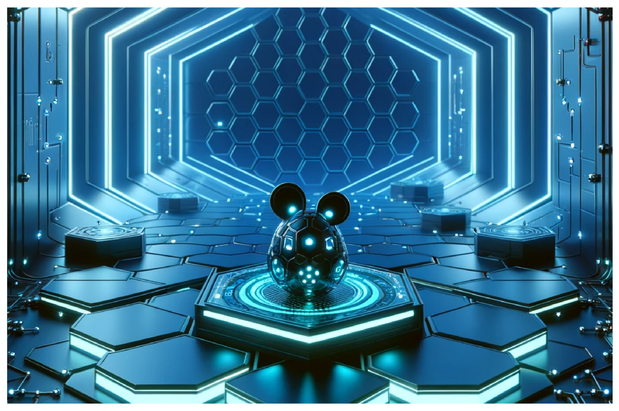
Testing Assumptions Underlying a Unified Theory for the Origin of Grid Cells

Rylan Schaeffer [1] * Mikail Khona [2] * Adrian Bertagnoli [3] Sanmi Koyejo [1] Ila Rani Fiete [3 4]

*Denotes Equal Contribution [1]Stanford Computer Science [2]MIT Physics [3]MIT Brain and Cognitive Sciences [4]McGovern Institute for Brain Research

Summary

- Mammals have specialized neural representations that encode space.
- Grid cells, whose discovery earned a Nobel prize, are a striking example: a grid cell is a neuron that fires if and only if the animal is spatially located at the vertices of a regular triangular lattice that tiles all explored two-dimensional environments.
- Recent work has proposed a "unified theory for the computational and mechanistic origin of grid cells" [7, 8], claiming the mammalian lineage has learned grid cells due to properties of place cells.
- We identify two critical mathematical assumptions about place cells made by the Unified Theory & find that neither of the assumptions are likely true in biological neural representations.



Background: The Mammalian Neural Code for Physical Space

Grid Cells in Medial Entorhinal Cortex

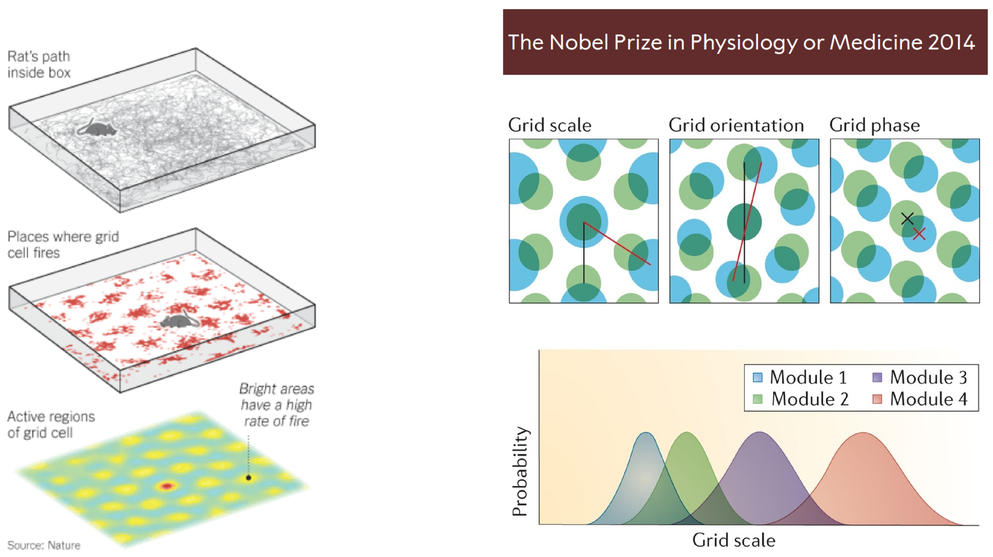


Figure 1. Figures from [6].

Place Cells in Hippocampus

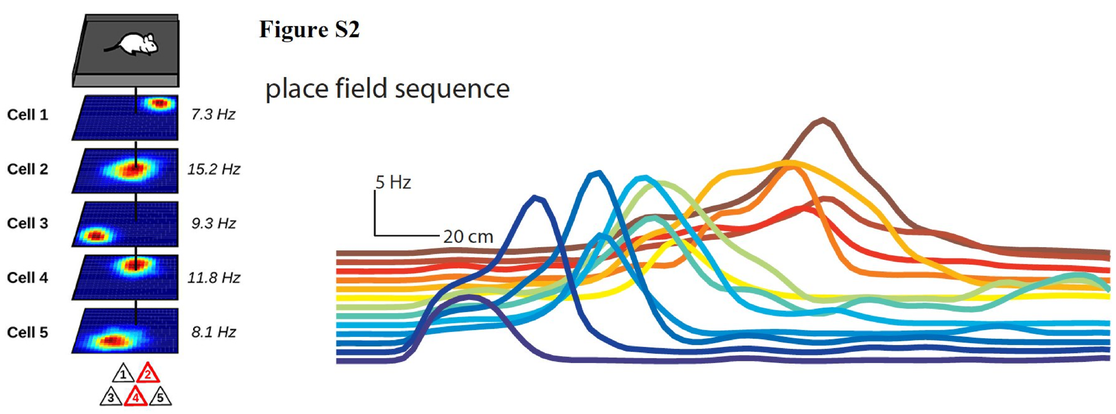


Figure 2. Figures from [4, 2].

Mathematical Assumptions of "Unified Theory for Origin of Grid Cells" [7, 8]

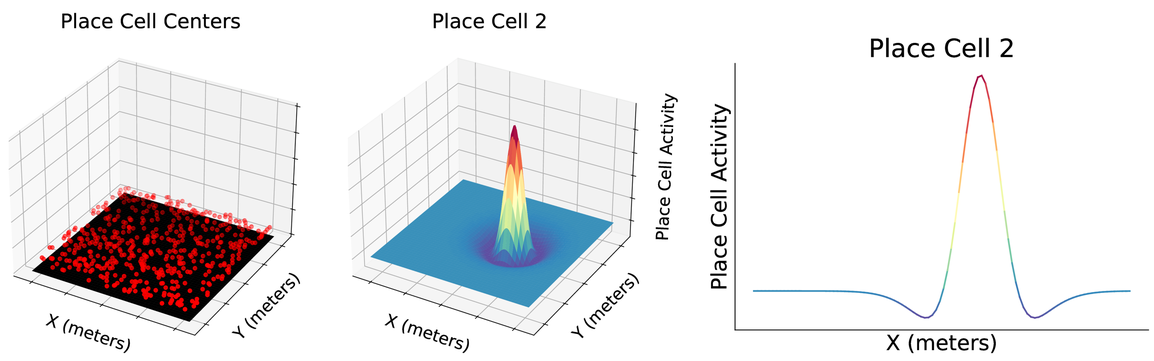


Figure 3. **Two key assumptions of the Unified Theory.** Left: Readouts, as a population, must be translationally invariant. Center and Right: Readouts, individually, must have center-surround tuning curves.

1. The readouts, as a population, must be translationally invariant.
2. The readouts, individually, must have carefully tuned center-surround tuning curves: either Difference-of-Softmaxes (DoS) or a particular Difference-of-Gaussians (DoG) tuning curve shape, i.e, a vector in $\mathbb{R}^{N_p}$ whose $i-$th entry is:

$$Softmax\Big(-\frac{1}{2\sigma_E^2}||x_t - p_i||^2\Big) - Softmax\Big(-\frac{1}{2\sigma_I^2}||x_t - p_i||^2\Big)$$

Place Cells Likely Lack Translational Invariance

Quantify how "translationally invariant" a population of place cells is: $\Pi_{\mathcal{T}}(\Sigma_i) \stackrel{\text{def}}{=} \arg\min_{T \in \mathcal{T}} ||T - \Sigma_i||_F^2$

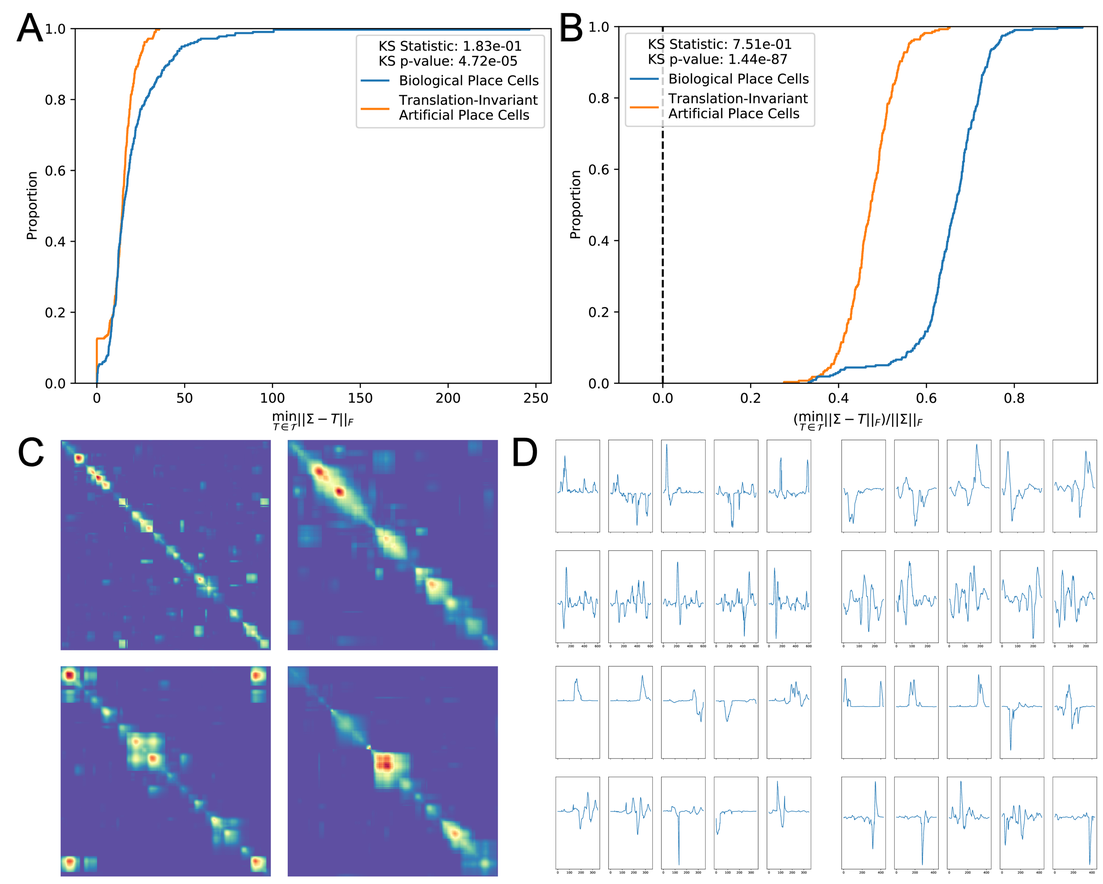


Figure 4. **Biological place cells populations are likely not translationlly invariant, as required mathematically by the Unified Theory.** Biological place cell populations from 320 recorded sessions [5] deviate significantly from a null distribution of translation-invariant artificial place cell population used to train deep networks, as measured by (A) the matrix distance (Kolmogorov-Smirnov 2-sample: $p = 4.72\text{e}{-5}$) and the (B) matrix absolute percent error (Kolmogorov-Smirnov 2-sample: $p = 1.44\text{e} - 87$). (C) Place cell spatial autocorrelation matrices Σ_i do not visually display constant diagonals of Toeplitz matrices (shown: 4 randomly chosen sessions). (D) Corresponding sessions' spatial autocorrelation matrices have non-periodic eigenvectors.

Place Cells Likely Lack Center-Surround Tuning

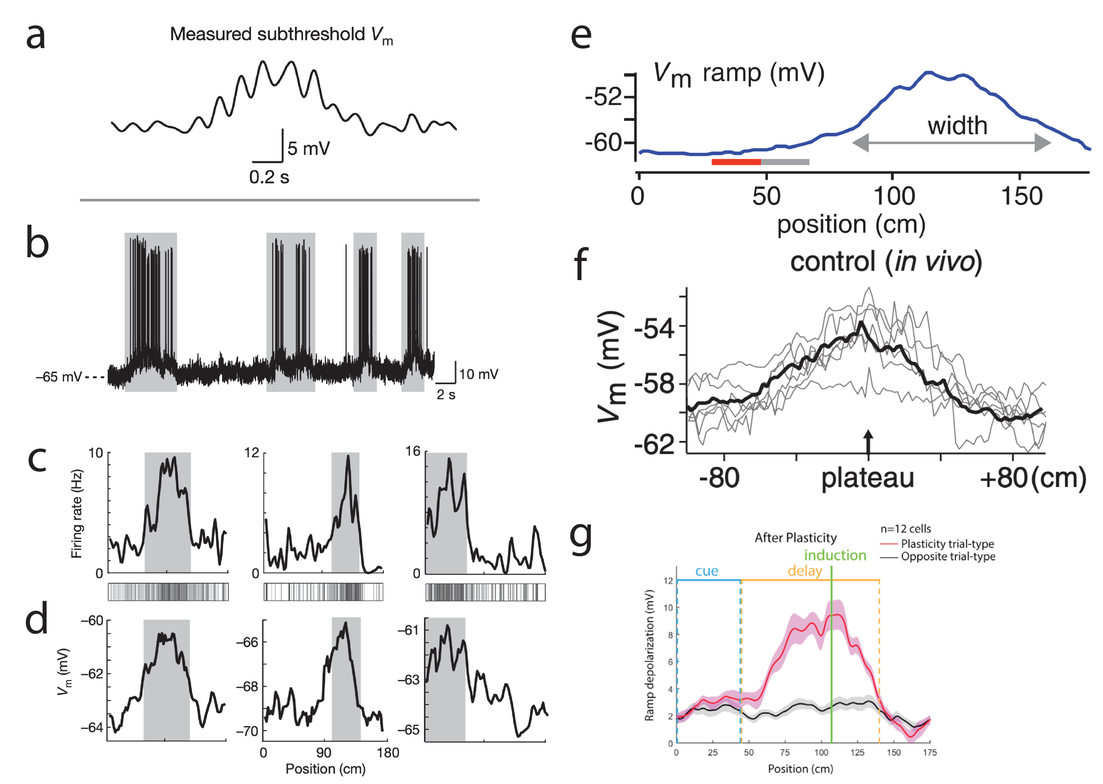


Figure 5. **Subthreshold voltages of place cells do not display DoG/DoS tuning.** (a-g) Measured membrane voltage from a selection of place cells in various conditions. Panels a-d reproduced from [3], e-f reproduced from [1] and g reproduced from [9] with permission.

References

[1] Katie C Bittner, Aaron D Milstein, Christine Grienberger, Sandro Romani, and Jeffrey C Magee. Behavioral time scale synaptic plasticity underlies ca1 place fields. *Science*, 357(6355):1033–1036, 2017.

[2] Kamran Diba and György Buzsáki. Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, 10(10):1241–1242, October 2007. Number: 10 Publisher: Nature Publishing Group.

[3] Christopher D Harvey, Forrest Collman, Daniel A Dombeck, and David W Tank. Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature*, 461(7266):941–946, 2009.

[4] Patrick Latuske, Olga Kornienko, Laura Kohler, and Kevin Allen. Hippocampal remapping and its entorhinal origin. *Frontiers in behavioral neuroscience*, 11:253, 2018.

[5] Jae Sung Lee, John J Briguglio, Jeremy D Cohen, Sandro Romani, and Albert K Lee. The statistical structure of the hippocampal code for space as a function of time, context, and value. *Cell*, 183(3):620–635, 2020.

[6] Edvard I Moser, Yasser Roudi, Menno P Witter, Clifford Kentros, Tobias Bonhoeffer, and May-Britt Moser. Grid cells and cortical representation. *Nature Reviews Neuroscience*, 15(7):466–481, 2014.

[7] Ben Sorscher, Gabriel C Mel, Surya Ganguli, and Samuel A Ocko. A unified theory for the origin of grid cells through the lens of pattern formation. *NeurIPS*, 2019.

[8] Ben Sorscher, Gabriel C Mel, Surya Ganguli, and Samuel A Ocko. A unified theory for the computational and mechanistic origins of grid cells. *Neuron*, 2022.

[9] Xinyu Zhao, Ching-Lung Hsu, and Nelson Spruston. Rapid synaptic plasticity contributes to a learned conjunctive code of position and choice-related information in the hippocampus. *Neuron*, 110(1):96–108, 2022.