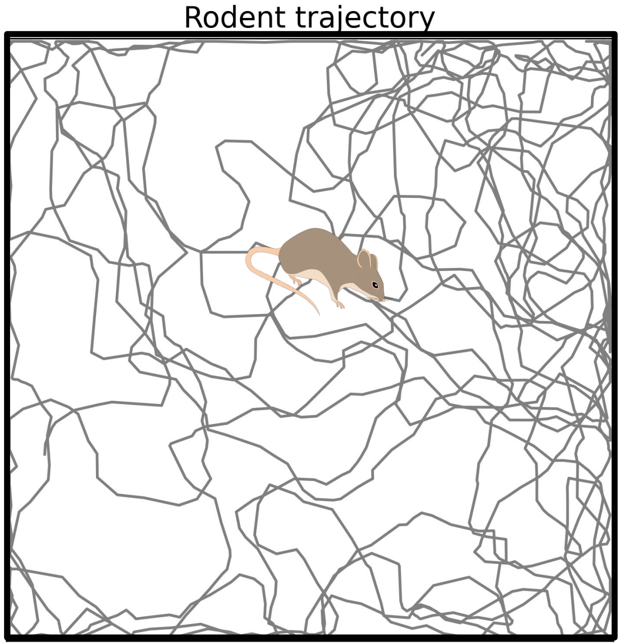
Binding in hippocampal-entorhinal circuits enables compositionality in cognitive maps

Chris Kymn, **Sonia Mazelet**, Anthony Thomas, Denis Kleyko, E.Paxon Frady, Friedrich T. Sommer, Bruno A. Olshausen

NeurIPS@Paris - December 4th 2024

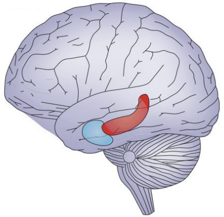
Context: neural representation of spatial position



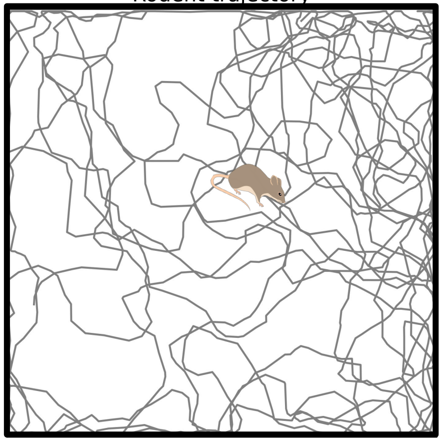
Spatial position encoding should have the following properties:

- **compositionality**
- robustness to noise
- high spatial resolution
- similarity preserving encoding

Representation of self spatial position

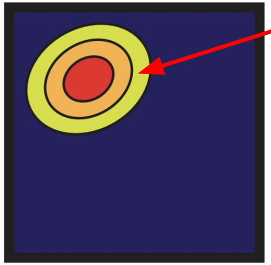


Rodent trajectory



Hippocampus

Place cells



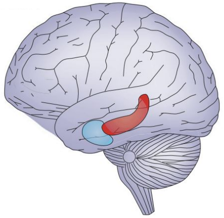
(O'keefe & Nadel 1978)

spatial location where
one cell fires

neurons fire in one
particular spatial location

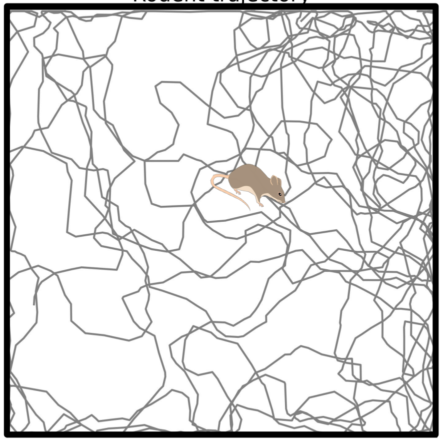
Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. Neuron, 100(2), 490-509.

Representation of self spatial position



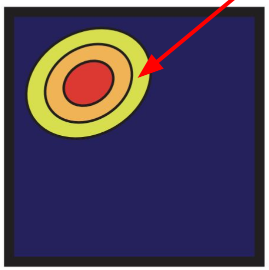
spatial locations where one cell fires

Rodent trajectory



Hippocampus

Place cells

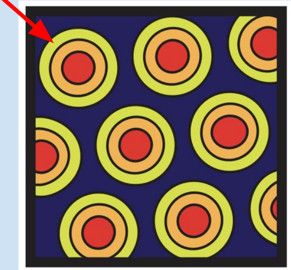


(O'keefe & Nadel 1978)

neurons fire in one particular spatial location

Medial entorhinal cortex

Grid cells



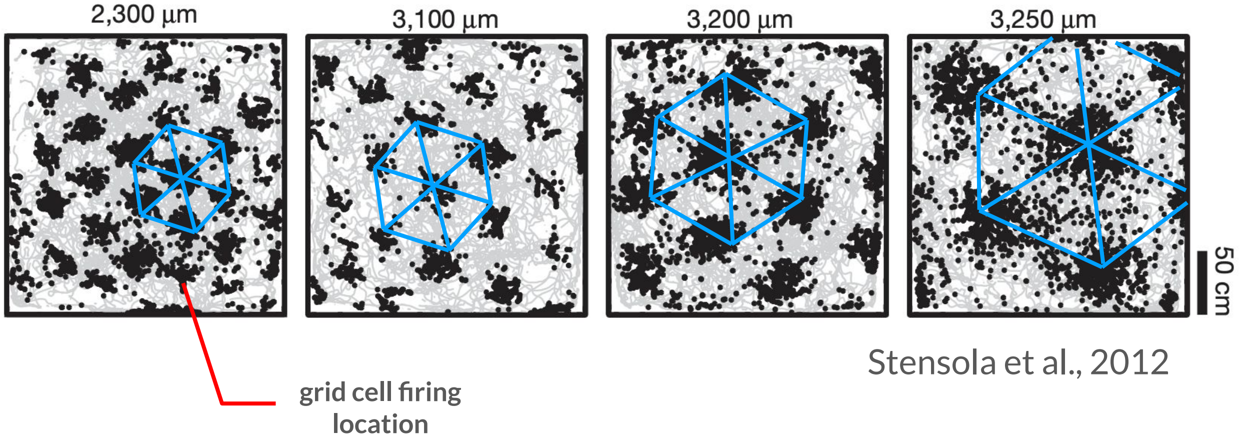
(Hafting et al., 2005)

neurons fire at many spatial locations on a triangular lattice

Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. Neuron, 100(2), 490-509.

Grid cells organisation in modules

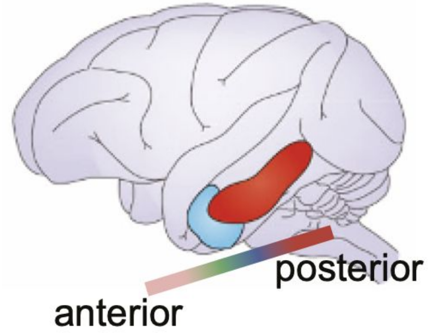
Grid cell activations in 2D space



Stensola et al., 2012

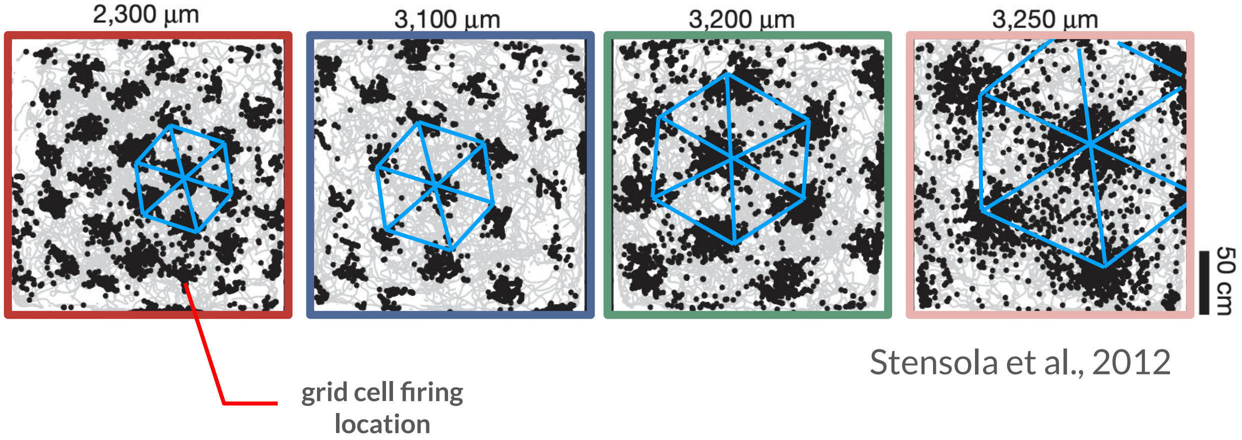
Grid cells are organized into modules. Each has many grid cells of (approximately) the same spatial scale.

Grid cells organisation in modules



Shpektor et al., 2024

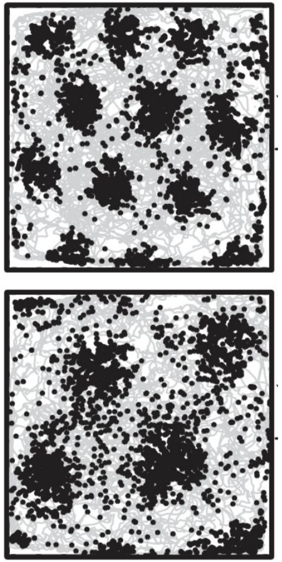
Grid cell activations in 2D space



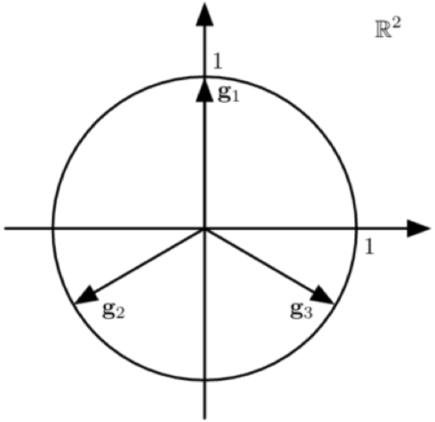
Stensola et al., 2012

Grid cells are organized into modules. Each has many grid cells of (approximately) the same spatial scale.

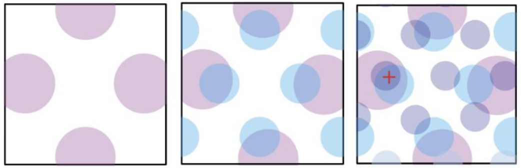
Questions



Why are grid cells organised into modules ?

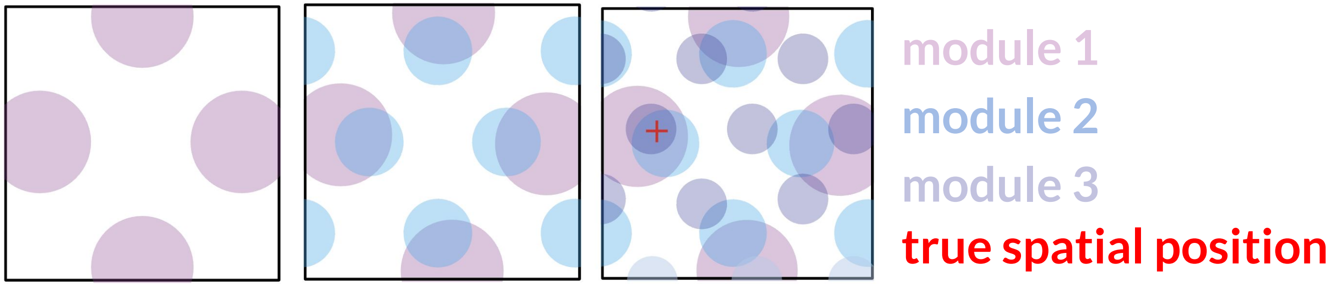


And why in a hexagonal lattice ?



How do grid cell modules coordinate their computations?

Grid cell modules give high precision spatial position



- Individual grid cell modules encode spatial position at varying scales.
- Combined, these modules enable accurate representation of precise 2D locations.

How to represent 1D position with random high dimensional vectors ?

- A representation of x is the binding (element wise multiplication) of high dimensional vectors of different spatial periodicities

$$p(x) = g_1(x) \odot g_2(x) \odot g_3(x)$$

High dimensional vectors of respective spatial periodicities m_1, m_2, m_3 (Random Fourier features)

How to represent 1D position with random high dimensional vectors ?

- A representation of x is the binding (element wise multiplication) of high dimensional vectors of different spatial periodicities

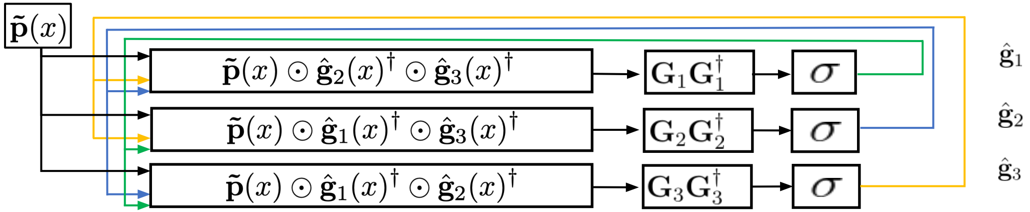
place cell

grid cell modules

$$\boxed{p(x)} = \boxed{g_1(x) \odot g_2(x) \odot g_3(x)}$$

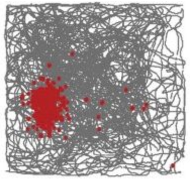
High dimensional vectors of respective spatial periodicities m_1, m_2, m_3 (Random Fourier features)

A Residue Number System Attractor Neural Network couples grid and place cells

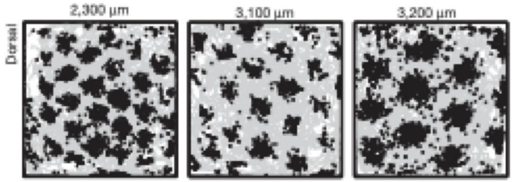


The resonator network iteratively factorizes the place cell representation into its grid cell representations according to different modules

$\mathbf{p}(x)$

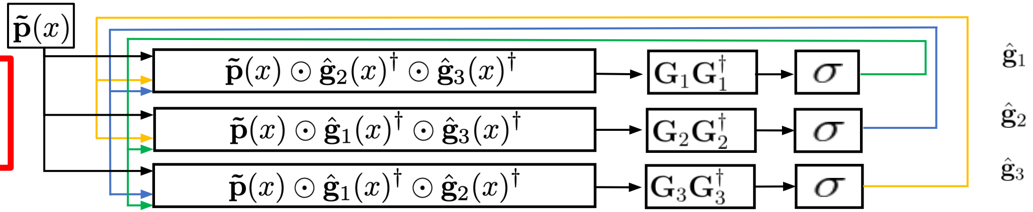


$\hat{\mathbf{g}}_1(x)$ $\hat{\mathbf{g}}_2(x)$ $\hat{\mathbf{g}}_3(x)$



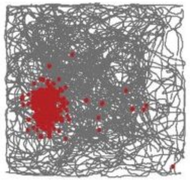
A Residue Number System Attractor Neural Network couples grid and place cells

0. Hippocampal input

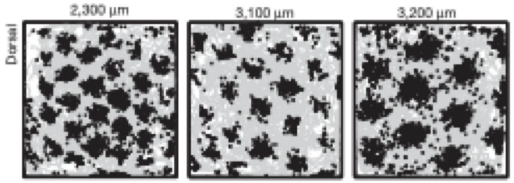


The resonator network iteratively factorizes the place cell representation into its grid cell representations according to different modules

$\mathbf{p}(x)$

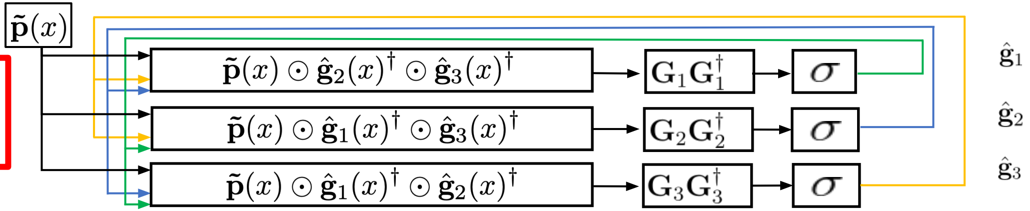


$\hat{\mathbf{g}}_1(x)$ $\hat{\mathbf{g}}_2(x)$ $\hat{\mathbf{g}}_3(x)$



A Residue Number System Attractor Neural Network couples grid and place cells

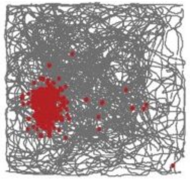
0. Hippocampal input



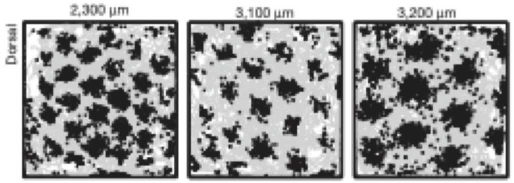
1. Unbinding the estimates of other modules

The resonator network iteratively factorizes the place cell representation into its grid cell representations according to different modules

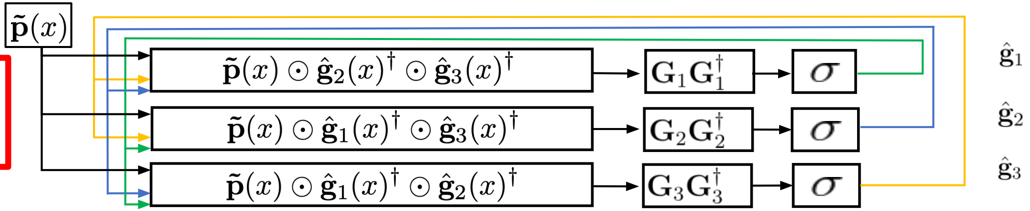
$\mathbf{p}(x)$



$\hat{\mathbf{g}}_1(x)$ $\hat{\mathbf{g}}_2(x)$ $\hat{\mathbf{g}}_3(x)$



A Residue Number System Attractor Neural Network couples grid and place cells



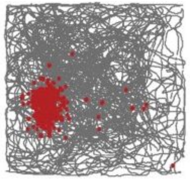
0. Hippocampal input

1. Unbinding the estimates of other modules

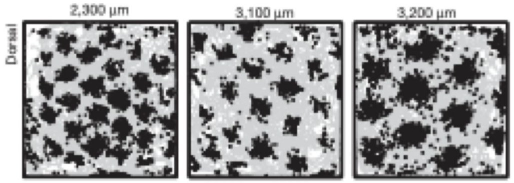
2. Intra-module attractor dynamics

The resonator network iteratively factorizes the place cell representation into its grid cell representations according to different modules

$\mathbf{p}(x)$



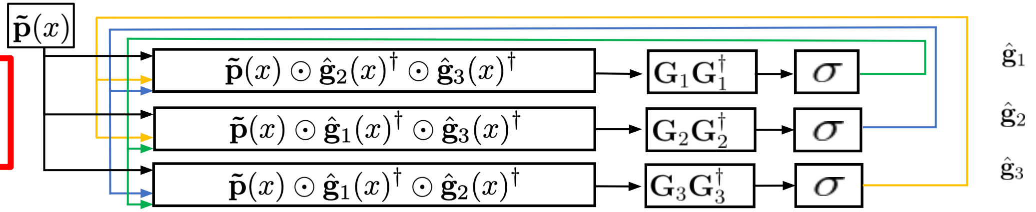
$\hat{\mathbf{g}}_1(x)$ $\hat{\mathbf{g}}_2(x)$ $\hat{\mathbf{g}}_3(x)$



A Residue Number System Attractor Neural Network couples grid and place cells

3. Repeat steps 1 and 2

0. Hippocampal input

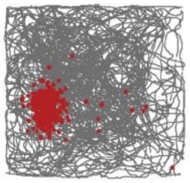


1. Unbinding the estimates of other modules

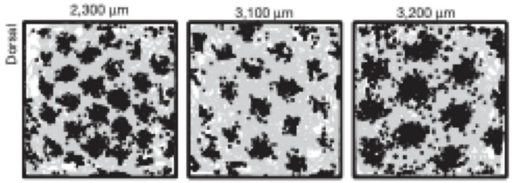
2. Intra-module attractor dynamics

The resonator network iteratively factorizes the place cell representation into its grid cell representations according to different modules

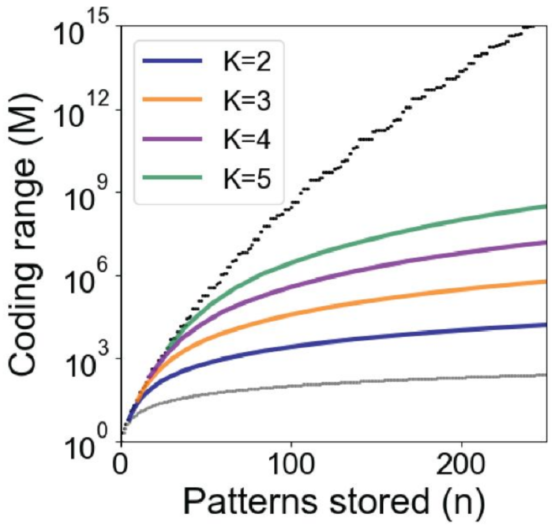
$\mathbf{p}(x)$

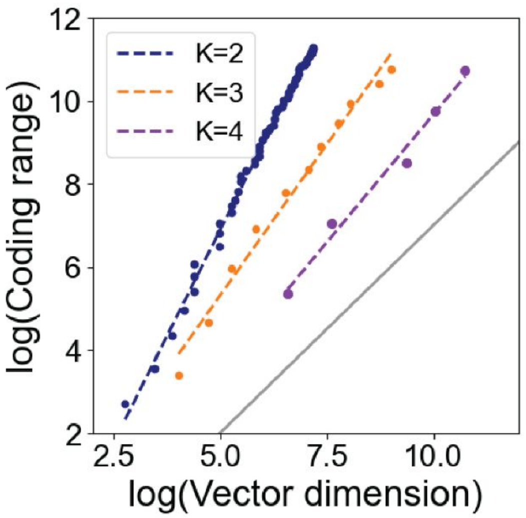


$\hat{\mathbf{g}}_1(x)$ $\hat{\mathbf{g}}_2(x)$ $\hat{\mathbf{g}}_3(x)$



Scaling laws of RNS attractor networks





Coding range scales as

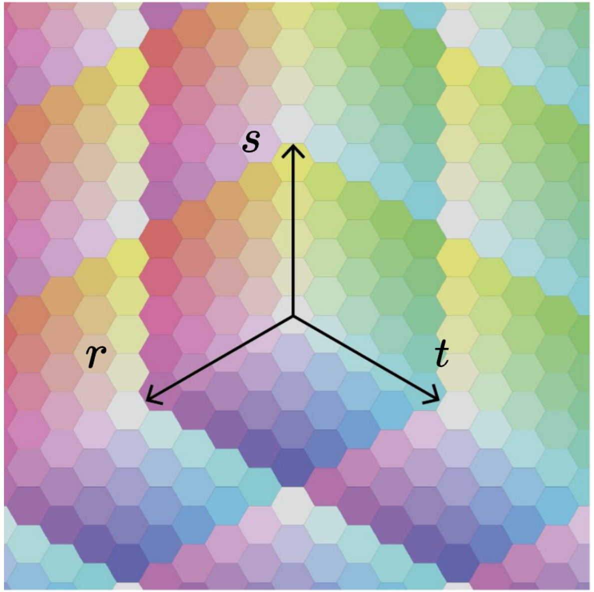
- e^K **in number of modules** K
- $\sim \exp\left(\sqrt{n}\ \ln n\right)$ **in number of patterns** n

Attractor network capacity scales superlinearly in vector dimension

Coding range: product space of patterns

Capacity : maximum number of patterns that can be stored and correctly recovered by resonator with high probability

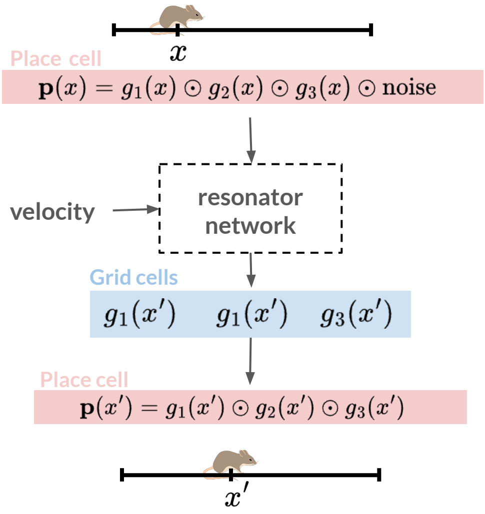
What about 2D ?

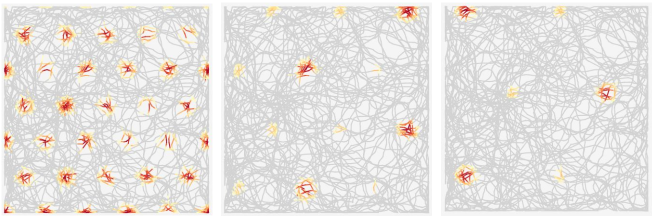


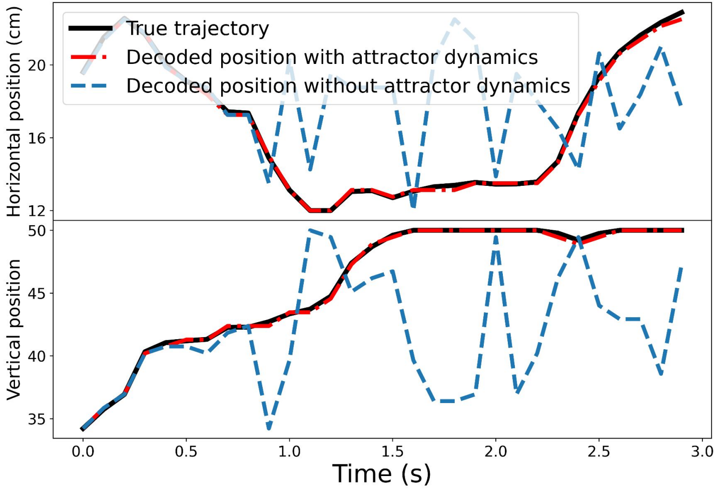
- Composition (binding) of the 1D representations along 3 axis.
- Voronoi tesselation of triangular frame has $3m^2 - 3m + 1$ states, compared to m^2 for square frame: **higher spatial resolution**
- 2D place cell representation is the binding of the 1D representations along the 3 axis

$$p(r, s, t) = p_r(r) \odot p_s(s) \odot p_t(t)$$

Path integration results

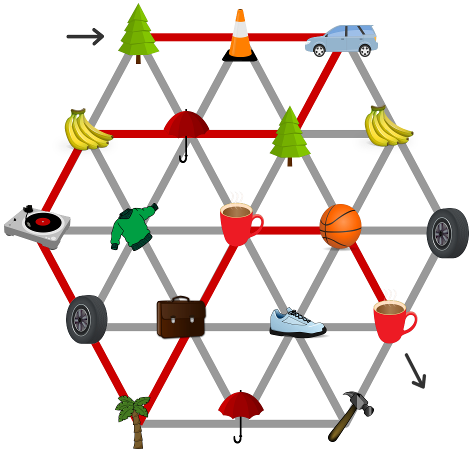


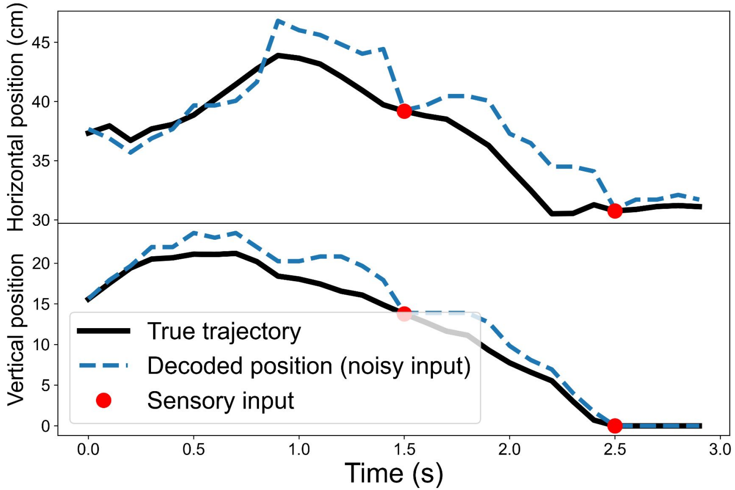




- Attractor dynamics **limit noise accumulation** along the 2D spatial trajectory.
- Encoding with triangular coordinate systems produce **hexagonal receptive fields**.

Path integration in conceptual spaces





Sequence retrieval via path integration in a conceptual space

Composition of information from sensory inputs and position corrects drift

Thank you !