Spotlight

DenseAM with Epanechnikov Energy

Unlocking simultaneous **creativity** and **memorization** in *Dense Associative Memories* (*DenseAMs*)

Benjamin **Hoover**

Zhaoyang **Shi**

Krishna **Balasubramanian**

Dmitry **Krotov**

Parikshit **Ram**

More info!

We want models that can

1. **Memorize** training data (near zero training loss)
2. **Generalize** to new data

Models that memorize training data still generalize well

Dense Associative Memories (DenseAMs) are great at memorizing…

*but **novel generation** sacrifices retrievability of **stored patterns***

We want DenseAMs that can **perfectly memorize** and **generate novelty** *at the same time*

DenseAM is unnormalized KDE

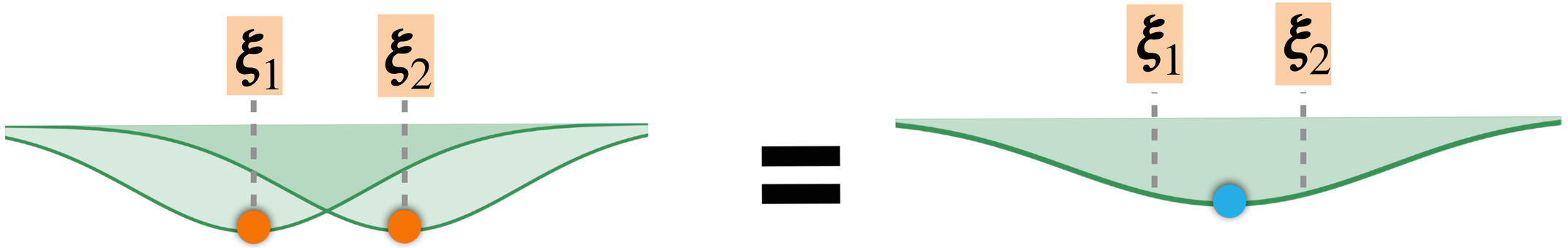
LogSumExp (LSE) energy sacrifices **memorization** for **generalization**

Gaussian Kernel

$$E^{\mathrm{LSE}} = -\frac{1}{\beta}\log\sum_{\mu=1}^{M}\exp\left(-\frac{\beta}{2}\|\mathbf{x}-\boldsymbol{\xi}_\mu\|^2\right)$$

M **stored patterns** $\boldsymbol{\xi}_\mu$,

$\mu = \{1 \ldots M\}$



Novel generation *sacrifices retrievability of* ***stored patterns***

DenseAM is unnormalized KDE

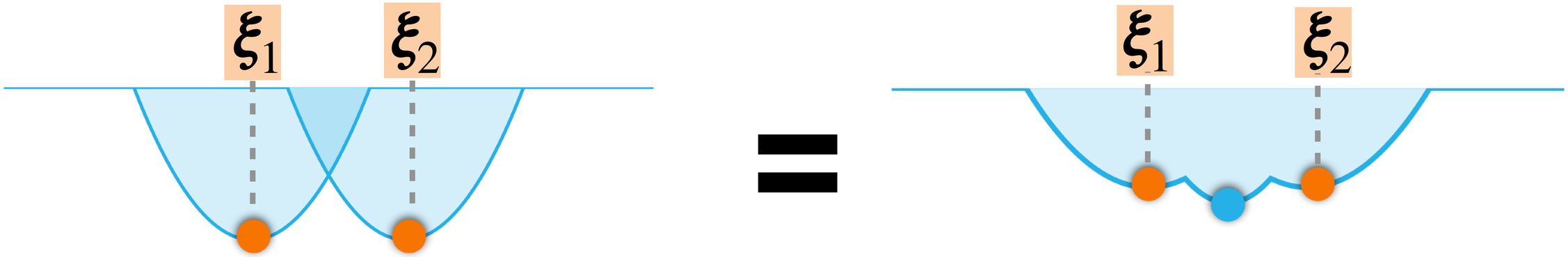
LogSumReLU (LSR) energy both **memorizes** *and* **generalizes**

Epanechnikov Kernel

$$E^{\mathrm{LSR}} = -\frac{1}{\beta}\log\sum_{\mu}\mathrm{ReLU}\left(1-\frac{\beta}{2}\|\mathbf{x}-\boldsymbol{\xi}_{\mu}\|^{2}\right)$$

M **stored patterns** $\boldsymbol{\xi}_{\mu}$,

$\mu = \{1 \ldots M\}$



Novel generation *coexists with* ***exact memorization***

DenseAM is unnormalized KDE

LogSumReLU (LSR) energy both **memorizes** *and* **generalizes**

Epanechnikov Kernel

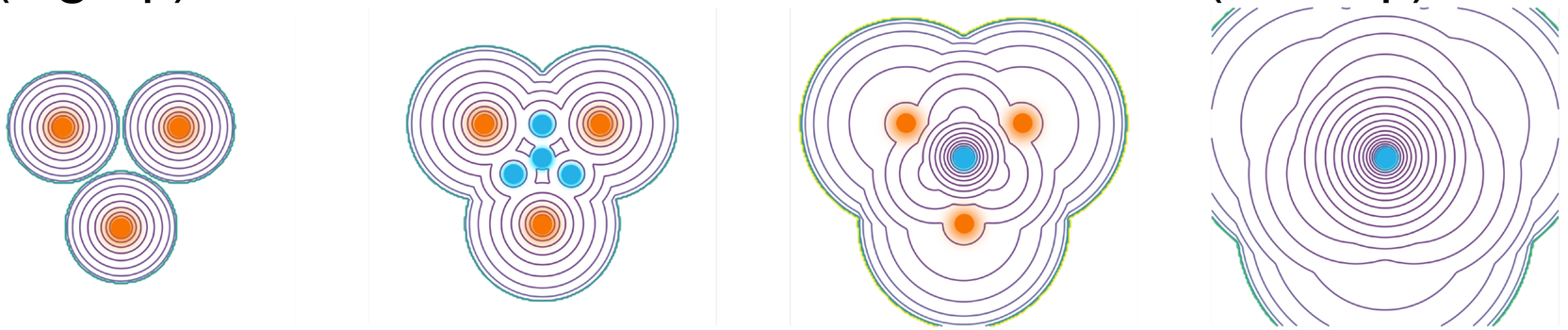
$$E^{\mathrm{LSR}} = -\frac{1}{\beta}\log\sum_{\mu}\mathrm{ReLU}\left(1-\frac{\beta}{2}\|\mathbf{x}-\boldsymbol{\xi}_{\mu}\|^2\right)$$

M **stored patterns** $\boldsymbol{\xi}_{\mu}$,

$\mu = \{1 \ldots M\}$

Narrow basins
(high β)

Wide basins
(small β)

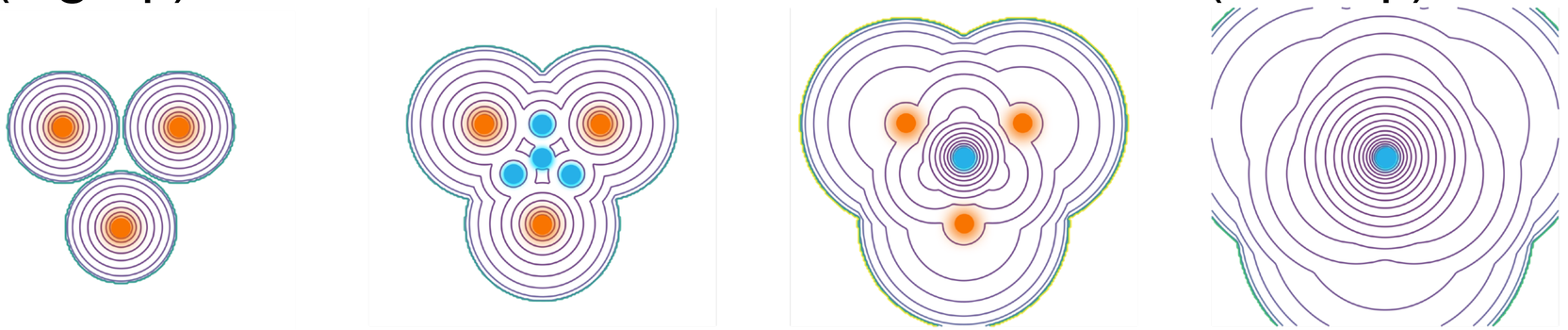


Emergence *happens* ***only*** *at critical β values*

Emergent memories form when stored patterns interact to form a new memory, and those interacting patterns remain memories themselves

Narrow basins
(high β)

Wide basins
(small β)



Emergence *happens* ***only*** *at critical β values*

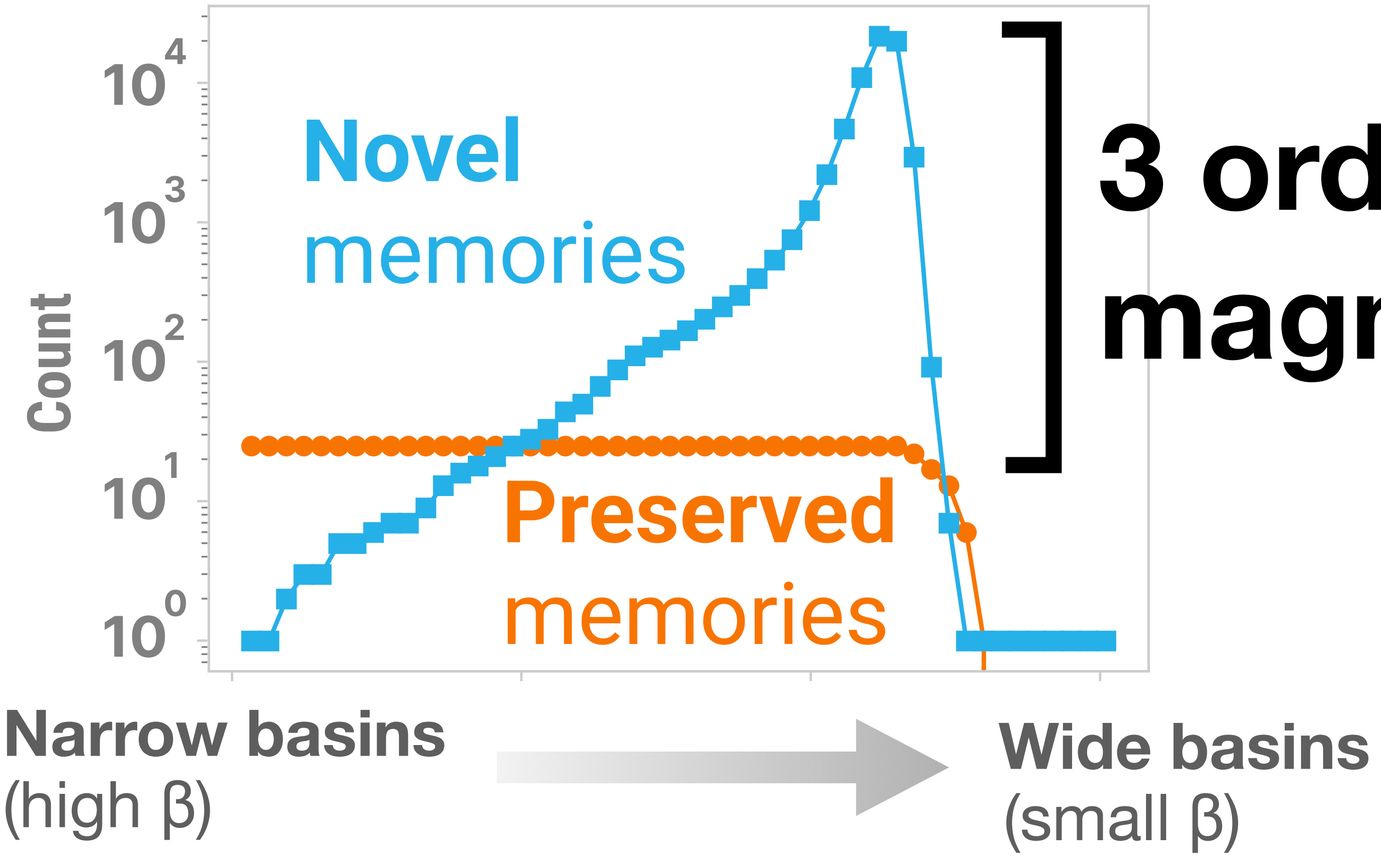
Three research questions

1. How **abundant** are emergent memories?
2. How **meaningful** are emergent memories?
3. Can emergent memories **scale** to real data?

How many emergent memories can LSRDAM create?

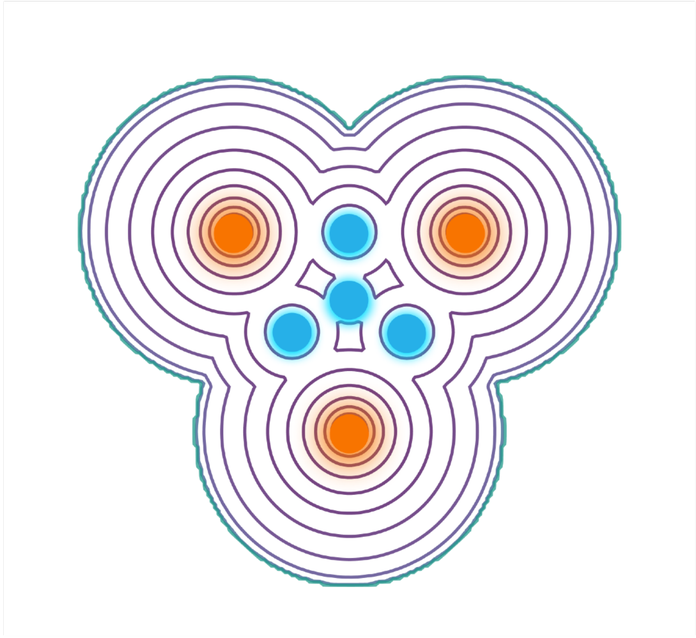
Create *orders-of-magnitude* more emergent memories than stored patterns

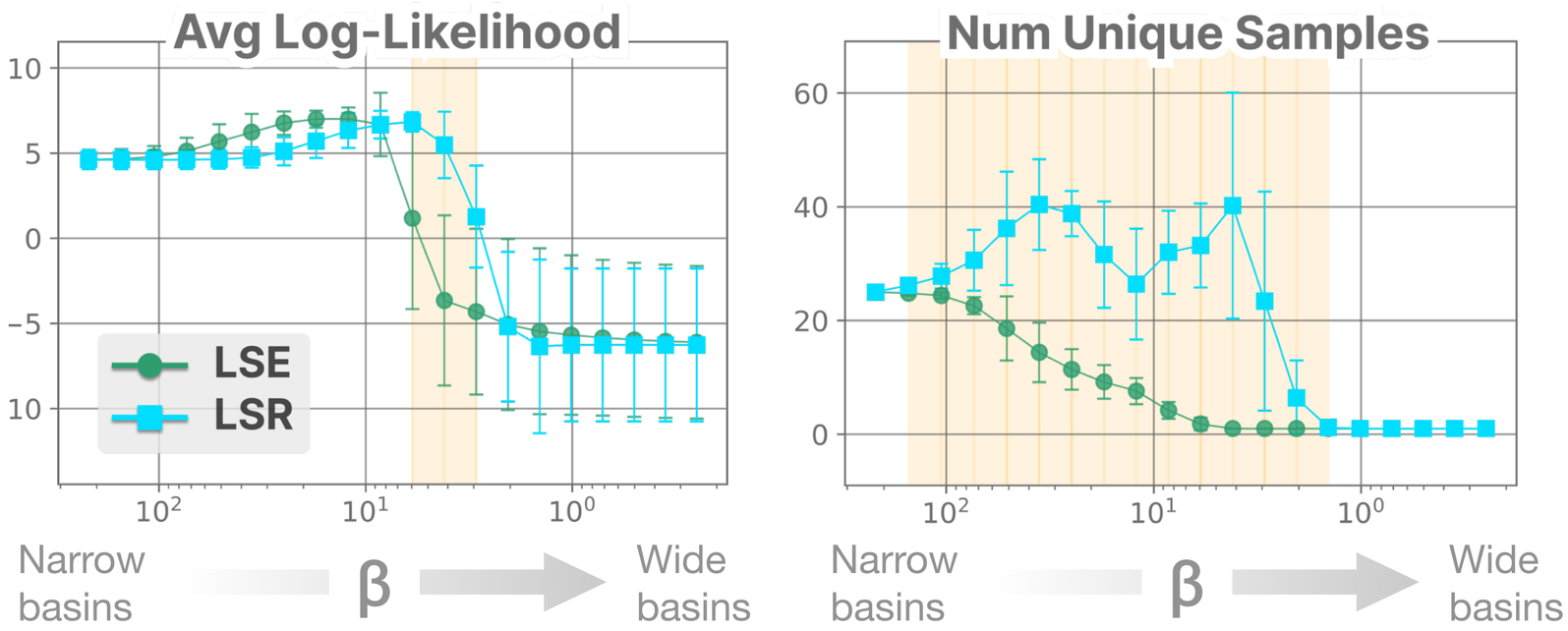
Storing **25** random patterns on the **32**-dim unit hypercube



How meaningful can they be?

Emergent memories must be *centroids of subsets of patterns*





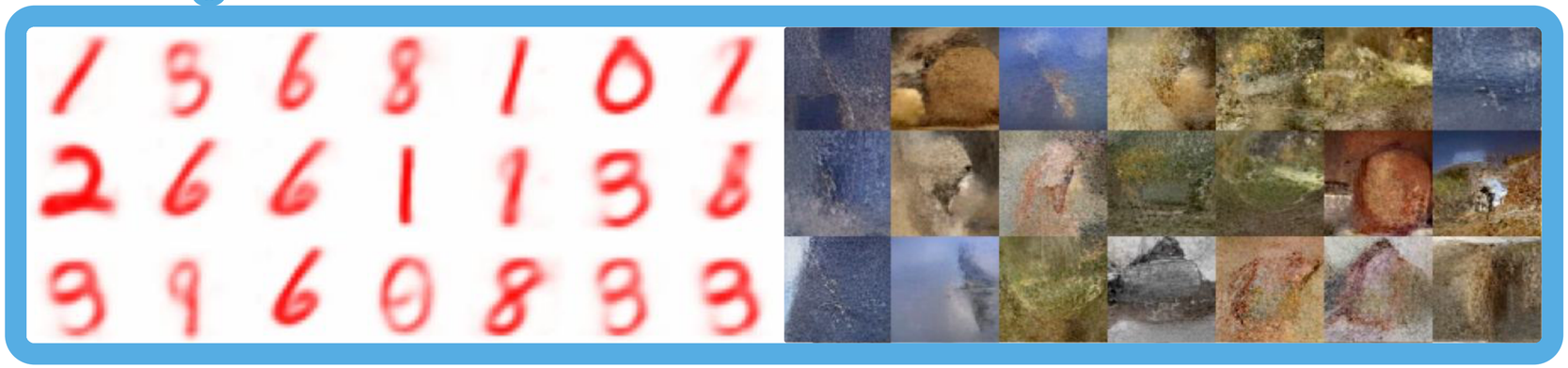
Stored patterns are 100 I.I.D. samples from a mixture of five 8-dim Gaussians, avged across 5 seeds

Can emergent memories scale?

Emergent memories capture novel regions of **latent space**

Stored patterns

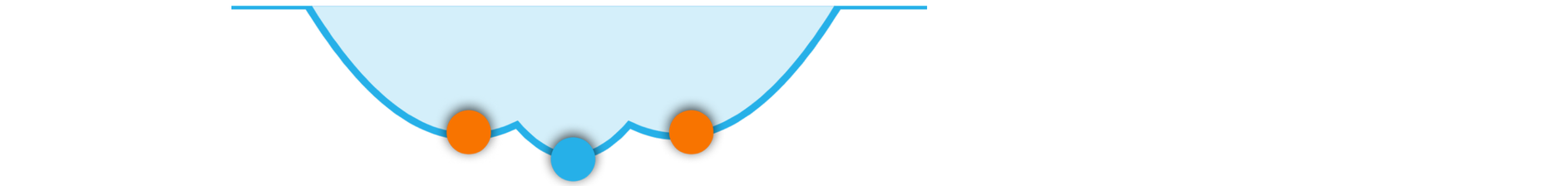
Emergent memories

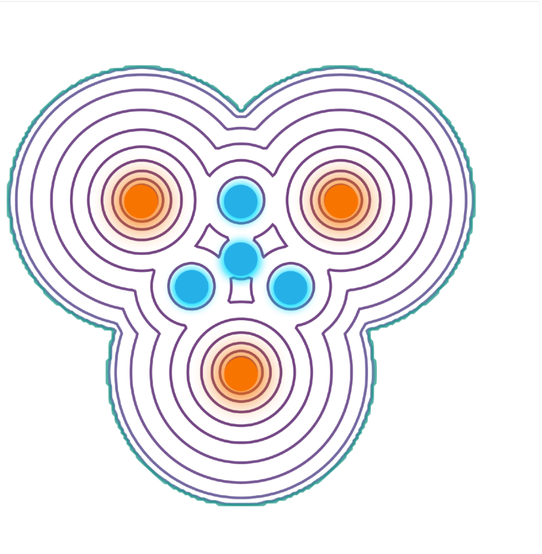


Thanks!

DenseAM with Epanechnikov Energy

Unlocking simultaneous **creativity** and **memorization** in *DenseAMs via* **emergent memories**





Georgia Tech

HARVARD UNIVERSITY

UCDAVIS

IBM Research

Get in touch!