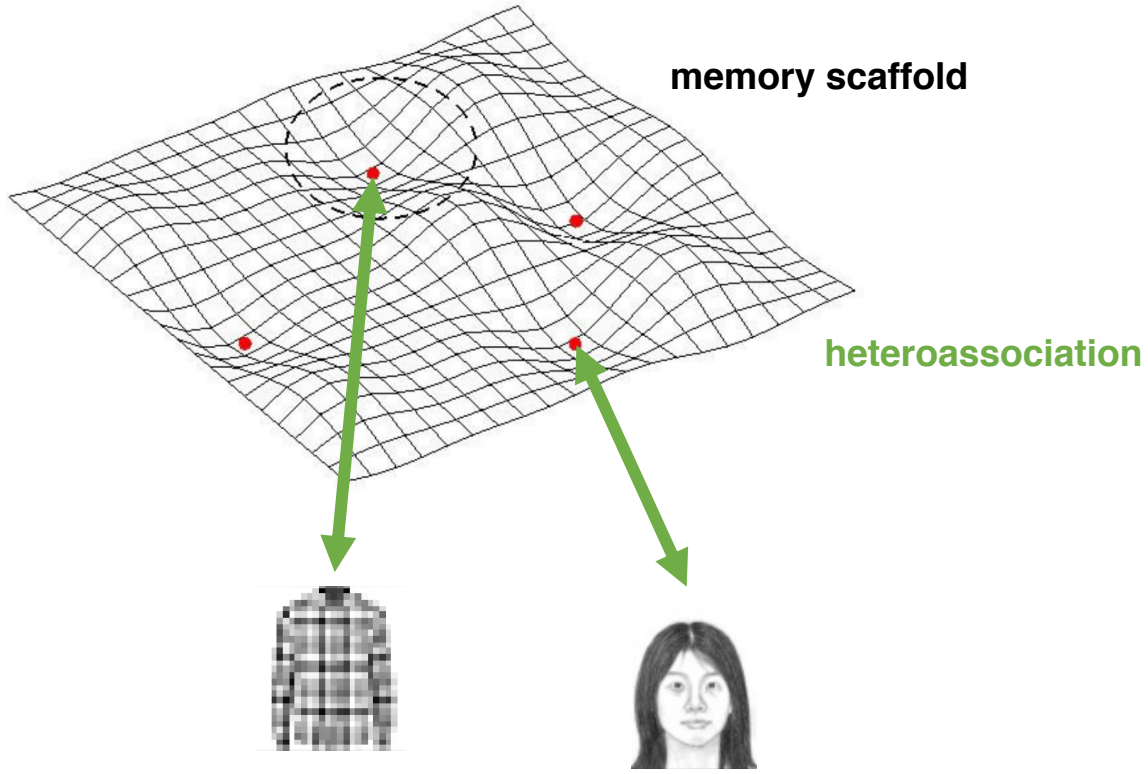
$N_H \sim O(1)$ $N_F \sim O(N_L^k)$

Total number of states: $\binom{N_L}{k} \sim \exp(dN_L)$

Memory **S**caffold with **H**eteroassociation (MESH)

Separation of input features and persistence of memory



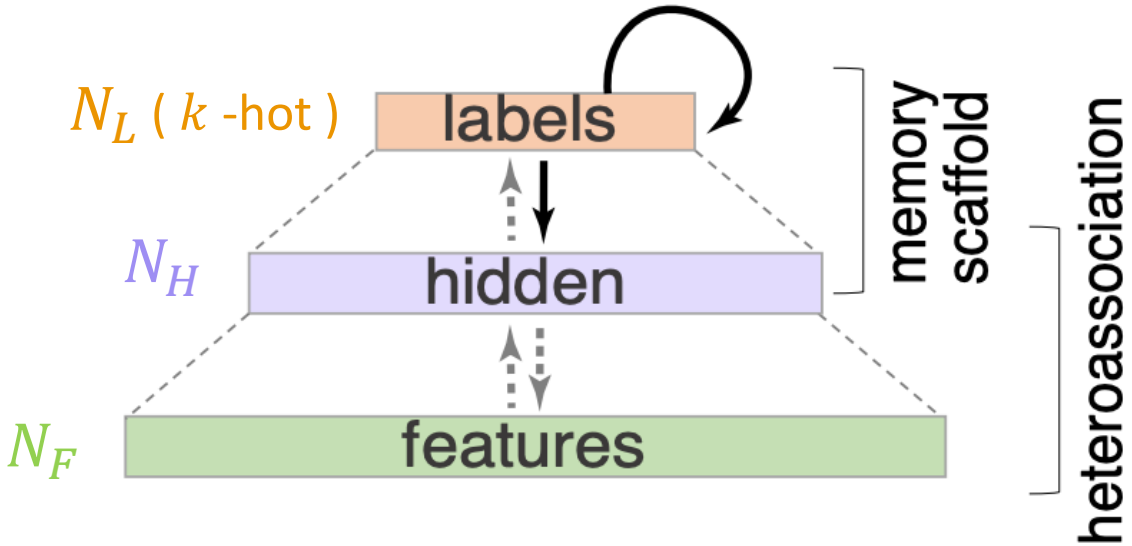


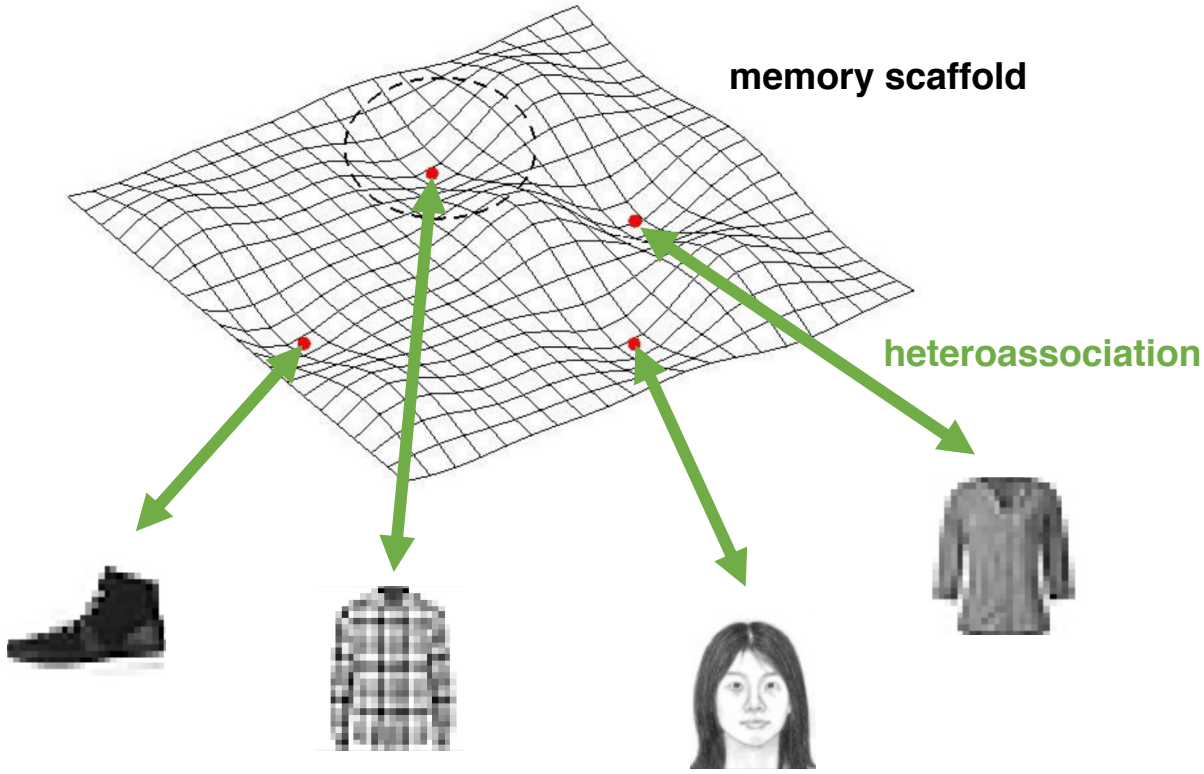
$N_H \sim O(1)$ $N_F \sim O(N_L^k)$

Total number of states: $\binom{N_L}{k} \sim \exp(dN_L)$

Memory Scaffold with Heteroassociation (MESH)

Separation of input features and persistence of memory

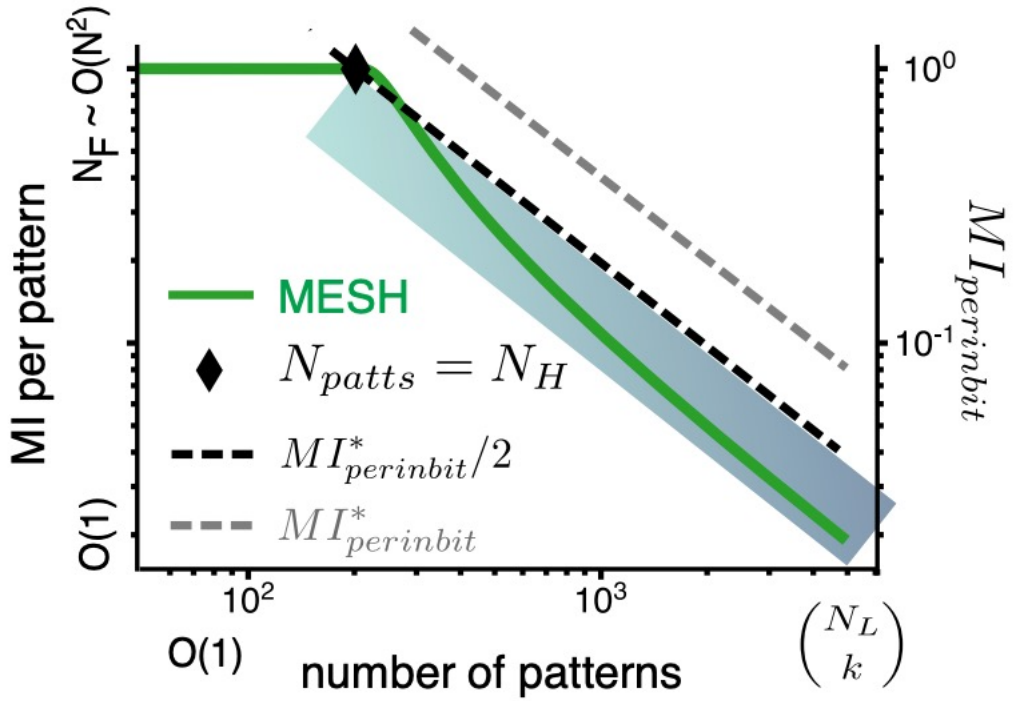


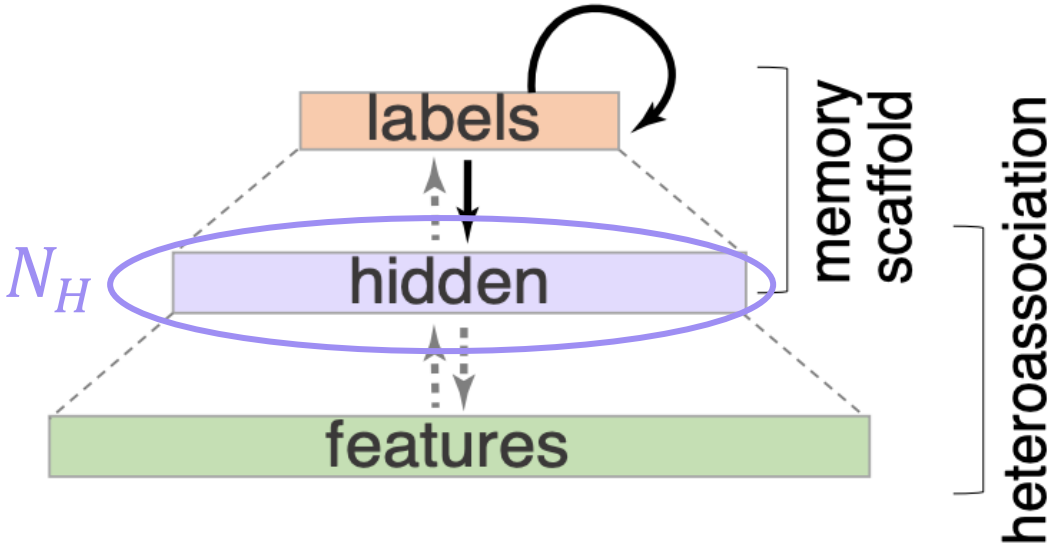


$N_H \sim O(1)$ $N_F \sim O(N_L^k)$

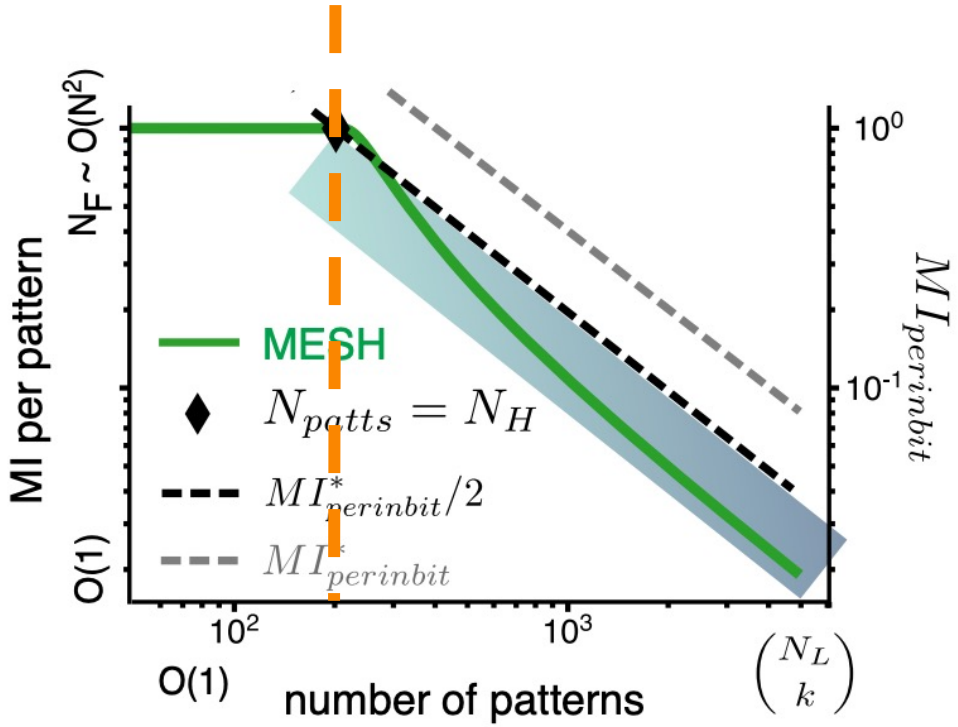
Total number of states: $\binom{N_L}{k} \sim \exp(dN_L)$

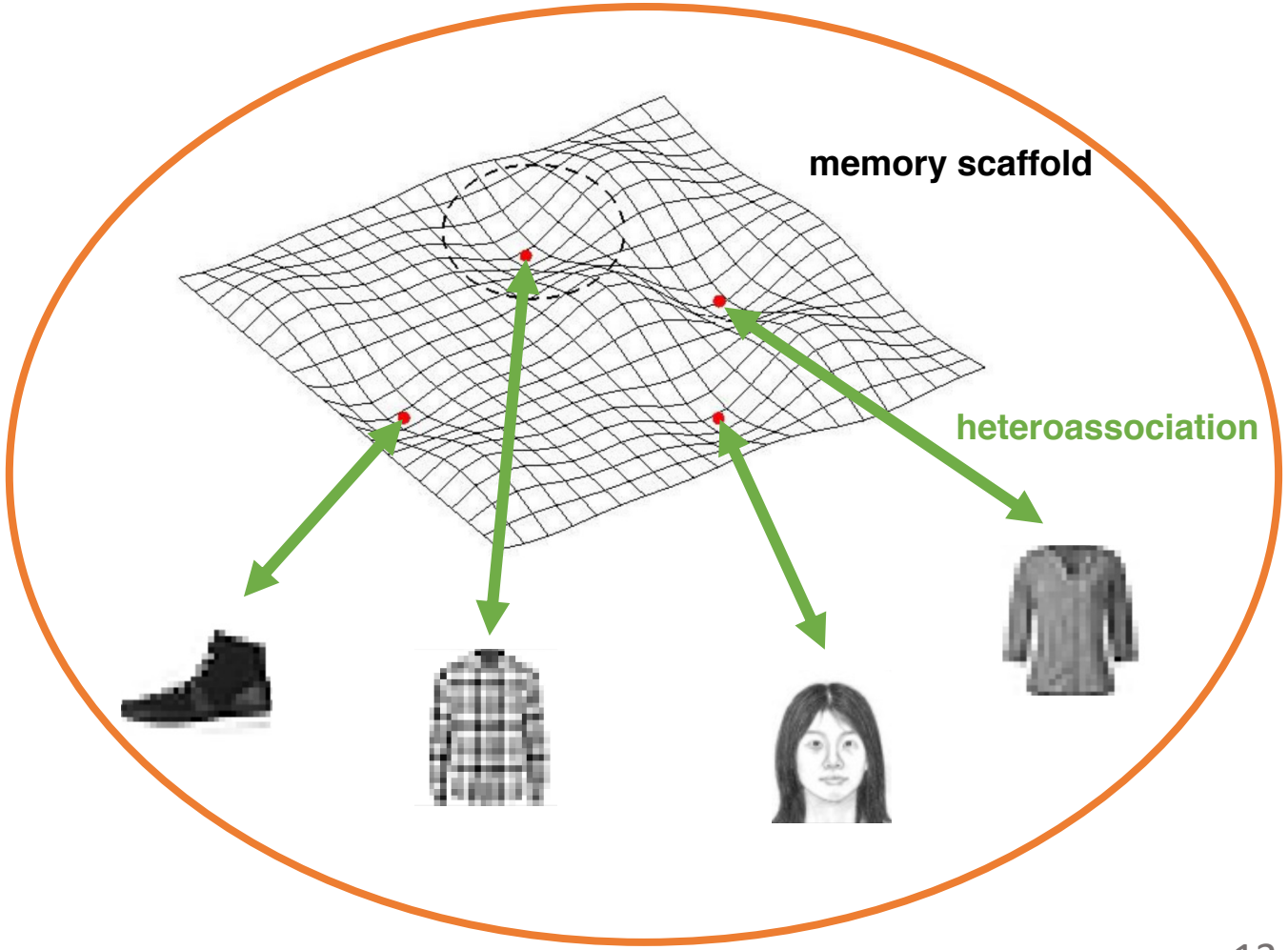
MESH exhibits a near-optimal CAM continuum



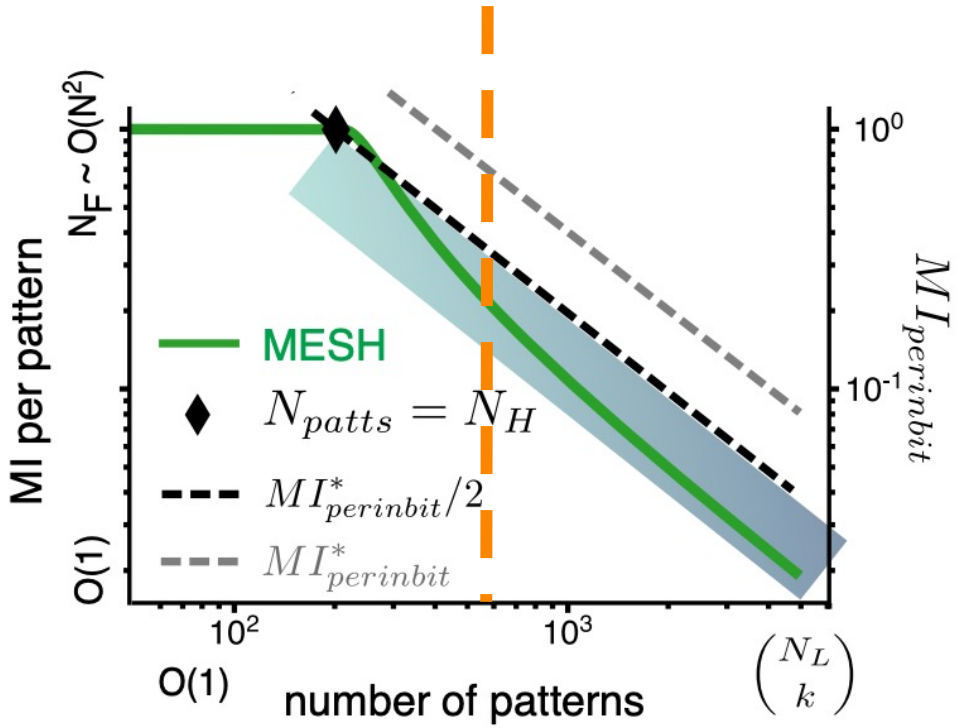


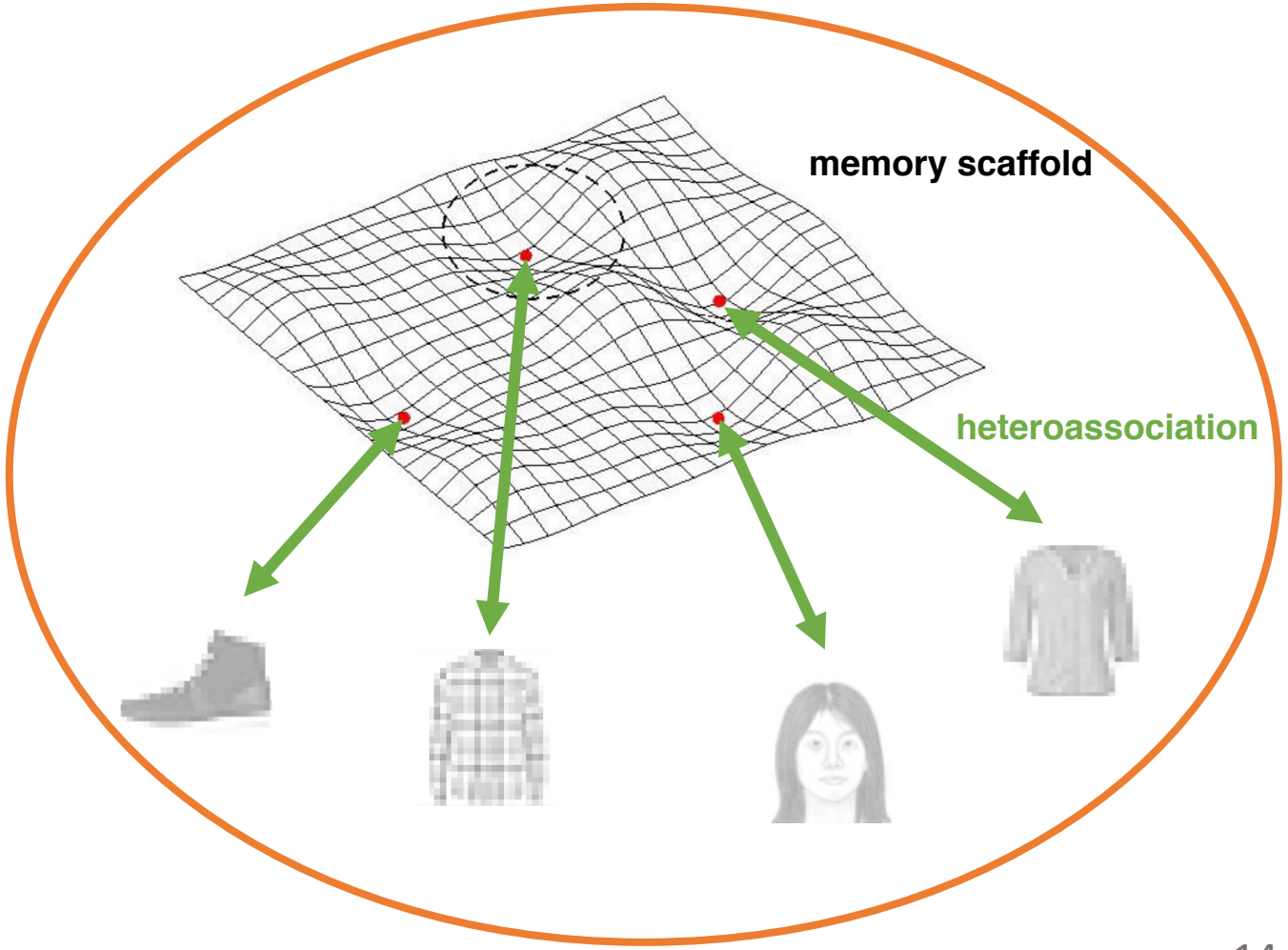
Complete recovery of up to N_H memories



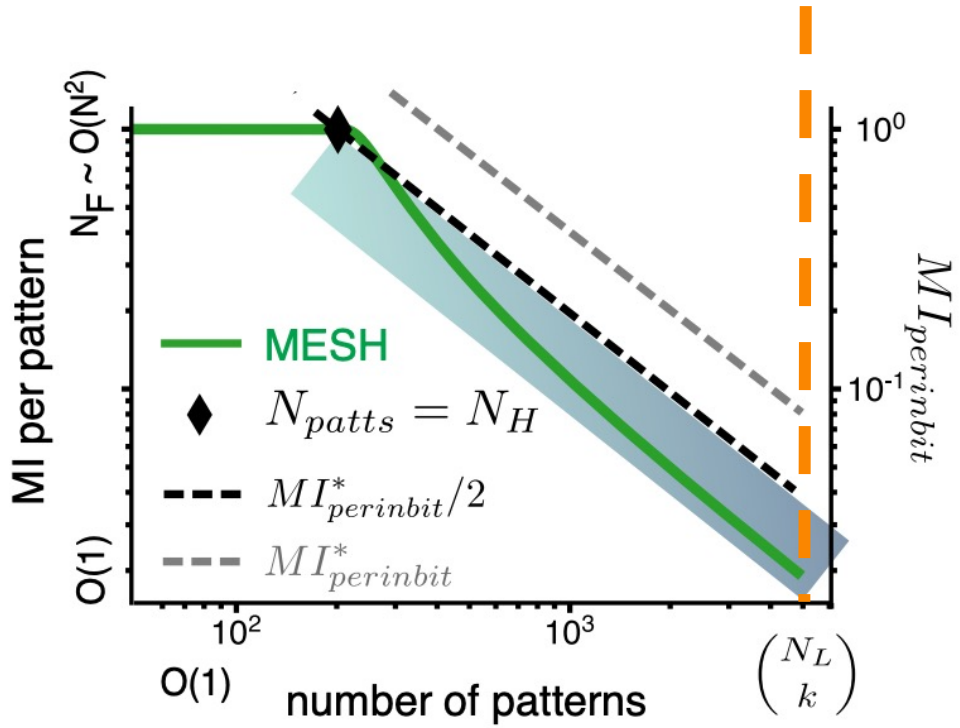


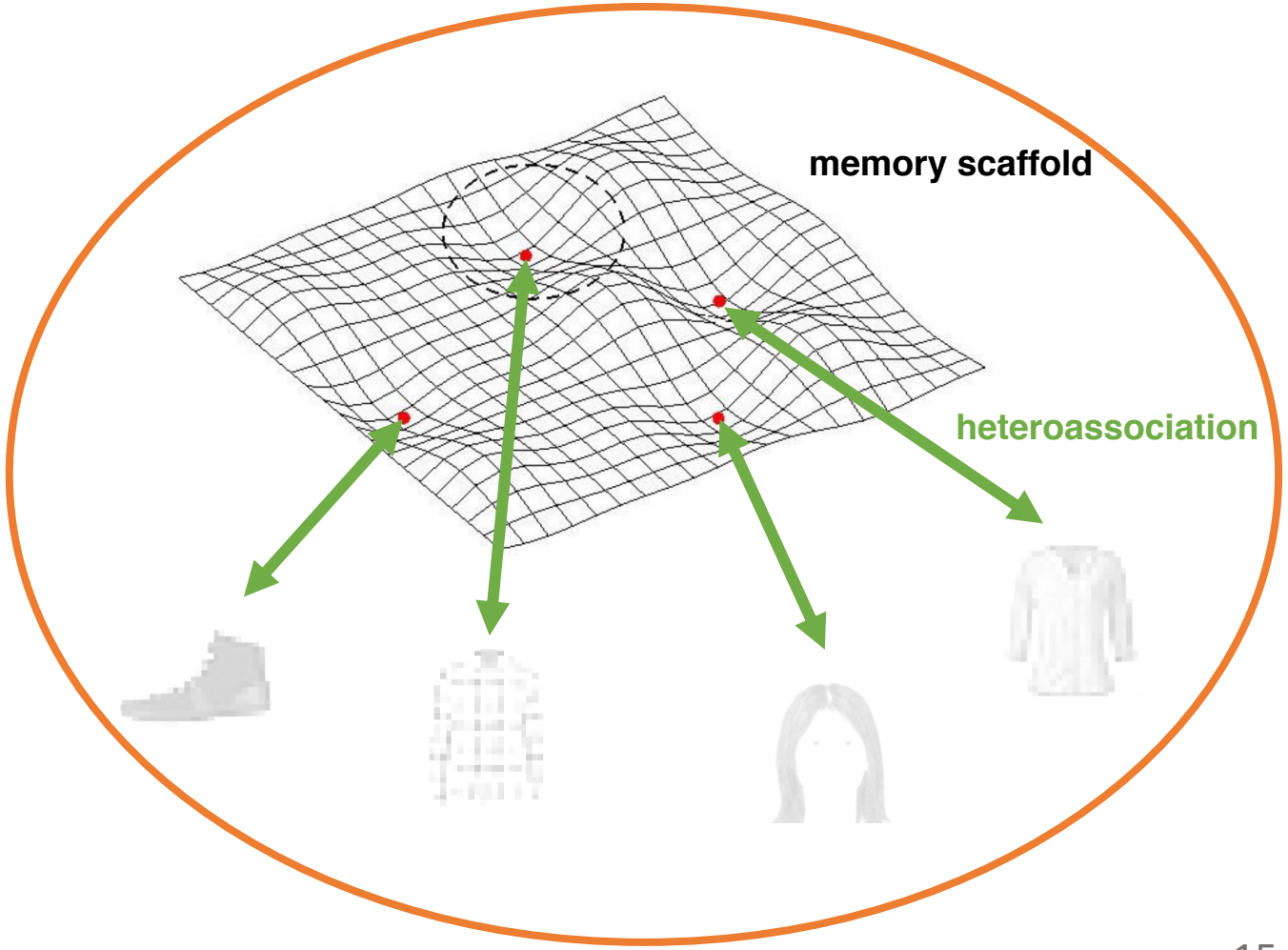
Partial recovery of each of the memories > N_H





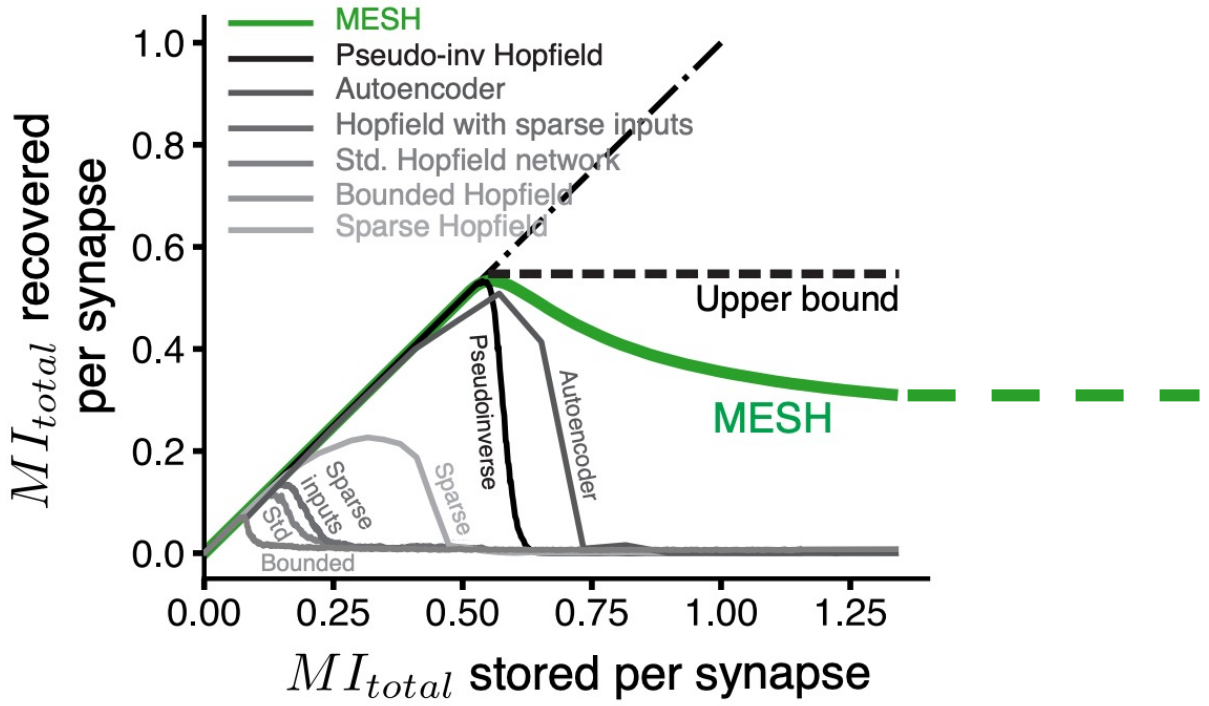
Gradual degradation of information per memory





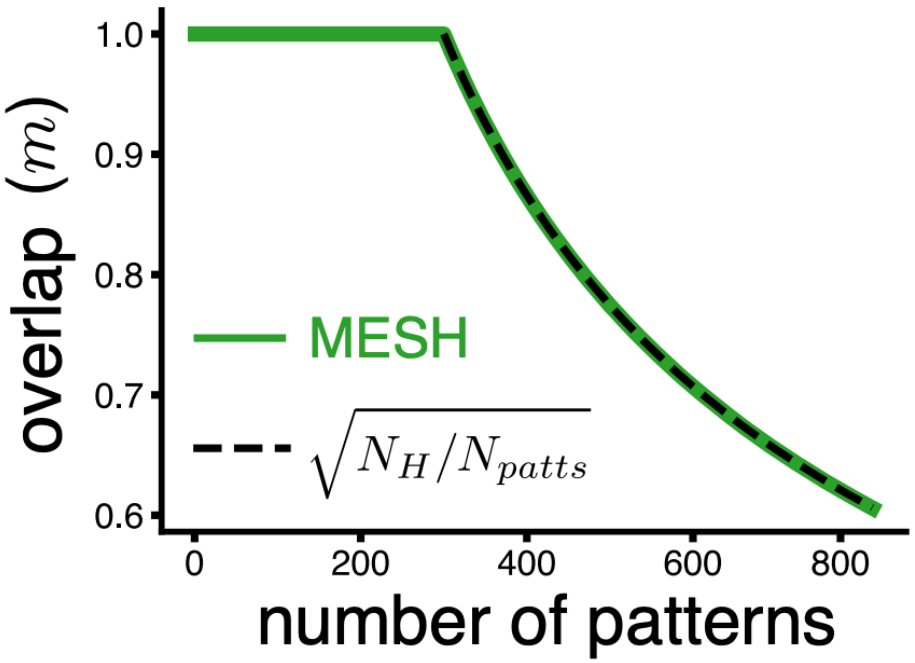
Constant total information per synapse in MESH

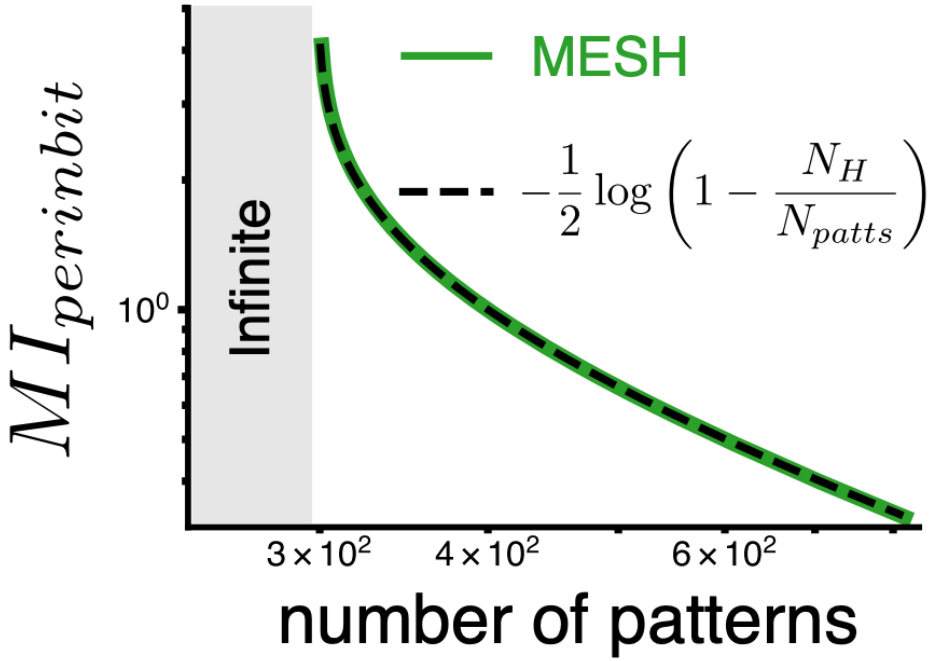
- Given a network of fixed size, the total information is invariant to the number of stored patterns
- Smooth trade-off between number of patterns and pattern richness.



Continuous valued patterns successfully stored

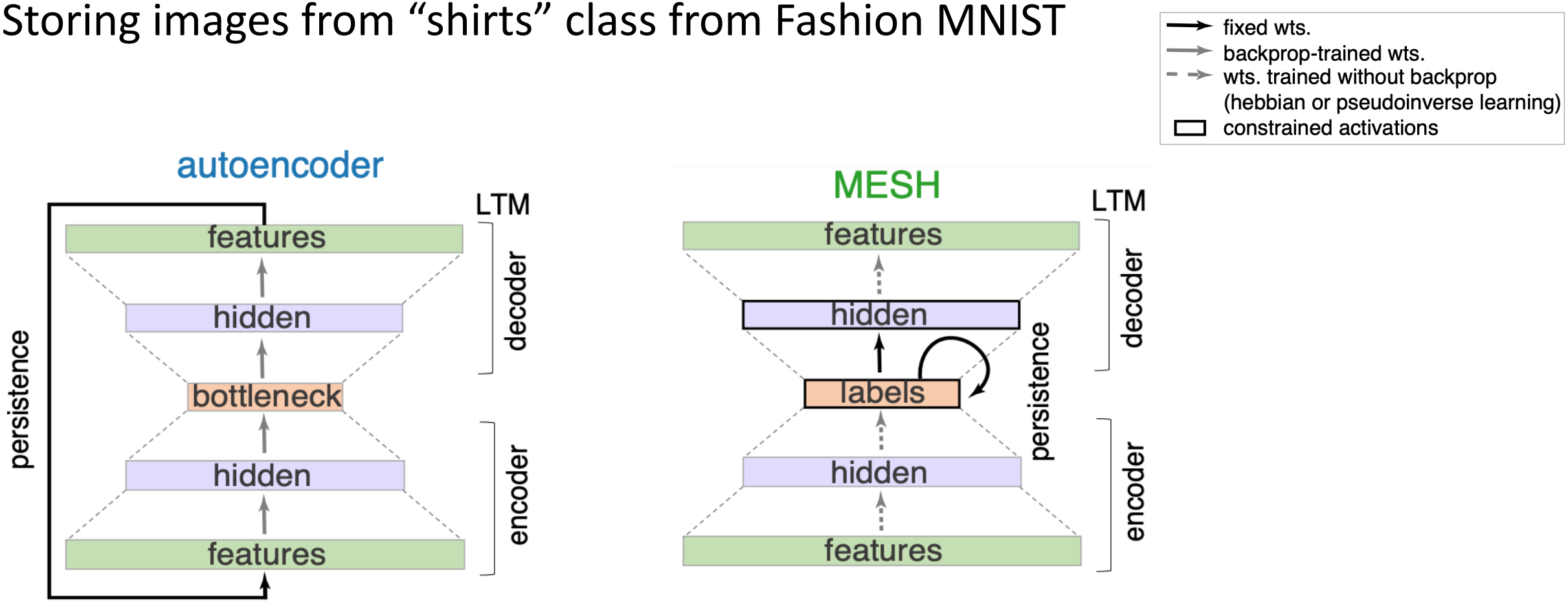
Random continuous valued patterns





Continuous valued patterns successfully stored

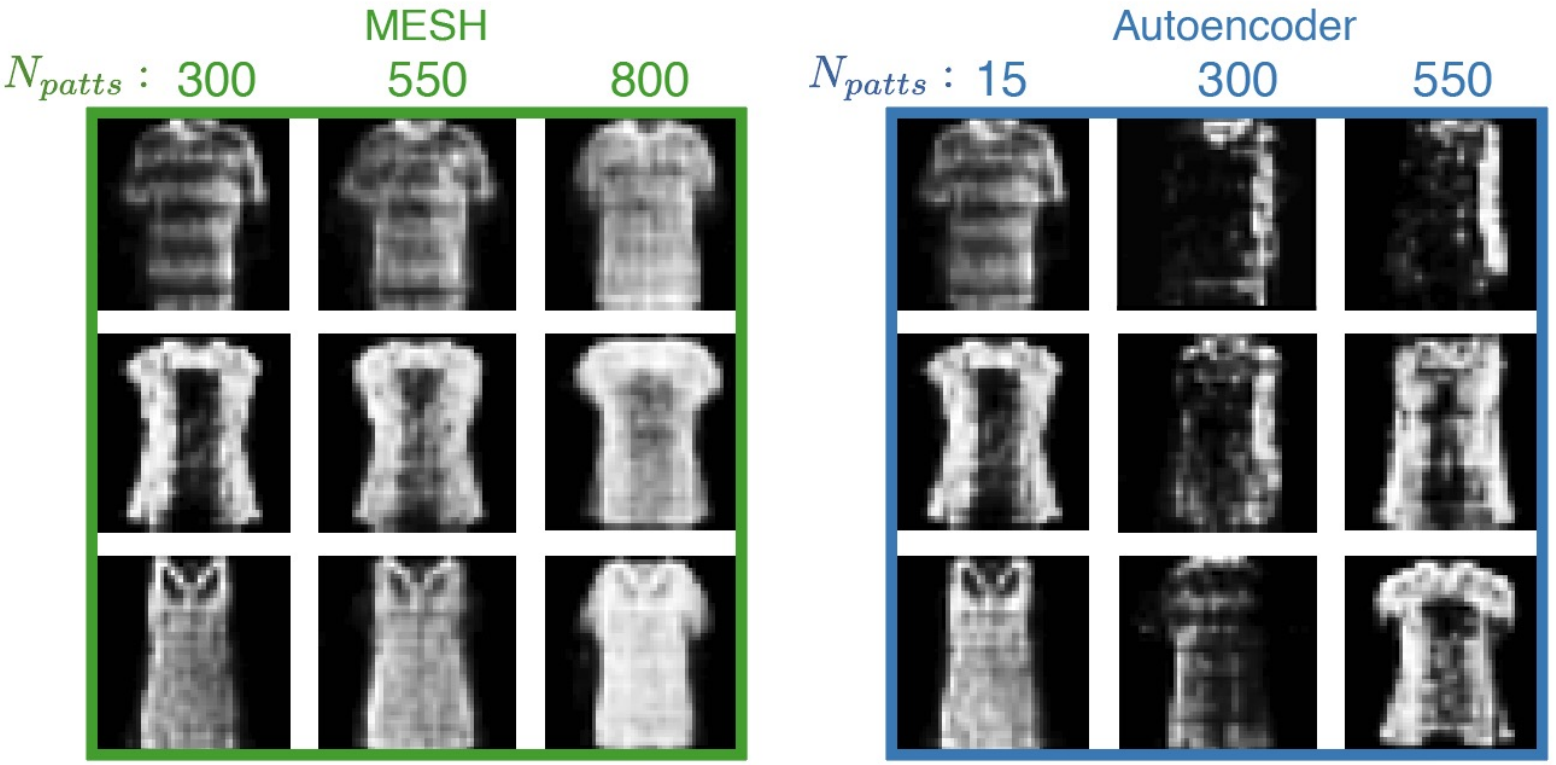
Storing images from “shirts” class from Fashion MNIST



Radhakrishnan, Belkin, Uhler (2020)

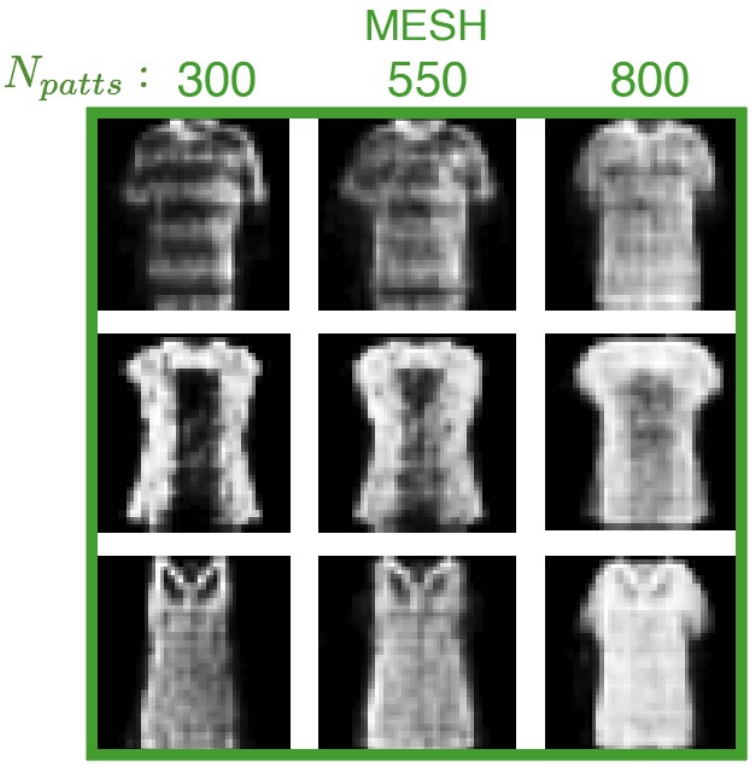
Continuous valued patterns successfully stored

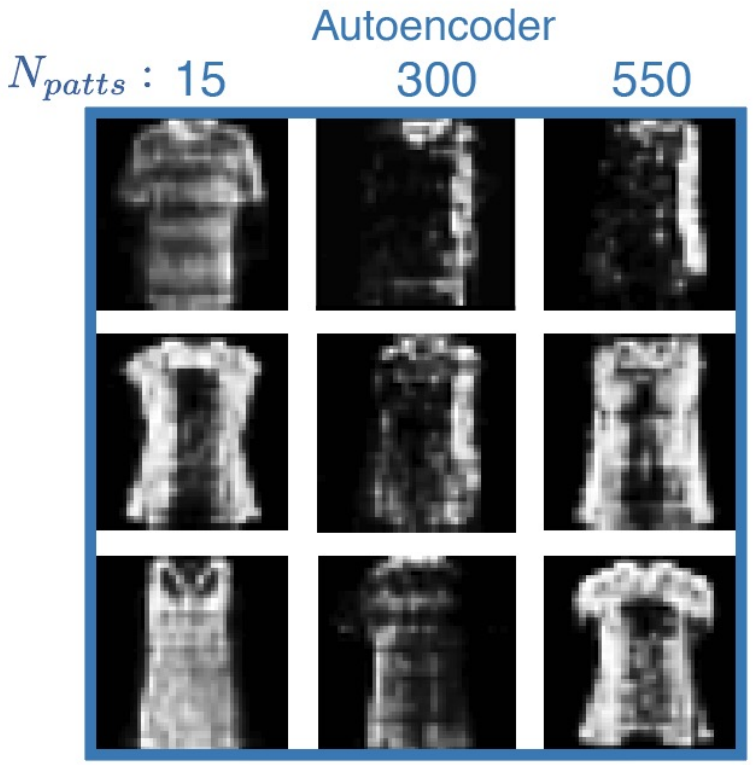
Storing images from “shirts” class from Fashion MNIST

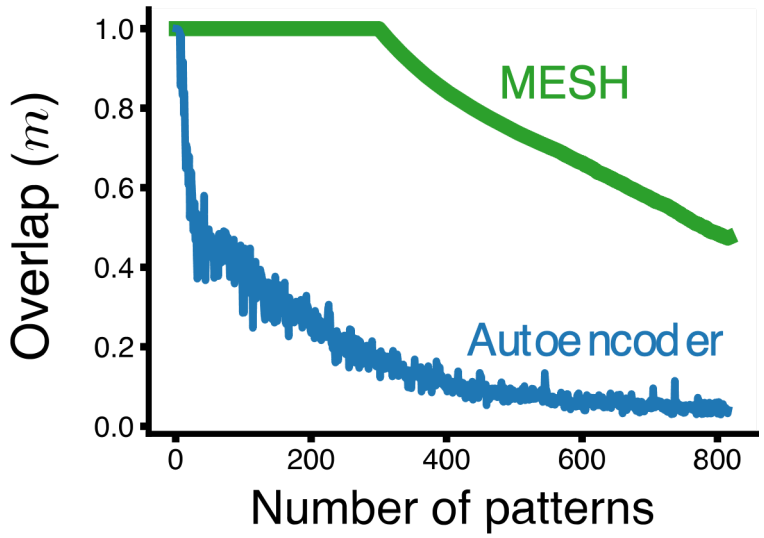


Continuous valued patterns successfully stored

Storing images from “shirts” class from Fashion MNIST







Conclusions

- MESH is a biologically plausible **associative memory model without a catastrophic memory cliff** (maps onto the entorhinal-hippocampal system in the brain)
- **Memory scaffold** generates an exponentially large number of fixed points; **Heteroassociation** labels arbitrary patterns to these fixed points
- MESH can be used for the storage, lookup, and retrieval of memory states through a **factorized key-value structure**; recognition/familiarity detection, locality sensitive hashing.