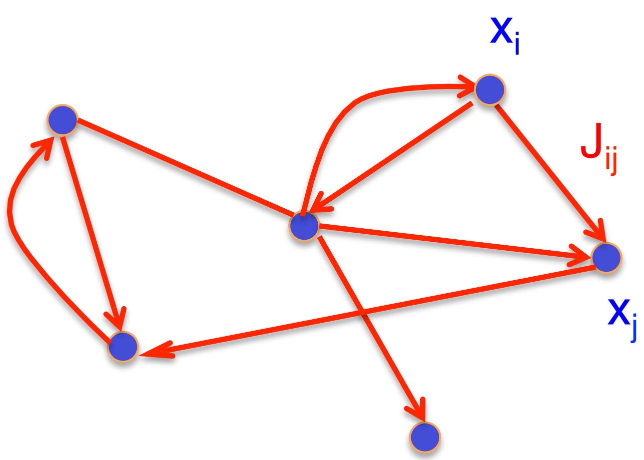
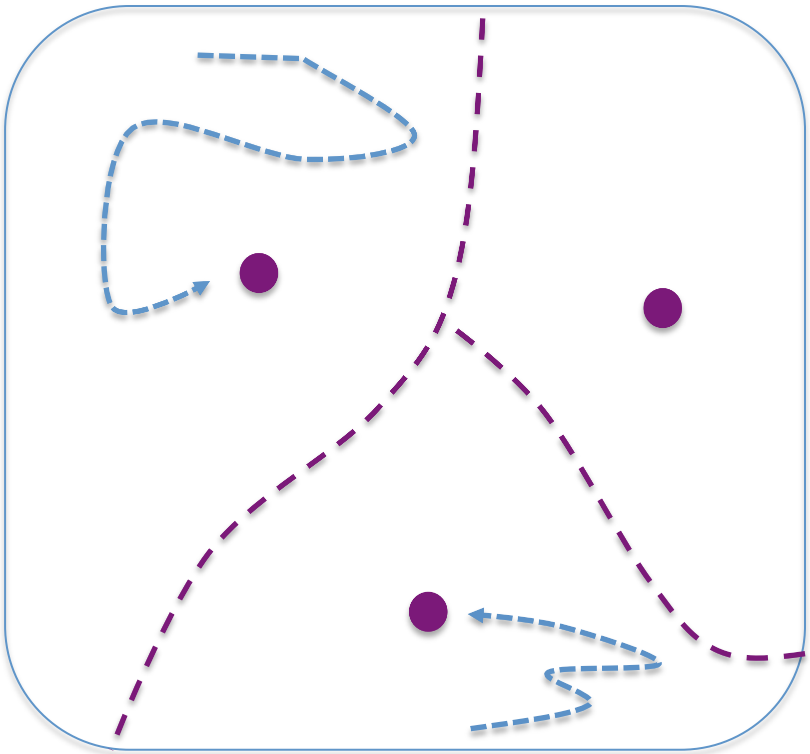
Embedding of low-dimensional manifolds in RNN

Part of Lecture 4 by R. Monasson



$$x_i(t+1) = sign\left(\sum_j J_{ij} x_j(t)\right)$$



Space of neural configurations x_i (N-dimensional)

Attractors in Recurrent Neural Nets

Set of neural configurations: $\left\{x_i^\mu\right\}, i = 1 \ldots N, \mu = 1 \ldots P$

Set of fixed point conditions: $x_i^\mu = sign\left(\sum_j J_{ij} x_j^\mu\right), \forall i, \mu$

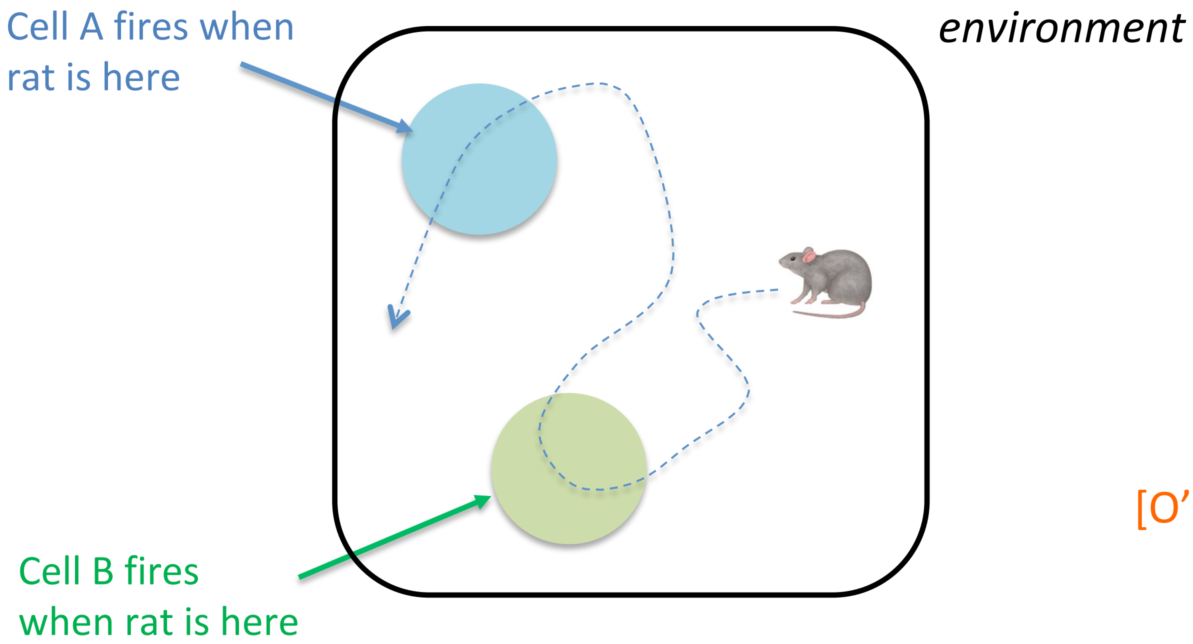
Equivalent to classification of P inputs/ouputs with N perceptrons (rows of J matrix)

Possible as long as $\alpha = \dfrac{P}{N} < 2$ if neural configurations are uncorrelated

NB: different from $y^\mu = sign\left(\sum_j J_j x_j^\mu\right), \forall \mu$

Place cells and place fields

Place cells in the hippocampus regions CA1 and CA3 present spatially localized firing fields



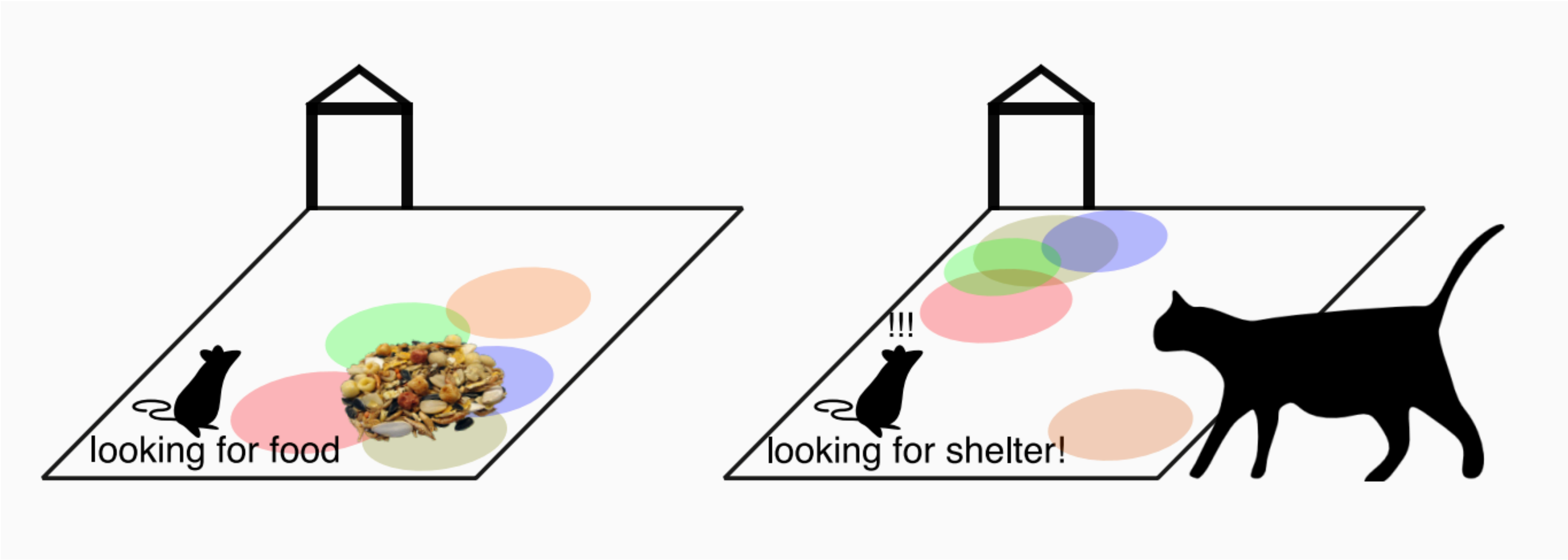
[O'Keefe et al. 1970's]

- Persists in the absence of input stimuli (dark)
- Place fields are retrieved when the animal is placed in the same environment after days

Multiple CANN

- Different representations of the same environment can be memorized and recalled upon contextual change on very short time scales

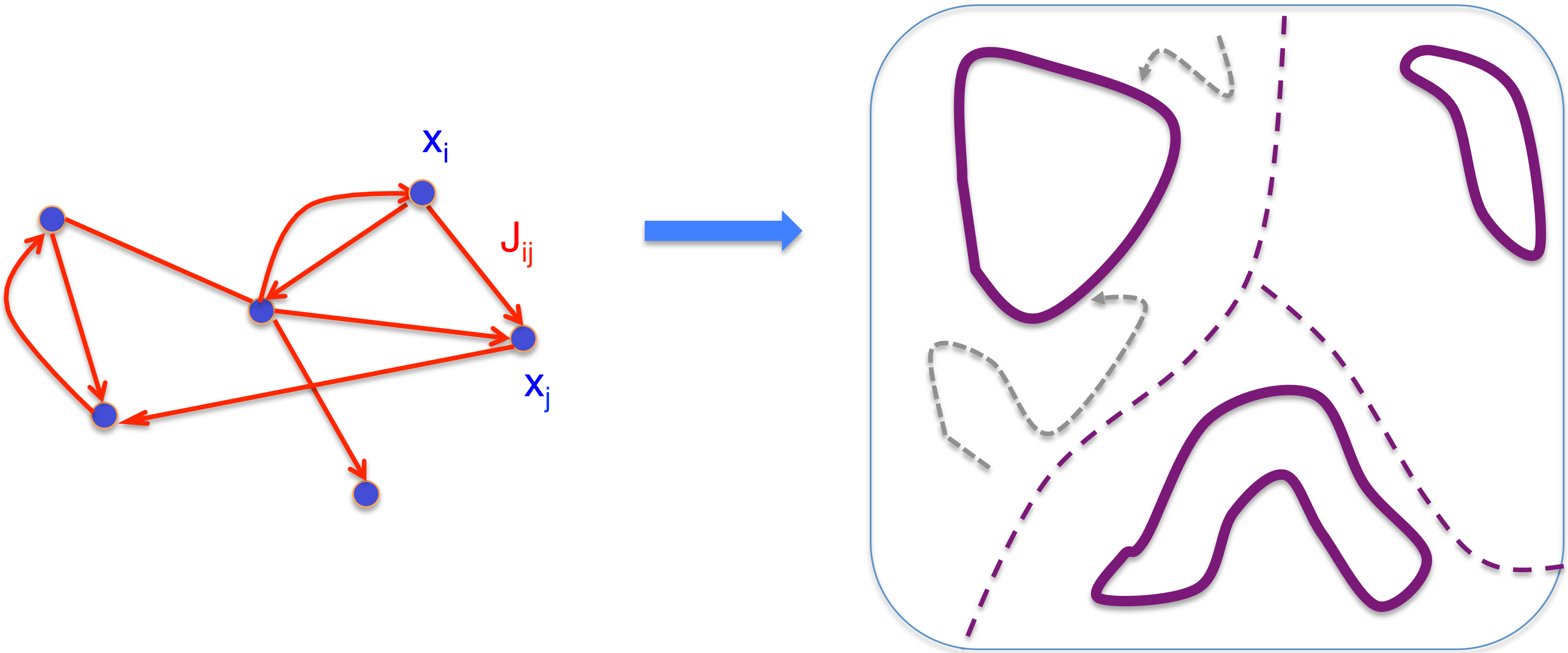
[Kelemen, Fenton, PLoS Biology 2010]



- Different environments need to be memorized in the same network; apparently random remapping of place fields across environment

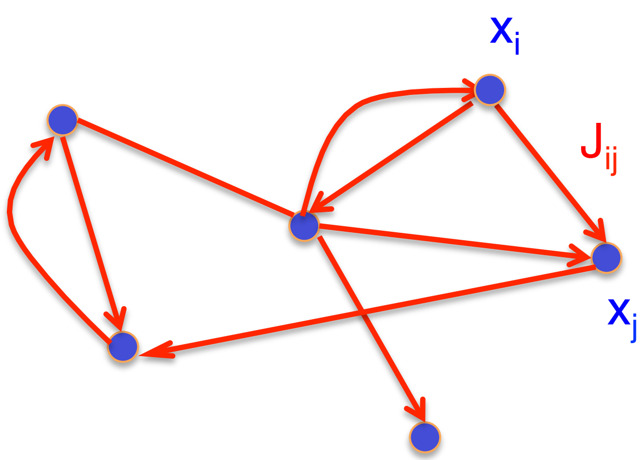
[Alme et al. “Place cells in the hippocampus: Eleven maps for eleven rooms.”, PNAS 2014]

Multiple Continuous Attractor Neural Networks

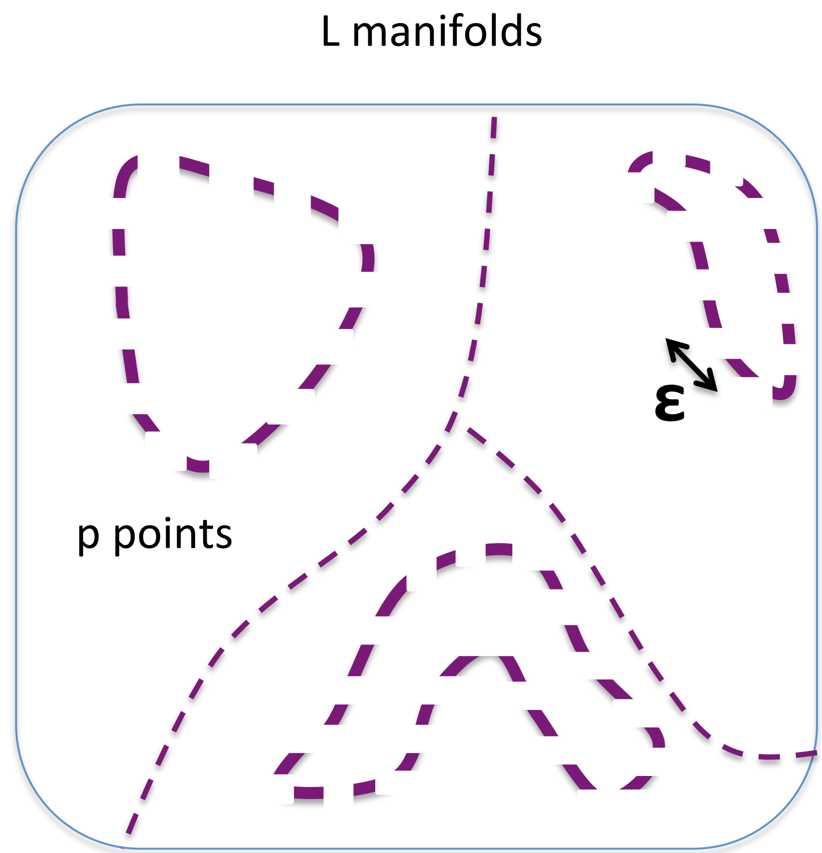


Space of neural
configurations x_i
(N-dimensional)

Multiple Continuous Attractor Neural Networks



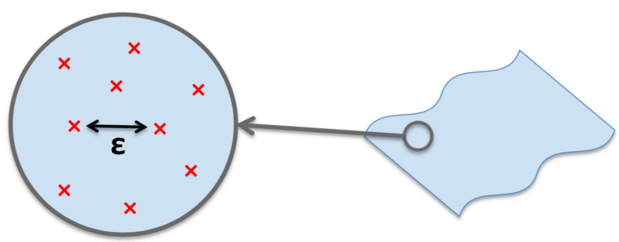
Trade-off between quantity (capacity) and quality (accuracy over « position » along attractor)



Space of neural configurations x_i (N-dimensional)

Formulation of the learning problem: data

- N neurons (i=1…N), L manifolds (μ=1…L) of dimension D
- Each neuron i has a place field in manifold μ centered in $\vec{r}_i^{(\mu)}$
- p anchoring points $\vec{r}^{(m)}$ (m=1...p) per manifold: controls accuracy ε of representation



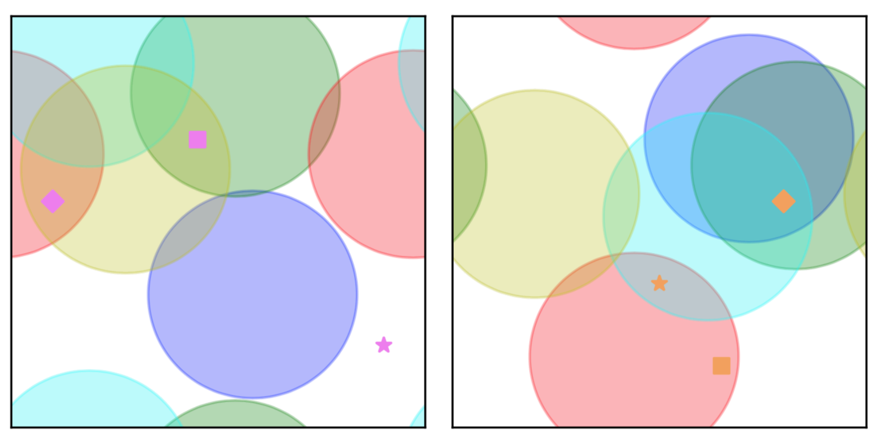
$$\varepsilon \approx p^{-1/D}$$

- L x p data configurations $$x_i^{(\mu,m)} = \Phi\left(\left|\vec{r}^{(m)} - \vec{r}_i^{(\mu)}\right|\right) = 0,1$$

Example:

N=5, L=2, D=2, p=3

(periodic boundary conditions)



★	■	◆	★	■	◆
0	0	0	0	0	1
0	0	0	1	0	1
0	1	0	0	0	1
0	0	1	1	1	0
0	1	1	0	0	0

Formulation of the learning problem: optimal couplings

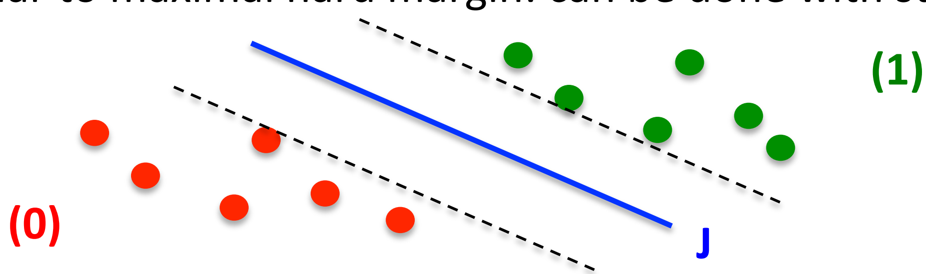
- RNN dynamics: $$\boxed{x_i(t+1) = f\left(\sum_j J_{ij}\ x_j(t)\right)}$$ with $$f(u) = 0 \text{ if } u < 0$$ $$1 \text{ if } u > 0$$

- Minimal conditions: data configurations should be fixed points

- Optimal set of couplings ***J*** of RNN maximizes

$$\boxed{\kappa(J) = \min_{\{i=1\ldots N,\mu=1\ldots L,m=1\ldots p\}} \left[(2x_i^{(\mu,m)} - 1) \sum_j J_{ij}\ x_j^{(\mu,m)} \right]}$$

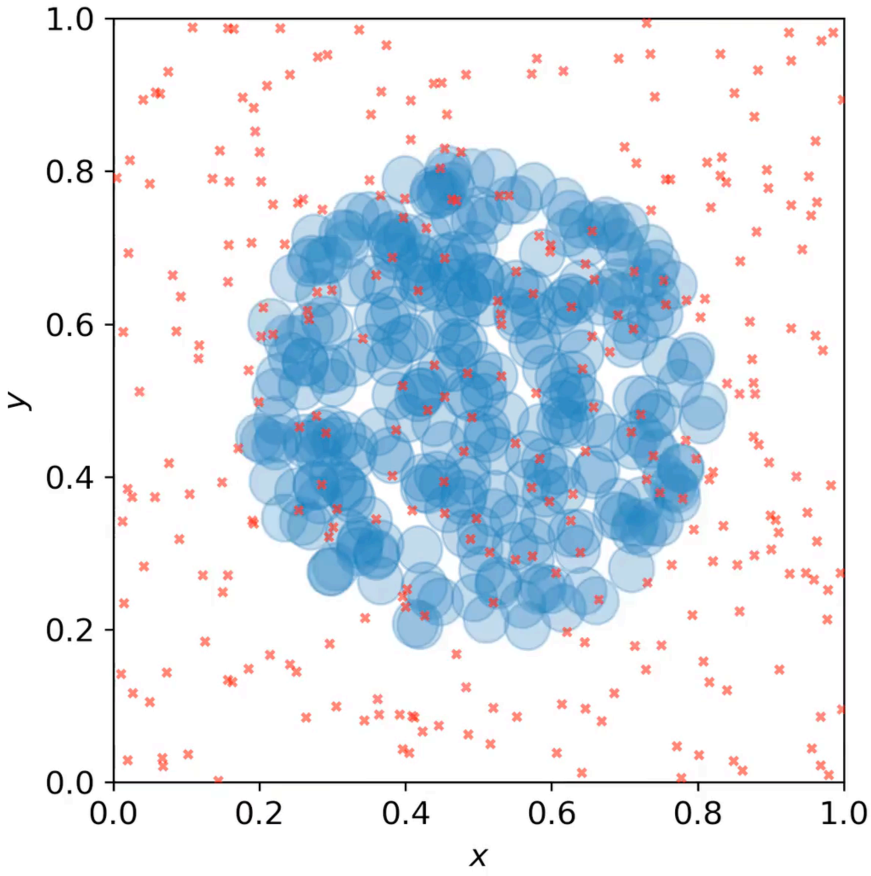
➢ Similar to maximal hard margin: can be done with standard SVM techniques



➢ Auto-associative mapping: unsupervised learning of the manifolds

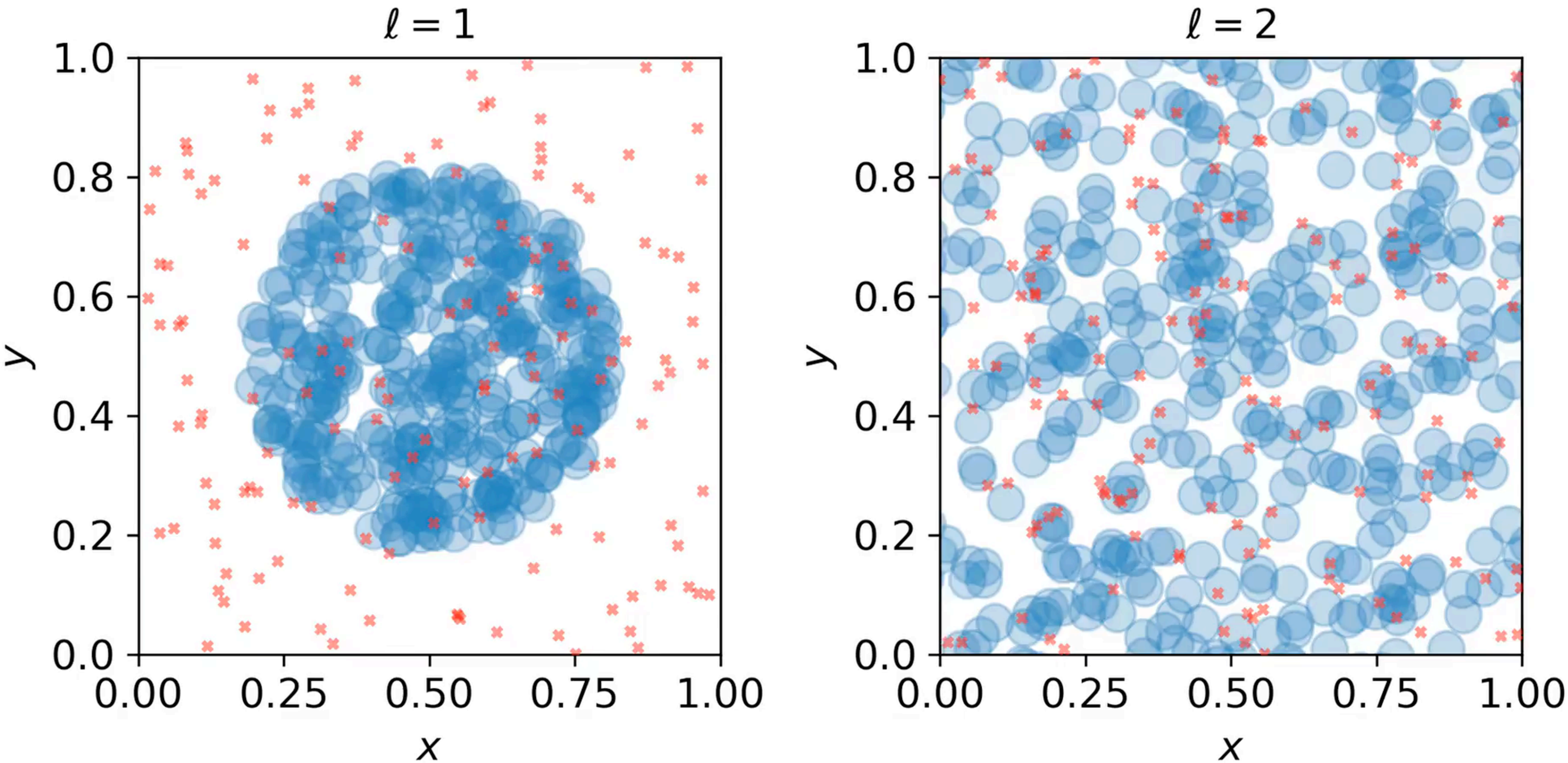
Results: bump spans manifold

N=1000
L=1
p=300



Here: noisy version of the dynamics of RNN

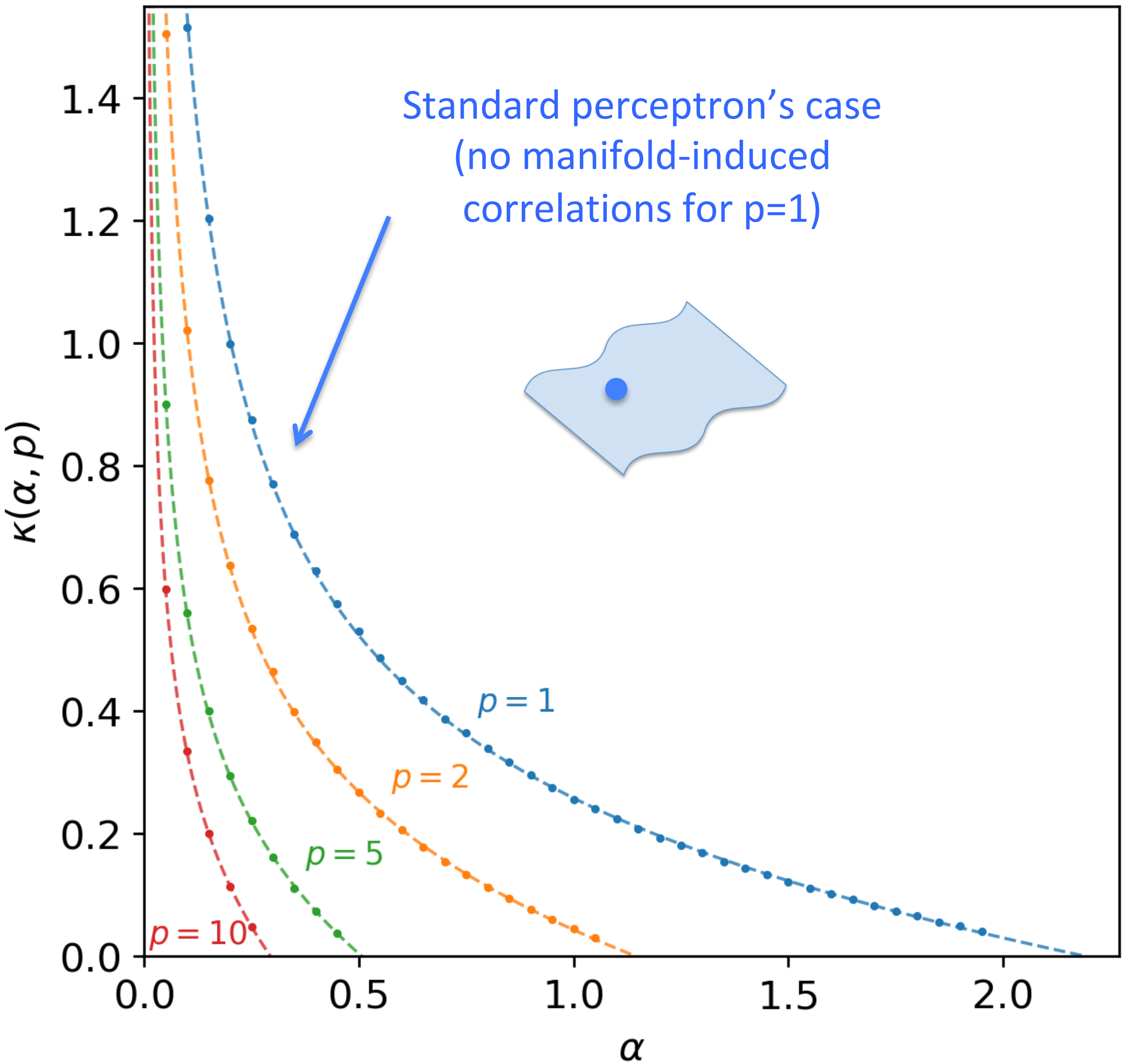
Results: transitions from manifold to manifold



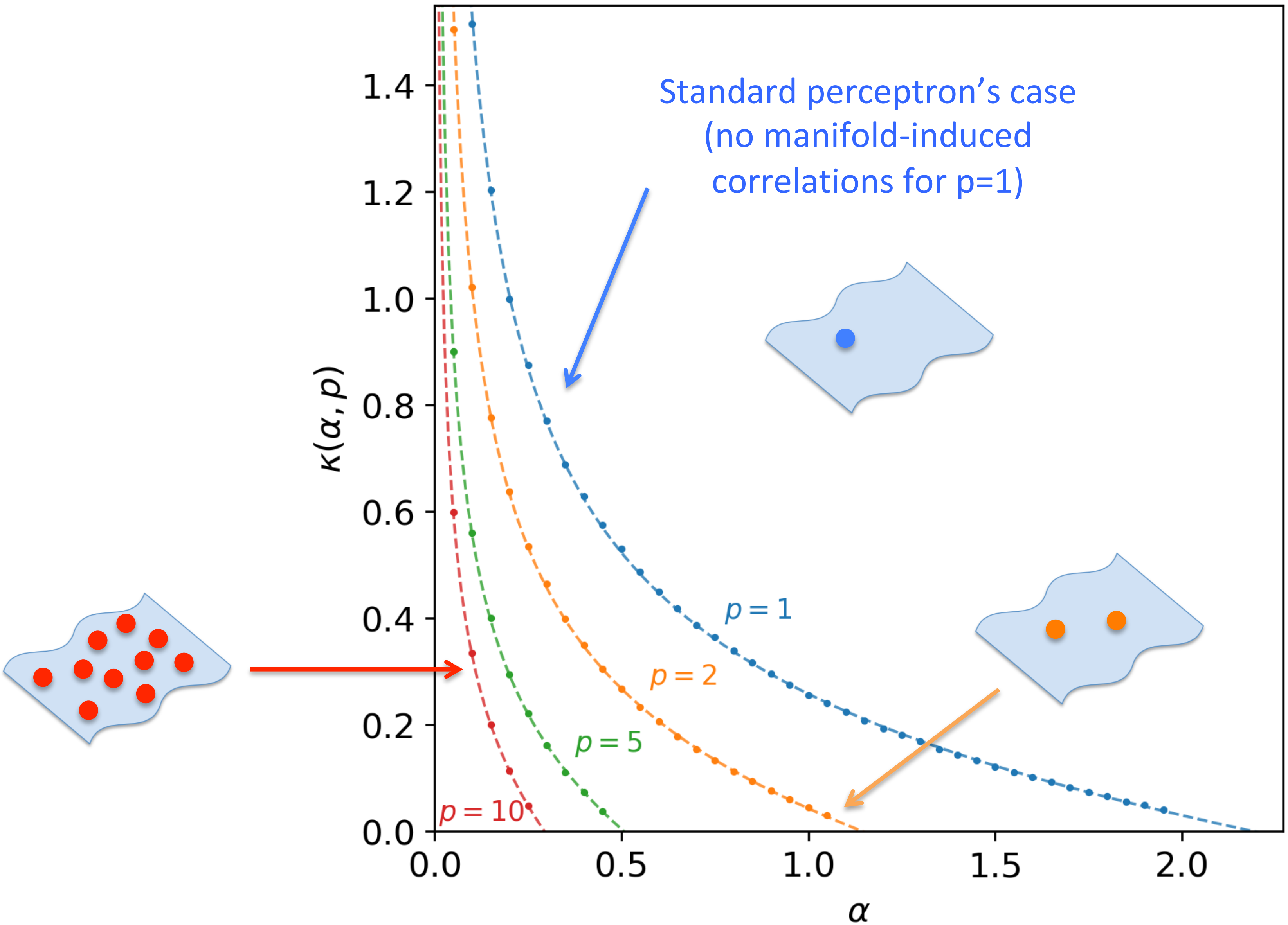
N=1000
L=2
p=150

Here: noisy version of the dynamics of RNN

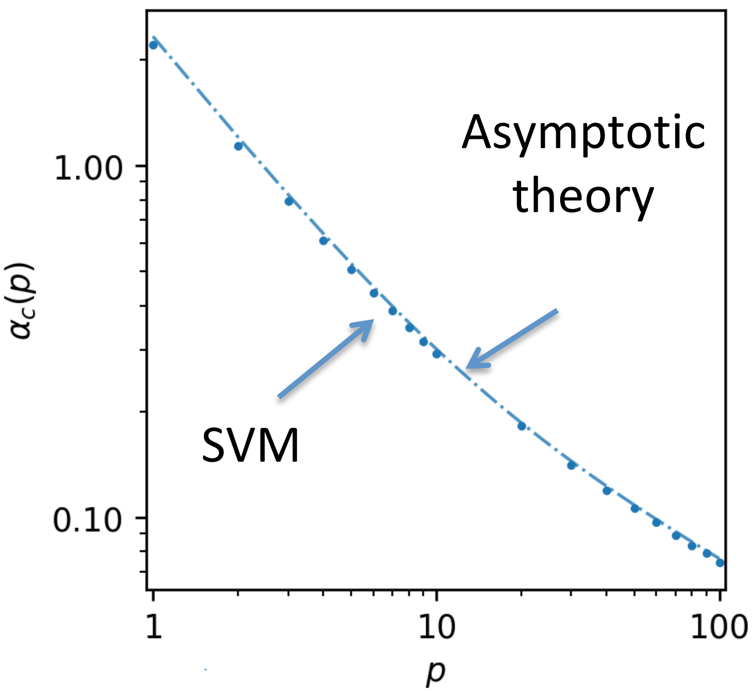
Optimal capacity



Optimal capacity



Optimal capacity: asymptotic theory



Large-p behavior:

$$\boxed{\alpha_c(p) \approx \frac{A(D,\Phi)}{(\log p)^D}}$$

Very slow decrease of capacity with spatial resolution $\varepsilon \approx p^{-1/D}$,

e.g.

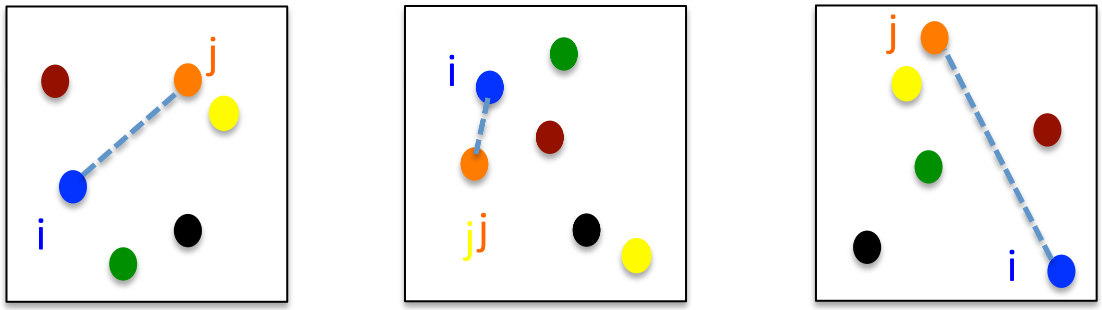
$$\varepsilon \to \varepsilon^2,\ \alpha_c \to \alpha_c / 2^D$$

[Battista & RM, Phys Rev Lett 2020]

Connection with Multi-space Euclidean Random Matrices

The result on the previous slide relies on the spectral properties of MERM:

[Battista & RM, Phys Rev E 2020]



...

$$C_{ij}\left(\left\{\vec{r}_i^{(\mu)}\right\}\right)=\frac{1}{L}\sum_{\mu=1}^{L}\gamma\left(\left|\vec{r}_i^{(\mu)}-\vec{r}_j^{(\mu)}\right|\right)$$

- Simple for L=1: high-density regime of ERM (eigenmodes ~ Fourier plane waves)
- Non trivial due to incoherent superimpositions of maps
- Self-consistent equation for the spectrum $\rho(\lambda;\alpha)$ (in fact, its resolvent) can be established with standard random matrix theory techniques