Identifying representational structure in CA1 to benchmark theoretical models of cognitive mapping.

Authors: J. Quinn Lee[1,*,†], Alexandra T. Keinath[1,2,†], Erica Cianfarano[3], Mark P. Brandon[1,3,*]

Affiliations:

[1] Department of Psychiatry, Douglas Hospital Research Centre, McGill University; Montreal, Quebec, Canada

[2] Department of Psychology, University of Illinois Chicago; Chicago, Illinois, USA

[3] Integrated Program in Neuroscience, McGill University; Montreal, Quebec, Canada

[*] Corresponding authors: jquinnlee@gmail.com and mark.brandon@mcgill.ca

[†] These authors contributed equally to this work.

Abstract: Decades of theoretical and empirical work have suggested the hippocampus instantiates some form of a cognitive map, yet tests of competing theories have been limited in scope and largely qualitative in nature. Here, we develop a novel framework to benchmark model predictions against observed neuronal population dynamics as animals navigate a series of geometrically distinct environments. In this task space, we show a representational structure that generalizes across brains, effectively constrains biologically viable model parameters, and discriminates between competing theories of cognitive mapping.

Main Text

The hippocampus and associated regions of the neocortex are thought to support diverse learning and memory processes through cognitive maps instantiated in the activity of principal neurons [1]. Growing evidence suggests that such maps are formed across metric spaces to support flexible goal-directed navigation and long-term memory [2, 3]; functions that are impaired following hippocampal damage [4-7]. These findings have recently motivated an increasing number of theories expressed in computational models to explain how the hippocampal system instantiates cognitive maps [8]. However, there has been no consensus on how to compare predictions across models and, importantly, against empirical observation.

To further our understanding of the organization of neural codes throughout the brain, we require experiments to make quantitative comparisons across regions, assays, and against theoretical models. One efficient approach toward this end is to collect large datasets in common task spaces where changes in neural representation can be robustly quantified across subjects, and directly compared to model predictions [9]. While the hippocampus encodes a wide range of behavioral and sensory conditions, among the strongest determinants of hippocampal coding is environmental geometry [10]. Indeed, extensive behavioral studies have shown that geometry predicts navigation behavior across species [11, 12], and neurophysiological studies further demonstrate that spatial codes across the hippocampal system remap following changes to environment shape [10, 13-17]. Geometric deformation is a common experimental approach to examine these determinants of neural coding, in which environmental boundaries are added, removed, or shifted. Importantly, geometry can be parametrically varied to create condition-rich experimental designs in which theories of cognitive mapping make competing predictions.

We recorded from large populations in hippocampal subregion CA1 in a condition-rich geometric deformation paradigm (5,413 unique neurons across 207 sessions in 10 geometries, forming 69,744 rate maps). Leveraging a representational-similarity framework, we show these data discriminate between competing model predictions and constrain neurobiologically viable model parameters – providing the first quantitative benchmark for theoretical advances in cognitive mapping research.

Cognitive mapping of a dynamic grid environment across protracted experience

Geometric deformation procedures are typically designed to examine changes in neural coding and behavior between two geometries (e.g., square versus circle); requiring the navigator to explore

each environment or linear "morphs" between two reference geometries [10, 13, 18-20]. While this approach has been highly useful to uncover how geometry determines behavior and neural coding, model predictions can be more clearly distinguished by increasing the number and diversity of deformations performed in a single task.

To this end, we partitioned an open square (75 x 75 cm) into a 3 x 3 grid-space to allow systematic manipulations of environmental geometry. We recorded large neural populations in CA1 with miniscope calcium imaging while mice freely explored a sequence of 10 geometrically distinct environments (Fig. 1A-B, S1) [21, 22]. Importantly, we selected these distinct geometries in effort to induce strong geometric remapping within the same environment across the entire CA1 population. Following habituation to the square arena, animals were exposed to a random sequence of geometries across days that started and ended with the square. The same sequence was repeated up to three times within animals but was randomized across mice (Fig. 1C-D).

To measure the quality of spatial coding across recordings, we calculated the split-half spatial reliability for registered cells and position decoding error with a naive Bayesian method [23, 24]. Across one month of recording, we observed an increase in the split-half reliability of recorded CA1 populations, and a resulting increase in the proportion of cells identified as place cells (Fig. 1E, S2). In keeping with this result, we also found a significant decrease in position decoding error across recordings, achieving the theoretical maximum level of decoding accuracy as suggested in recent work (Fig. 1F, S2) [25]. These results demonstrate that our CA1 recordings offer a high degree of spatial reliability and thus motivated the inclusion of all cells in subsequent analyses (mean number of cells per animal = 773 ± 68 SE, minimum cells per animal = 515, maximum cells per animal = 952).

To confirm remapping in CA1 following geometric deformation [10, 17, 19], we generated vectoral field maps of the spatial bins containing maximum population vector correlations across all geometric deformations relative to the square environment (Fig. 1G). The resulting maps demonstrated a local influence of deformations on population similarity, with displacement of spatial coding that scaled with the distance to inserted walls (Fig. 1G). The same dynamics were not observed when rate maps from the square environment were masked to simulate stable spatial mapping with identical occlusions (Fig. S3). These results confirm strong geometric remapping within the same environment across the entire recorded CA1 population, and thus provide a rich basis to examine the geometric determinants of neural coding across brains and against model predictions.

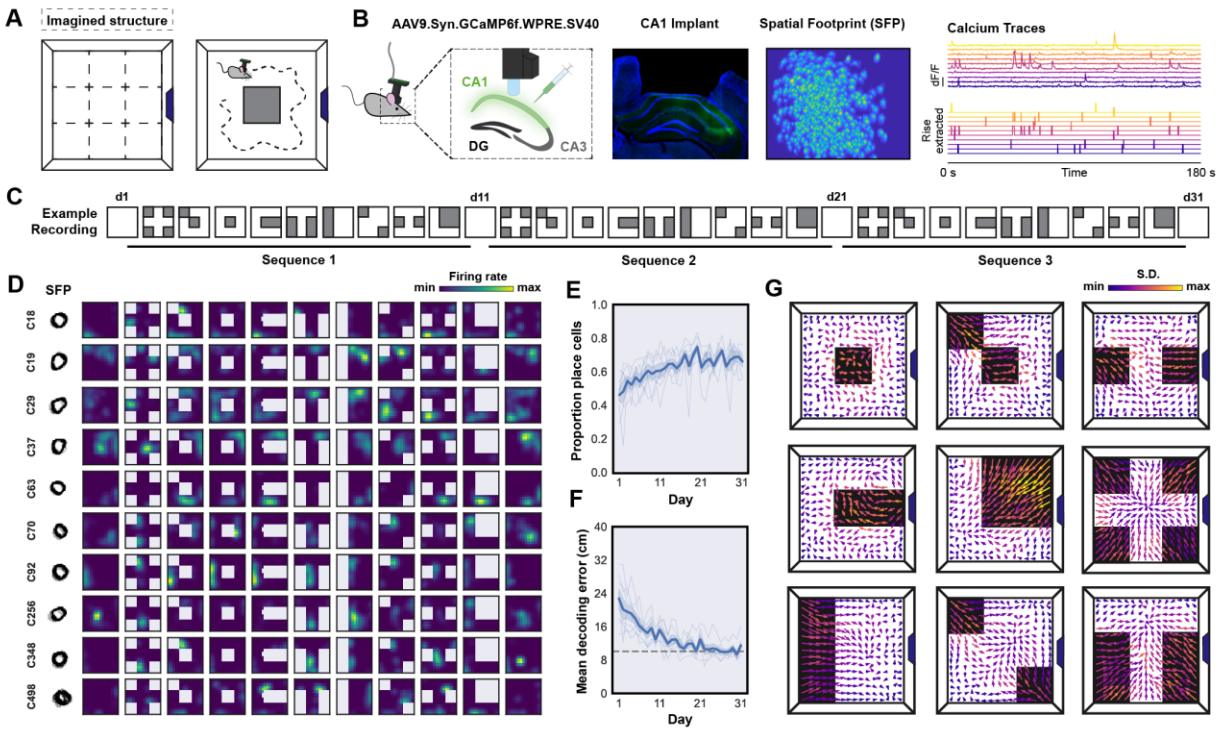

Fig. 1. Recording large neuronal populations in CA1 during geometric deformation. (**A**) To perform systematic manipulations of environmental geometry, we treated a square open field (75 x 75 cm) as an imagined 3 x 3 grid structure, blocking select partitions of the grid with 25 cm walls while mice freely navigated the open field. (**B**) To record large neuronal populations in CA1, mice were injected with a viral construct to express GCaMP6f in dorsal CA1 and implanted with a gradient-refractive index (GRIN) lens targeting CA1 (Methods). Spatial footprints and calcium transients were processed and binarized to create a firing rate vector for all recorded cells. (**C**) Following pre-exposure to the environment, we recorded a repeating sequence of randomly ordered geometries across days, starting and ending with the square environment for each sequence for up to three total repetitions (31 days). (**D**) Example rate maps of randomly selected cells registered on all days from a single sequence. The outline of aligned spatial footprints for each cell are shown (left) with rate maps for a target cell across deformations scaled to the maximum average firing across days (right). (**E**) The proportion of cells that were split-half reliable ($P<0.01$) for each animal and the group average across all recorded sessions. We observed a significant increase in the proportion of cells identified as place cells across sessions (ANOVA: $P<0.0001$; $F=37.5825$). The group average is shown in the bold blue trace with shaded SEM, while individual animals are shown with thinner blue traces. (**F**) Average Bayesian position decoding error (cm) from all registered cells and animals across all recorded sessions. The dashed gray line indicates the estimated minimum decoding error from recent computational studies (10 cm) [25]. We observed a significant decrease in decoding error across recorded sessions (ANOVA: $P<0.0001$; $F=7.9845$). (**G**) Population vector field maps show the direction and magnitude of displacement of the maximally correlated population vector in deformed environments relative to the same location in the square environment, while standard deviation of the vector across all cells is indicated in color. The vector field maps reveal the dynamics of remapping across all recorded neurons in each geometrically deformed environment compared to the square.

The representation of a geometrically dynamic environment is reliable across brains

We next asked whether the dynamics we observed in response to environmental deformations reflected a common representational structure that was consistent within and across brains. To do so, we leveraged a Representational Similarity Analysis (RSA) framework. RSA proposes that measurements of a neuronal representation be transformed into a distance matrix that captures the similarity of all pairwise comparisons between conditions in an experiment (Fig. 2A-B) [20, 26]. Such distance matrices, which define the structure of the representation [27], can then be quantitatively compared within and across subjects, and even across assays to uncover if and how representational structure differs. Notably, because distance matrices are agnostic to the measurements from which they were derived, they can be compared across subjects, experimental groups, and theoretical models, which has led to its widespread adoption in human fMRI, EEG, and model comparison research [26-28]. Recent advances in high-yield neuronal recording technology have also enabled its application to study large-scale neuronal dynamics with cellular resolution [20]. Importantly, this approach circumvents challenges to determine correspondence between empirically observed and simulated neuronal activity [26].

To this end, we partitioned each cell's rate map according to the 3 x 3 grid design in each geometric sequence (25 x 25 cm per partition). For each cell, we then calculated the Pearson correlation coefficient of rate maps across all comparisons of sampled partitions within and across geometries [20]. This yielded a large (75^2) and expressive representational similarity matrix (RSM) for comparison across repeated experiences within the same animal, across animals, and against model predictions (Fig. 2B) [27].

To visualize the differences in hippocampal representational structure across geometries, we first averaged our RSMs across all sequences and animals (Fig. 2B). Next, we transformed this averaged RSM into a set of points (one for each partition and geometry) in a two-dimensional space where proximity reflects the similarity between each partition's rate maps using a technique called non-metric Multidimensional Scaling (nMDS) [29]. This allows us to visualize a high-dimensional representation (RSM with 75^2 dimensions) in a digestible lower-dimensional space while preserving the relative similarity between points. This embedding revealed that occluding one partition resulted in nearby partitions becoming more similar to the occluded area, qualitatively mirroring the effects observed in the vectoral field maps (Figs. 1G, 2C). To quantify the effect of deformation in each partition, we measured the correlation of each partition relative to the square environment (Fig. 2D). This revealed a strong effect of partition on similarity,

wherein corners of the environment exhibited the greatest similarity, and the center partition showed the least similarity across geometries (Fig. 2D). Comparing all geometries to the first square of each sequence, we also found a strong effect of geometry on overall similarity, in which the cross- and T-shaped geometries showed the least similarity (i.e., the strongest deformation), while the rectangle and subsequent square geometries showed the greatest similarity (Fig. 2E-F). These results demonstrate that our population-level RSM effectively captures the impact of geometric deformations on hippocampal representational structure.

To compare the reliability of this representational structure within and across brains, we computed the average RSM across cells within each animal for each sequence. Comparison of these RSMs revealed a high degree of similarity in representational structure within animals across sequences (Fig. 2G-H), and across animals within sequence (Fig. 2H). We also found that the similarity of representations across animals significantly increased with experience (Fig. 2H). Next, we asked if we could decode animal identity from the RSM based on reliable individual differences, and whether the RSM could be accurately predicted across animals and sequences. We anticipated that if the RSM is highly reliable across brains, then we should not be able to predict an animal's identity from its RSM. Indeed, while the individual RSMs can be accurately predicted across animals and sequences, we were not able to decode animal identity (Fig. 2I-K). Further analysis of male and female subjects also confirmed there are no sex differences in the CA1 neural representation in our task (Fig. S4). These findings demonstrate that the geometric determinants of the CA1 representation are highly reliable within and across brains.

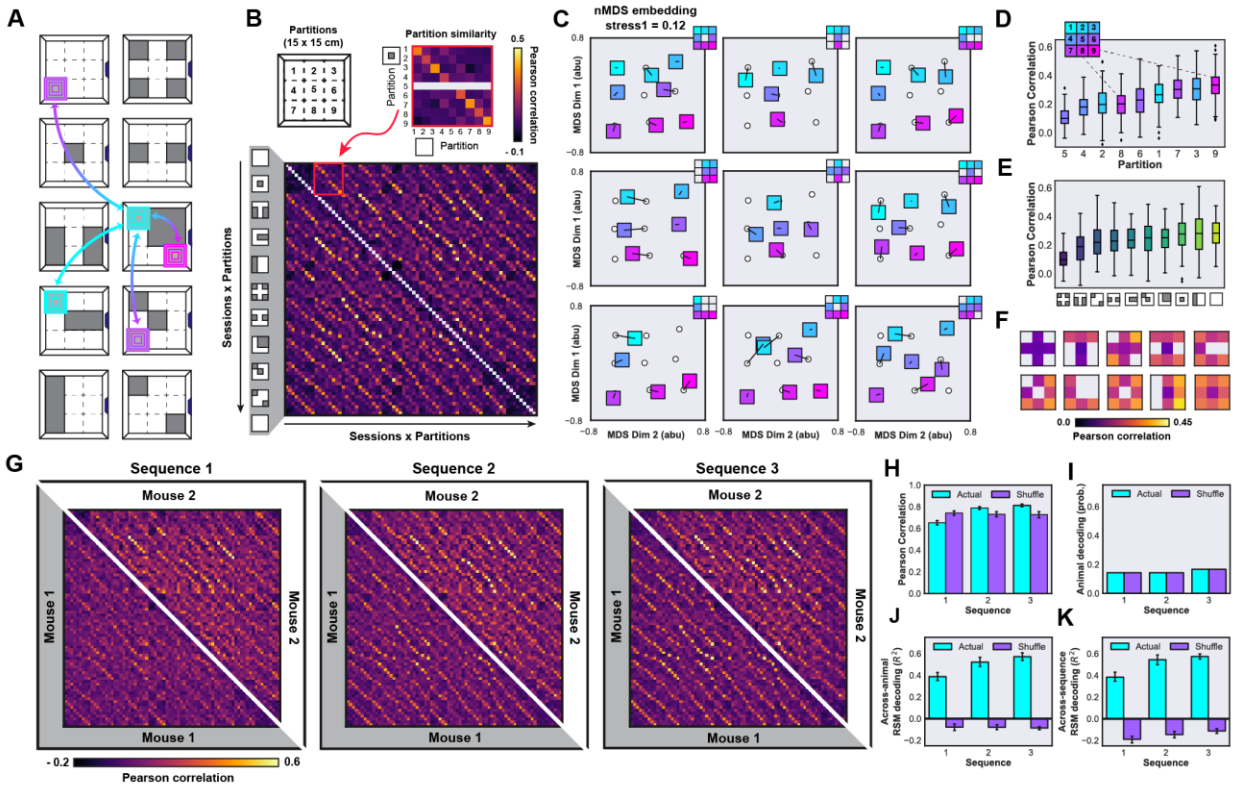

Fig. 2. An RSA-based framework to quantify representational structure in cognitive maps within and across brains. (**A**) The schematic illustrates how partitions within and across geometries are compared in each sequence. (**B**) To construct an RSM for each animal and sequence, we calculated the average Pearson correlation across all pairwise comparisons of partitioned rate maps from matched cells across registered sessions. For each geometry this comparison yields a 9 x 9 matrix of similarities (75^2 comparisons within a sequence for all sampled partitions). (**C**) The nMDS embeddings of the average RSM reveals the effect of geometric deformation on representational similarity among partitions (*stress1* = 0.12). The average result of the square environment at the start and end of each sequence is shown in white, while black lines show translation in the embedding space of the matched partition relative to the square. Colors indicate actual position of each partition across environments (inset shown at the top right of each panel). (**D**) Similarity of individual partitions of all deformed geometries to the matched partition in the square (ANOVA: P(Partition)<0.0001, F(Partition)=102.2067), and (**E**) all partitions in deformed geometries relative to the same partition in the square for each environment (ANOVA: P(Geometry)<0.0001, F(Geometry)=29.3338; P(Geometry X Partition)<0.0001, F(Geometry X Partition)=2.8847). (**F**) Average similarity of individual partitions relative to the matched partition in the square environment across all sequences and animals. (**G**) The RSM from two separate mice shown in upper and lower diagonal across sequences, wherein the white diagonal line separates the triangle of the RSM for each mouse. (**H**) Pearson correlations (mean ± SEM) for all across-animal RSM comparisons within each sequence, compared to correlations of shuffled sequence ordering. We observed a significant effect of sequence on animal-wise representational similarity compared to sequence-shuffled controls (ANOVA: P(Sequence)<0.0001, F(Sequence)=43.4980; P(Shuffle)=0.6045, F(Shuffle)=0.2690; P(Sequence X Shuffle)<0.0001, F(Sequence X Shuffle)=33.6484). (**I**) Decoding probability of animal identity from the respective RSM across sequences compared to a shuffled control. We did not observe a significant effect of shuffling

animal identity on RSM decoding (Chi-square: P=1.000, $\chi 2$=0.000). (**J**) Linear decoding of RSMs across sequences, wherein each animal's RSM is predicted from a left-out sequence (ANOVA: P(Sequence)<0.0001, F(Sequence)=12.9576; P(Shuffle)<0.0001, F(Shuffle)<1477.4415; P(Sequence X Shuffle)<0.001, F(Sequence X Shuffle)=9.8169). (**K**) Decoding of RSMs across animals for all partitions and geometries (ANOVA: P(Sequence)<0.0001, F(Sequence)=16.3418; P(Shuffle)<0.0001, F(Shuffle)=1651.4348; P(Sequence X Shuffle)<0.001, F(Sequence X Shuffle)=11.7739).

Benchmarking theoretical models of cognitive mapping with neuronal population data

Our experimental and analytic approach was designed to provide a highly reliable measure of the representational structure of hippocampal cognitive mapping. Having addressed this need, we explored the possibility that these data can be used to benchmark model predictions of cognitive mapping. Specifically, we sought: (1) quantification of model prediction accuracy against an observed neuronal representation in the brain, (2) discrimination between model predictions in a common task space, and (3) empirical constraints for biologically viable model parameters. While comparison among all models of cognitive mapping is beyond the present scope, we chose to assay several popular neurobiological models of cognitive mapping, including the grid-to-place (GC2PC) [30], boundary-tethered grid-to-place (bt-GC2PC) [15], boundary-vector-to-place (BVC2PC) [31-33], and boundary-vector to successor feature model (BVC2SF) [34, 35]. In addition, we provide a naïve stability model that assumes no geometric effect on mapping dynamics in CA1 (Methods).

According to each view, we simulated predicted neural dynamics in 500 model neurons as an artificial agent navigated our paradigm following the actual trajectories for each animal (Fig. 3A; Methods) [36]. These model neurons formed a basis set from which hippocampal model predictions were derived (Fig. 3B). Treating these model neurons as we had treated recorded populations in CA1, we computed model RSMs for each geometric sequence and animal (Fig. 3B, S5). *Prime facie*, model RSMs varied considerably in this task space, suggesting this approach provides an expressive characterization to adjudicate between predictions. We then measured the rank-order correlation of each model RSM against the recorded CA1 representation, which revealed differences in the ability of each model to account for the CA1 representational structure we observed (Fig. 3C). We found that the BVC2PC model provided the greatest level of prediction accuracy, achieving the theoretical maximum level of prediction (noise ceiling; Methods) (Fig. 3C).

Given that the BVC2PC model provided the highest quality fit to the CA1 representation in one parameterization, we next asked whether model fit could be used to constrain a biologically

plausible parameter space. To this end, we performed a large sweep of the parameters for the boundary vector features forming the hippocampal basis set in the BVC2PC model, including the parameters determining distance-based tuning and the shape of the distribution from which tunings were drawn (Methods). Comparison of different model RSMs to one another revealed that different parameterizations of the BVC2PC model led to disparate predictions (Fig. 3D), demonstrating that our approach is sensitive to model parameterization. We then compared all model parameters against CA1 data and found that only a subset of distance tuning and distribution parameters approached the theoretical maximum level of prediction (Fig. 3E, S6). These results demonstrate the effectiveness of this approach in constraining neurobiologically viable model parameterizations, and in turn allow us to constrain our models using these data prior to application in future experiments.

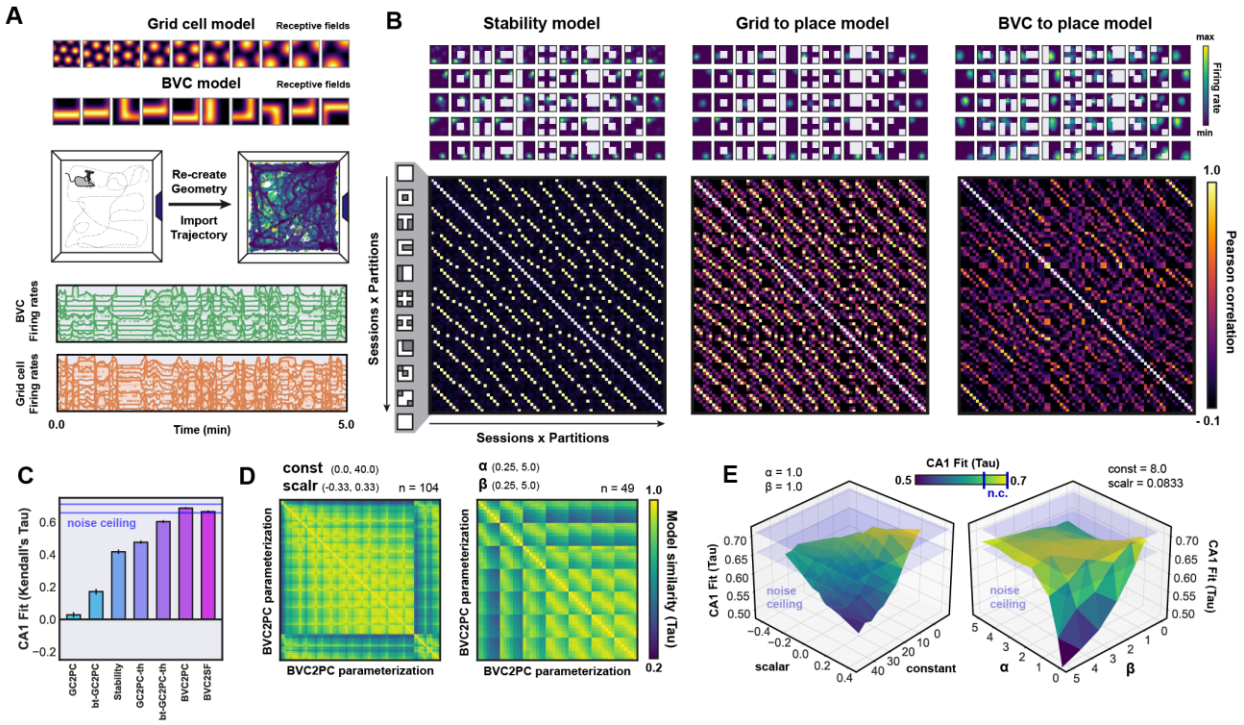

Fig. 3. Benchmarking model predictions against representational structure in CA1. (**A**) Illustration of firing rate simulation for model features in the parahippocampal system, wherein receptive fields are drawn according to each model, and firing rates are simulated with recorded animal trajectories using an open-source toolbox [36]. (**B**) Rate maps were constructed from model firing rates for all features and used for model RSM calculation identical to CA1 data. Example rate maps are shown for the naïve stability, GC2PC with inhibition (GC2PC-th), and BVC2PC models. Following the same process for partitioning and comparing rate maps across geometries, we generated RSMs for each model. (**C**) To evaluate model predictions against the observed representation in CA1, we measured the rank-order correlation between each model RSM and CA1 (Kendall's Tau). Each bar shows the fit of each model to CA1 data, while error bars show the

bootstrap standard deviation of each fit. Blue lines indicate the upper and lower bounds of noise ceiling, which is considered the margin of maximum prediction accuracy given noise in the CA1 data (Methods). GC2PC and bt-GC2PC models were tested both with and without thresholding (GC2PC-th and bt-GC2PC-th) to simulate inhibition, based on the original modeling studies. We observed a significant effect of model on predicting CA1 representation, with BVC2PC and BVC2SF models showing prediction accuracy within noise ceiling (ANOVA: P(Model)<0.0001, F(Model)=47336.1870). All model predictions with the exception of GC2PC (without inhibition) significantly fit CA1 representation (P<0.0001). (**D**) The model-wise RSM shows the measured similarity (Kendall's Tau) of grid-search parameterizations of the BVC distance constant and scalar (left; N=104), and BVC alpha and beta parameters that determined the BVC distance distribution (right; N=49). (**E**) The surface plots illustrate the CA1 prediction accuracy landscape of the BVC2PC model with the same parameters shown in (**D**), both for the BVC distance constant and scalar (left), as well as the BVC distance distribution (right). The figures illustrate that a specific range of BVC2PC parameters enters the upper and lower bounds of noise ceiling, revealing a constrained, biologically viable parameter space.

Conclusion

Here we present the first direct quantitative comparison of population-level neuronal representation in freely behaving animals against theoretical models of cognitive mapping. Leveraging a similarity-based framework, we observed that a constrained range of allocentric, BVC-based models accurately predict CA1 spatial representation (within maximum theoretical limits). The success of such models to predict CA1 representation suggest that an allocentric vector-based code strongly determines population-level neuronal dynamics in CA1 [10, 37, 38]. These are the first results to our knowledge demonstrating that large-scale predictions from neurobiological models of cognitive mapping can be directly evaluated against population dynamics in freely behaving animals. The present dataset and framework provides a benchmark for theoretical innovations in cognitive mapping research and, more generally, establishes a novel approach to compare representational structure across brain regions, assays, species, and theoretical models.

By comparing model predictions in this paradigm, we were able to accurately discriminate between population-level representations with distinct (across-model) and similar bases (within-model parameterizations), suggesting the same approach can be applied to determine the relationship between neural codes across brain regions. For example, the representation we observed in CA1 could be quantitatively compared to populations in the entorhinal, subicular, retrosplenial, and further association cortices following a similar approach [39]. This framework holds the promise to allow us to quantitatively derive the transformation of representational structure across brain-wide networks. When combined with causal circuit manipulations, these

experiments can further implicate particular mechanisms thought to be selectively responsible for instantiating specific aspects of this representational structure.

A growing emphasis of recent work in systems neuroscience is to determine how behavioral states and task objectives shape neuronal representation. While the hippocampus has been long appreciated for its role in spatial coding, navigation, and memory, an increasing number of studies report hippocampal coding of non-spatial, behaviorally relevant task features, such as rewards, tones, and evidence accumulation [2, 3, 40, 41]. We suggest that future studies build on this approach to compare neuronal codes across regions and against model predictions under distinct behavioral states and task objectives. Developments upon the present dataset and framework will offer a novel testbed for such models within and beyond the hippocampus.

References

1. O'Keefe, J. and L. Nadel, *The Hippocampus as a Cognitive Map*. 1978, Oxford, U.K.: Oxford University Press.
2. Nieh, E.H., et al., *Geometry of abstract learned knowledge in the hippocampus.* Nature, 2021. **595**(7865): p. 80-84.
3. Lee, J.S., et al., *The Statistical Structure of the Hippocampal Code for Space as a Function of Time, Context, and Value.* Cell, 2020. **183**(3): p. 620-635 e22.
4. Astur, R.S., et al., *Humans with hippocampus damage display severe spatial memory impairments in a virtual Morris water task.* Behavioural Brain Research, 2002. **132**(1): p. 77-84.
5. Sutherland, R.J., B. Kolb, and I.Q. Whishaw, *Spatial mapping: definitive disruption by hippocampal or medial frontal cortical damage in the rat.* Neuroscience Letters, 1982. **31**(3): p. 271-276.
6. Morris, R.G.M., et al., *Place navigation impaired in rats with hippocampal lesions.* Nature, 1982. **297**(5868): p. 681-683.
7. Scoville, W.B. and B. Milner, *LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS.* Journal of Neurology, Neurosurgery & Psychiatry, 1957. **20**(1): p. 11-21.
8. Whittington, J.C.R., et al., *How to build a cognitive map.* Nat Neurosci, 2022. **25**(10): p. 1257-1272.
9. Schrimpf, M., et al., *Integrative Benchmarking to Advance Neurally Mechanistic Models of Human Intelligence.* Neuron, 2020. **108**(3): p. 413-423.
10. O'Keefe, J. and N. Burgess, *Geometric determinants of the place fields of hippocampal neurons.* Nature, 1996. **381**(6581): p. 425-428.
11. Keinath, A.T., et al., *Environmental deformations dynamically shift human spatial memory.* Hippocampus, 2021. **31**(1): p. 89-101.
12. Keinath, A.T., et al., *Environmental Geometry Aligns the Hippocampal Map during Spatial Reorientation.* Current Biology, 2017. **27**(3): p. 309-317.
13. Leutgeb, S., et al., *Independent Codes for Spatial and Episodic Memory in Hippocampal Neuronal Ensembles.* Science, 2005. **309**(5734): p. 619-623.

14. Muller, R. and J. Kubie, *The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells.* The Journal of Neuroscience, 1987. **7**(7): p. 1951-1968.
15. Keinath, A.T., R.A. Epstein, and V. Balasubramanian, *Environmental deformations dynamically shift the grid cell spatial metric.* Elife, 2018. **7**.
16. Krupic, J., et al., *Grid cell symmetry is shaped by environmental geometry.* Nature, 2015. **518**(7538): p. 232-235.
17. Lever, C., et al., *Long-term plasticity in hippocampal place-cell representation of environmental geometry.* Nature, 2002. **416**(6876): p. 90-94.
18. Wills, T.J., et al., *Attractor Dynamics in the Hippocampal Representation of the Local Environment.* Science, 2005. **308**(5723): p. 873-876.
19. Leutgeb, J.K., et al., *Pattern Separation in the Dentate Gyrus and CA3 of the Hippocampus.* Science, 2007. **315**(5814): p. 961-966.
20. Keinath, A.T., C.A. Mosser, and M.P. Brandon, *The representation of context in mouse hippocampus is preserved despite neural drift.* Nat Commun, 2022. **13**(1): p. 2415.
21. Ghosh, K.K., et al., *Miniaturized integration of a fluorescence microscope.* Nat Methods, 2011. **8**(10): p. 871-8.
22. Aharoni, D., et al., *All the light that we can see: a new era in miniaturized microscopy.* Nature Methods, 2019. **16**(1): p. 11-13.
23. Zhang, K., et al., *Interpreting Neuronal Population Activity by Reconstruction: Unified Framework With Application to Hippocampal Place Cells.* Journal of Neurophysiology, 1998. **79**(2): p. 1017-1044.
24. Tanni, S., W. de Cothi, and C. Barry, *State transitions in the statistically stable place cell population correspond to rate of perceptual change.* Curr Biol, 2022. **32**(16): p. 3505-3514 e7.
25. Hazon, O., et al., *Noise correlations in neural ensemble activity limit the accuracy of hippocampal spatial representations.* Nat Commun, 2022. **13**(1): p. 4276.
26. Kriegeskorte, N., M. Mur, and P. Bandettini, *Representational similarity analysis - connecting the branches of systems neuroscience.* Front Syst Neurosci, 2008. **2**: p. 4.
27. Kriegeskorte, N. and R.A. Kievit, *Representational geometry: integrating cognition, computation, and the brain.* Trends in Cognitive Sciences, 2013. **17**(8): p. 401-412.
28. Nili, H., et al., *A Toolbox for Representational Similarity Analysis.* PLoS Computational Biology, 2014. **10**(4): p. e1003553.
29. Kruskal, J.B., *Nonmetric multidimensional scaling: A numerical method.* Psychometrika, 1964. **29**(2): p. 115-129.
30. Solstad, T., E.I. Moser, and G.T. Einevoll, *From grid cells to place cells: a mathematical model.* Hippocampus, 2006. **16**(12): p. 1026-31.
31. Hartley, T., et al., *Modeling place fields in terms of the cortical inputs to the hippocampus.* Hippocampus, 2000. **10**(4): p. 369-79.
32. Barry, C., et al., *The Boundary Vector Cell Model of Place Cell Firing and Spatial Memory.* Reviews in the Neurosciences, 2006. **17**(1-2).
33. Grieves, R.M., É. Duvelle, and P.A. Dudchenko, *A boundary vector cell model of place field repetition.* Spatial Cognition & Computation, 2018. **18**(3): p. 217-256.
34. de Cothi, W. and C. Barry, *Neurobiological successor features for spatial navigation.* Hippocampus, 2020. **30**(12): p. 1347-1355.
35. Stachenfeld, K.L., M.M. Botvinick, and S.J. Gershman, *The hippocampus as a predictive map.* Nat Neurosci, 2017. **20**(11): p. 1643-1653.
36. George, T.M., et al., *RatInABox: A toolkit for modelling locomotion and neuronal activity in continuous environments*. 2022, Cold Spring Harbor Laboratory.

37. Ormond, J. and J. O'Keefe, *Hippocampal place cells have goal-oriented vector fields during navigation.* Nature, 2022. **607**(7920): p. 741-746.
38. Nagelhus, A., et al., *Object-centered population coding in CA1 of the hippocampus.* Neuron, 2023. **111**(13): p. 2091-2104.e14.
39. Maisson, D.J., et al., *Making Sense of the Multiplicity and Dynamics of Navigational Codes in the Brain.* J Neurosci, 2022. **42**(45): p. 8450-8459.
40. Gauthier, J.L. and D.W. Tank, *A Dedicated Population for Reward Coding in the Hippocampus.* Neuron, 2018. **99**(1): p. 179-193 e7.
41. Aronov, D., R. Nevers, and D.W. Tank, *Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit.* Nature, 2017. **543**(7647): p. 719-722.

Acknowledgments: We would like to thank Tianmeng Xu, Raphael Lavoie, and Kay Cui for their technical assistance with this project.

Funding:

Canadian Institutes of Health Research grant #367017, #377074 (MPB)

Natural Sciences and Engineering Research Council of Canada grant #74105 (MPB)

Canadian Institutes of Health Research Postdoctoral Fellowship (JQL)

Fonds de Recherche du Québec Postdoctoral Fellowship (JQL)

Natural Sciences and Engineering Research Council of Canada Banting Fellowship (ATK)

Canadian Institutes of Health Research Postdoctoral Fellowship (ATK)

Author contributions:

Conceptualization: JQL, ATK, MPB

Methodology: JQL, ATK

Investigation: JQL, ATK, EC

Visualization: JQL, ATK

Funding acquisition: JQL, ATK, MPB

Project administration: JQL, ATK, MPB

Supervision: MPB

Writing – original draft: JQL, ATK

Writing – review & editing: JQL, ATK, MPB

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: All data, code, and materials information will be made freely available upon publication.

Materials and Methods

Subjects

Naive male (4) and female (3) mice (C57Bl/6, Charles River) were housed in pairs on a 12-hour light/dark cycle at 22°C and 40% humidity with food and water ad libitum. All experiments were carried out during the light portion of the light/dark cycle, and in accordance with McGill University and Douglas Hospital Research Centre Animal Use and Care Committee (protocol #20157725) and with Canadian Institutes of Health Research guidelines.

Surgeries

During all surgeries, mice were anesthetized via inhalation of a combination of oxygen and 4% Isoflurane before being transferred to the stereotaxic frame (David Kopf Instruments), where anesthesia was maintained via inhalation of oxygen and 0.5-2.5% Isoflurane for the duration of the surgery. Body temperature was maintained with a heating pad and eyes were hydrated with gel (Optixcare). Carprofen (10 ml kg^{-1}) and saline (0.5 ml) were administered subcutaneously at the beginning of each surgery. Preparation for recordings involved three surgeries per mouse.

First, at the age of six to ten weeks, each mouse was transfected with a 500 nl injection of the calcium reporter GCaMP6f via the viral construct AAV9.syn.GCaMP6f.WPRE.SV40. The original titre of the AAV9.syn.GCaMP6f.WPRE.SV40 construct, sourced from University of Pennsylvania Vector Core, was 3.26e14 GC-ml and was diluted in sterile PBS at a 1:30 ratio before surgical microinjection.

Between five to six weeks post-injection, a 1.8mm diameter gradient refractive index (GRIN) lens (Edmund Optics) was implanted above dorsal CA1 (Referenced to bregma: +/-1.45 mm ML, -1.95 mm AP; -1.37 mm DV from skull surface). Implantation required aspiration of intervening cortical tissue. In addition to the GRIN lens, three stainless steel screws were threaded into the frontal, occipital and parietal bone (above the contralateral hippocampus) to stabilize the implant. Dental cement (C&B Metabond) was applied to secure the GRIN lens and anchor screws to the skull. A silicone adhesive (Kwik-Sil, World Precision Instruments) was applied to protect the top surface of the GRIN lens until the next surgery.

Two weeks after lens implantation, an aluminum baseplate was affixed via dental cement (C&B Metabond) to the skull of the mouse, which would later secure the miniaturized fluorescent endoscope (miniscope) in place during recording. The miniscope/baseplate was mounted to a stereotaxic arm for lowering above the implanted GRIN lens until the field of view contained visible cell segments and dental cement was applied to affix the baseplate to the skull. A polyoxymethylene cap with a metal nut weighing ~3 g was affixed to the baseplate when the mice

were not being recorded, to protect the baseplate and lens, as well as to simulate the weight of the miniscope.

After surgery, animals were continuously monitored until they recovered. For the initial three days after surgery mice were provided with a soft diet supplemented with Carprofen for pain management (MediGel CPF, ~5mg kg^{-1} each day). Familiarization with the open square environment began at least 1 week following the baseplate surgery.

Data acquisition

In vivo calcium videos were recorded with a UCLA miniscope (v3; miniscope.org) containing a monochrome CMOS imaging sensor (MT9V032C12STM, ON Semiconductor) connected to a custom data acquisition (DAQ) box (miniscope.org) with a lightweight, flexible coaxial cable. The DAQ was connected to a PC with a USB 3.0 SuperSpeed cable and controlled with Miniscope custom acquisition software (miniscope.org). The outgoing excitation LED was set to between 2-8% (~0.05-0.2 mW), depending on the mouse to maximize signal quality with the minimum possible excitation light to mitigate the risk of photobleaching. Gain was adjusted to match the dynamic range of the recorded video to the fluctuations of the calcium signal for each recording to avoid saturation. Behavioral video data were recorded by a webcam mounted above the environment. The DAQ simultaneously acquired behavioral and cellular imaging streams at 30 Hz as uncompressed AVI files and all recorded frames were timestamped for post-hoc alignment.

All recording environments were constructed of white Lego base plates (Lego, Inc) and white painted plywood walls. The full square environment was 75 cm x 75 cm. The external walls were 60 cm tall, while inserted partition walls were 25 cm tall. A blue Lego baseplate was affixed at the top-center of the Eastern wall to serve as an orienting visual feature. During recording, the environment was dimly lit. All sessions were 40 min, and one session was recorded per day to avoid photobleaching. The order of the environmental configurations was randomized for each mouse, but repeated across sequences such that repeated configurations were always equidistant in time.

Data preprocessing

Calcium imaging data were preprocessed prior to analyses via a pipeline of open-source MATLAB (MathWorks; version R2015a) functions to correct for motion artifacts (*42*), segment cells and extract transients (*43, 44*). To extract the rising phase of transients from each filtered calcium trace, we proceeded as follows. First, we computed the derivative of the calcium signal, smoothed with a gaussian kernel with a standard deviation of 5 frames. Next, because calcium transients around the baseline can only be positive, we estimated the variance in the derivative of the smoothed calcium signal on the basis of noise via a half-normal distribution such that:

$$NOISE = \frac{std(\Delta t(t < 0))}{\sqrt{1 - \frac{2}{\pi}}}$$

Where Δt is the smoothed time-derivative of the median-subtracted calcium trace t. We then z-scored Δt on the basis of this noise distribution. The final binarized rising-phase vector was then set to 1 whenever this z-scored Δt vector exceeded 2.5, and 0 otherwise. This binary vector was treated as the *firing rate* in all further analyses. The motion-corrected calcium imaging data were manually inspected to ensure that motion correction was effective and did not introduce additional artifacts. Following this preprocessing pipeline, the spatial footprints of all cells were manually verified to remove lens artifacts. Position data were inferred from the onboard miniscope red LED offline following recording using a custom written MATLAB (MathWorks) script and were manually corrected if needed. Cells were tracked across sessions on the basis of prominent brain surface landmarks, their spatial footprints, and/or centroids.

Data analysis

All analyses were conducted using the binary vector of the rising phases of transients, treating this vector as if it were the firing rate of the cell (henceforth *firing rate*). Similar results are observed when the likelihood of spiking was inferred via a second-order autoregressive deconvolution instead of transient rising-phase extraction. All statistical comparisons were performed with Scipy and Statsmodels libraries in Python 3.8.

Place cell identification

To detect place cells, we generated rate maps for each session half (first and second 20 min), and compared the resulting Pearson correlation to the maps generated from 1000 circular shuffles or

the corresponding position data. Place cells were identified as those whose split-half rate map correlation exceeded the 99th percentile of the corresponding shuffled distribution for each cell.

Bayesian decoding of animal position

To decode animals' position from calcium traces within each session, we performed a 5-fold split of spatially binned position and trace data and transformed the binned positions to a one-hot vector. Using the Gaussian Naïve Bayes method in the scikit-learn Python library, we predicted position on withheld data from maximum likelihood estimation data following:

$$P(x_i \mid y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} exp\left(-\frac{\left(x_i - \mu_y\right)^2}{2\sigma_y^2}\right)$$

where x_i corresponds to the predicted position at i and y to the respective calcium traces. We assumed a flat prior (equal likelihood at all positions) and the scikit-learn default σ value of 1e-9. Decoding error was then estimated as the Euclidean distance between the predicted and actual spatial bin position of the animal on withheld data.

Vector field map generation

To visualize the directional flow of spatial mapping in CA1 across environments, we calculated the rate map correlation across all cells (i.e., the population vector) for each spatial bin in all geometries relative to the square environment. We then measured the vector between each spatial bin in the square environment and the maximally correlated bin in each geometry, averaged across animals and sequences. As a measure of variability, we calculated the standard deviation of the resulting vector for each spatial bin across animals and sequences.

Representational similarity matrix (RSM) generation

Rate maps were first constructed by spatially binning position data into pixels corresponding to a 5cm x 5cm grid of locations and smoothed with a 5cm standard deviation isometric Gaussian kernel. We then divided rate maps into 9 partitions according to the grid-design of the environment and calculated the pairwise Pearson correlation between all rate maps registered across all pairs of environments and computed the average rate map correlation across all registered cells in each geometry and partition. The resulting RSM was then organized to match the order of geometries and partitions across animals based on the sequence presented to the first subject and in ascending order of partitions from left to right (West to East), and top to bottom (North to South), including

each geometry from a respective sequence, starting and ending with the square (Sequence 1: Session 1 – 11; Sequence 2: Session 11 – 21; Sequence 3: Session 21 – 31).

Non-metric multi-dimensional scaling (nMDS)

Embeddings of the population representation via nMDS were computed as follows: First, we computed the mean rate map correlation across cells for each pairwise comparison of partitions, as described in the main text. Next, we transformed this correlation matrix into a distance matrix by computing one minus this correlation matrix, with the diagonal set to 0 distance. Finally, this distance matrix was reduced to a two-dimensional embedding via nonmetric multidimensional scaling, implemented by the MDS function in the scikit-learn Python library with conservative parameterizations (*eps*=1e-12, *max_iter*=1e9) and minimizing Kruskal's normalized stress1 cost function:

$$stress1 = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(F\left(d_{ij}\right) - d_{ij}\right)^2}{\sum_{i=1}^{n}\sum_{j=i+1}^{n} {d_{ij}}^2}}$$

where d_{ij} is the distance between i and j in the measured distance matrix and $F\left(d_{ij}\right)$ is a nonparametric monotonic function of the measured distances fit via isotonic regression.

Noise ceiling calculation

To estimate the suggested theoretical maximum level of model prediction we calculated the upper and lower bounds of the noise ceiling across animals and sequences. To calculate the lower bound, we excluded one RSM at a time from the averaged RSM calculated across animals and sequences, and then computed the rank-order correlation between the excluded RSM and the remaining RSMs averaged across animals and sequences. To calculate the noise ceiling upper bound, we calculated the average rank-order correlation of each RSM to the average of all RSMs across animals and sequences. This range of values between the upper and lower bound thus indicate the average correlation of an unseen animal with the group representation.

Modelling of spatial cell types

To model spatial features in the parahippocampal system according to each theoretical view, we utilized a recently developed open-source toolbox (*36*) and custom functions in Python 3.8. Parameterizations for each model were based on those defined in previous work (*30* – *35*). In

addition, for each simulated cell type we drew firing rates from the modelled spatial receptive fields using animal trajectories from each recorded session recorded at 30 Hz.

Boundary-vector cells (BVC)

We generated a population of BVCs according to same methods used in prior studies (*31 – 34*). As described previously (*34*), the firing rate of the ith BVC, with preferred distance d_i and angle φ_i to a boundary at distance r and direction θ subtending at angle $\delta\theta$ is given by:

$$\delta f_i = g_i(r,\theta)\delta\theta$$

where,

$$g_i(r,\theta) \propto \frac{exp[-(r-d_i)^2/2\sigma_{rad}^2(d_i)]}{\sqrt{2\pi\sigma_{rad}^2(d_i)}} \times \frac{exp\left[-(\theta-\varphi_i)^2/2\sigma_{ang}^2\right]}{\sqrt{2\pi\sigma_{ang}^2}}$$

in which the angular tuning width σ_{ang} is constant and the radial tuning increases with the preferred distance tuning $\sigma_{rad}(d_i) = \frac{d_i}{\beta + \xi}$ for constants β and ξ. To define the tuning distances of BVCs, we randomly drew values from 0 and 75 cm from a β distribution with parameters α and β between 0.25 and 5.0, wherein α and β both equal to 1.0 generate a random uniform distribution of tuning distances.

Boundary-vector to place cells (BVC2PC)

According to the boundary-vector to place model (*31 – 33*), we calculated the firing of a model place cell from the input from at least 2 and up to 16 model BVCs. Specifically, we calculated the geometric mean of randomly selected BVC inputs as in (*22*):

$$F(x) = \Theta\left(u\left(\left(\prod_{i=1}^{n} f_i(x)/max_x\right)^{\frac{1}{n}}\right) - T\right)$$

where T is the cell's threshold, defined as 80 percent of the maximum within-session summed

inputs, and

$$\Theta(\mathrm{x}) = \left\{ \begin{matrix} x & if\ x > 0 \\ 0 & else \end{matrix} \right\}$$

Boundary-vector to successor features (BVC2SF)

As described previously (*34*), we generated BVC2SFs from a population of 500 model BVCs randomly drawn for each animal and learned predictive weights among BVC firing vectors observed through time, according to the temporal difference learning rule:

$$\widetilde{M} \leftarrow \widetilde{M} + a_{\widetilde{M}}[f(s_t) + \gamma\tilde{\psi}(s_{t+1}) - \tilde{\psi}(s_t)]f(s_t)^{\mathrm{T}}$$

where $a_{\widetilde{M}}$ is the learning rate for the successor representation $\widetilde{M}$ weight vector, and $\tilde{\psi}(s)$ is the expected sum of future BVC population firing rate vectors, discounted exponentially into the future with the parameter $\gamma \in [0, 1]$. Based on a grid search for approximately optimal parameterizations with a uniform BVC tuning distribution, we set $a_{\widetilde{M}} = 0.0167$ and $\gamma = 0.999$. The firing rate of each simulated feature F_i in a location s was proportional to the threshold, weighted sum of the BVCs connected to it:

$$F_i(s) \propto \Theta\left(\sum\nolimits_j \widetilde{M}(i,j)f_j(s) - T\right)$$

where *T* is the cell's threshold, and

$$\Theta(\mathrm{x}) = \left\{ \begin{matrix} x & if\ x > 0 \\ 0 & else \end{matrix} \right\}$$

Grid cells

We modelled the firing of populations of grid cells as the summation of three shifted cosines, as described previously (*30*). Briefly, grid cell functions were constructed from the sum of three two-dimensional sinusoidal gratings with 60- and 120-degree angular differences, and firing rates

between 0 and g_w^{max}, given by:

$$g_w(x,y) = g_w^{max}\frac{2}{3}\left(\frac{1}{3}\sum_{i=1}^{3}\cos\big(k_i(r-r_0)\big)+\frac{1}{2}\right)$$

where r defines the spatial phase, and k_i defines the wave vectors at i.

Grid to place cell (GC2PC model)

As described previously, we simulated individual place cells from the input of 50 randomly generated grid cells with a single phase, and logarithmic grid-spacing between 28 cm to 73 cm and uniformly random grid orientations. Grid inputs to each place cell were then defined according to the weight coefficient:

$$A_{w0} = \frac{\exp\left(-\frac{4}{3}\pi^2\sigma^2/\lambda^2\right)}{\lambda^2}$$

where λ is the grid spacing of the cell's respective inputs, and the constant σ = 12 cm, as in (*30*). While the primary results from the prior study did not explicitly define the extent inhibition is to be applied to the grid cell inputs, we modelled local inhibition with the summed a threshold T, defined as 80% of the maximum within-session firing rates of grid inputs. Inhibition was then applied to the same GC2PC prediction (GC2PC-th), and place cell firing was estimated as:

$$F_i(s) \propto \Theta\left(\sum_{n=1}^{N} A_w^n g_w^n(s) - T\right)$$

where,

$$\Theta(\mathrm{x}) = \left\{\begin{matrix} x & if\ x > 0 \\ 0 & else \end{matrix}\right\}$$

For model features without inhibition, the same procedure was applied without the subtraction of the threshold T.

Boundary-tethered grid to place cells (bt-GC2PC)

To simulate a grid to place cell model with dynamic boundary-tethering (*15*), we adopted a

parameter-free approach to boundary tethering of the same GC2PC features described above, as in (*30*). To this end, we first generated boundary-tethered receptive fields for the North, South, East, and West boundaries by translating the GC2PC fields generated in the square environment by their relative distance to the walls of each geometry. Following contact with the respective wall (animal within 6 cm of boundary), we sampled from the respective boundary-tethered receptive field. bt-GC2PC features were calculated as the average of the sampled boundary-tethered maps, and inhibition was applied identically as described above (bt-GC2PC-th).

References

42. E. A. Pnevmatikakis, A. Giovannucci, NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *Journal of Neuroscience Methods* **291**, 83-94 (2017).
43. Eftychios *et al.*, Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron* **89**, 285-299 (2016).
44. P. Zhou *et al.*, Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *eLife* **7**, (2018).
45. L. Sheintuch *et al.*, Tracking the Same Neurons across Multiple Days in Ca2+ Imaging Data. *Cell Reports* **21**, 1102-1115 (2017).
46. T. Nath, A. Mathis, A.C. Chen, *et al.* Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nat Protoc* **14**, 2152–2176 (2019). https://doi.org/10.1038/s41596-019-0176-0

Supplementary Figures

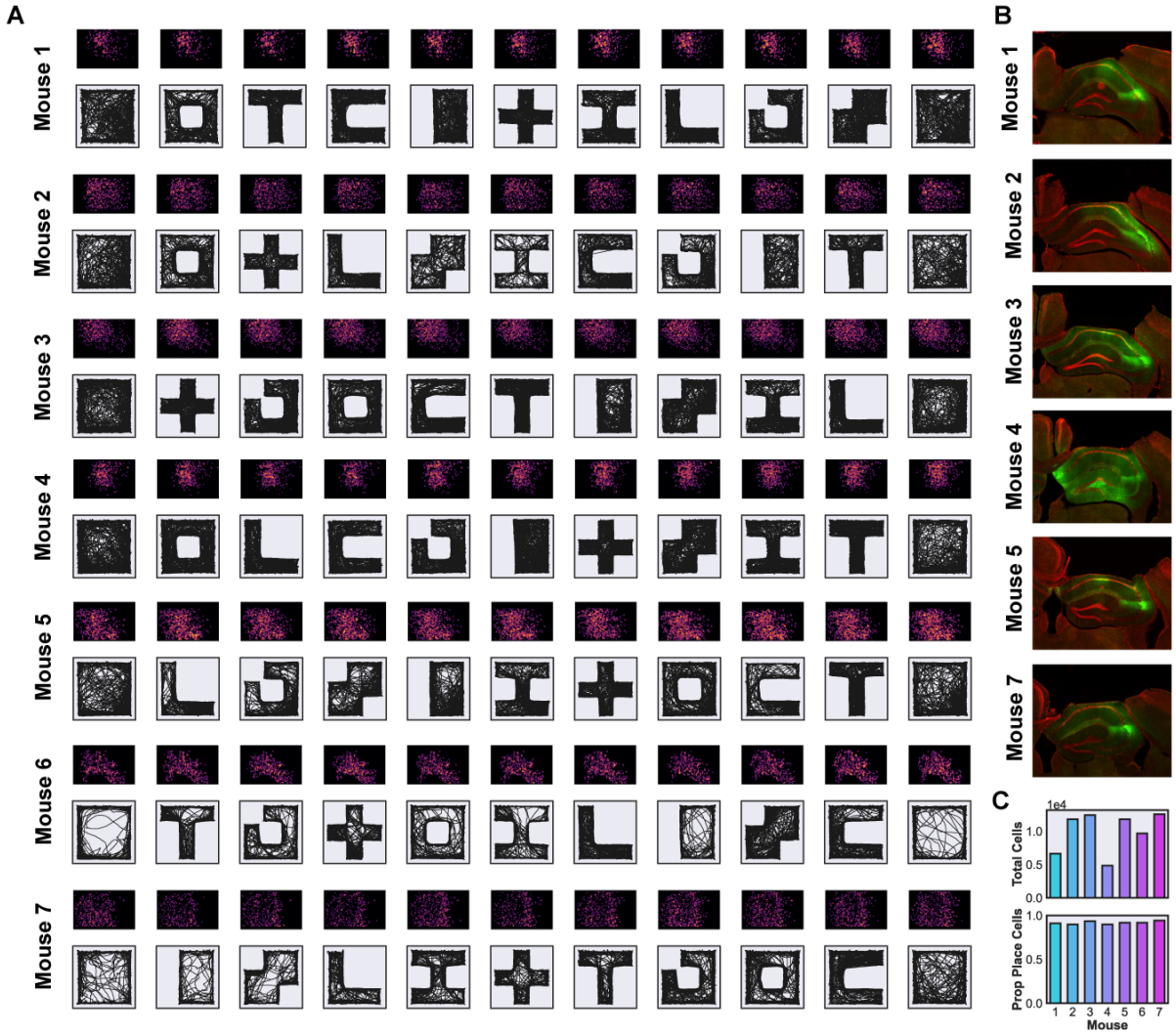

Fig. S1. Recording large neuronal populations with miniscope calcium imaging in freely behaving mice. (**A**) Examples of spatial footprint (SFP) max-projections across a single geometric sequence in each animal following source extraction and across session image alignment (top), along with the respective animal's estimated position using Deeplabcut (head center; bottom) [46]. (**B**) Histological confirmation of GRIN lens implantation above dorsal CA1 (DAPI in red, GCaMP6f in green). Tissue damage following post-mortem lens extraction prevented histological confirmation of CA1 implant location in Mouse 6, but subsequent analyses confirmed this mouse was similar to the remaining subjects in the group (see below). (**C**) Top panel shows the total number of cells detected across all sessions in each animal (non-unique cells prior to across-session registration), and the bottom panel shows that the detected proportion of place cells did not differ across mice, confirming similar quality of spatial coding across subjects in CA1.

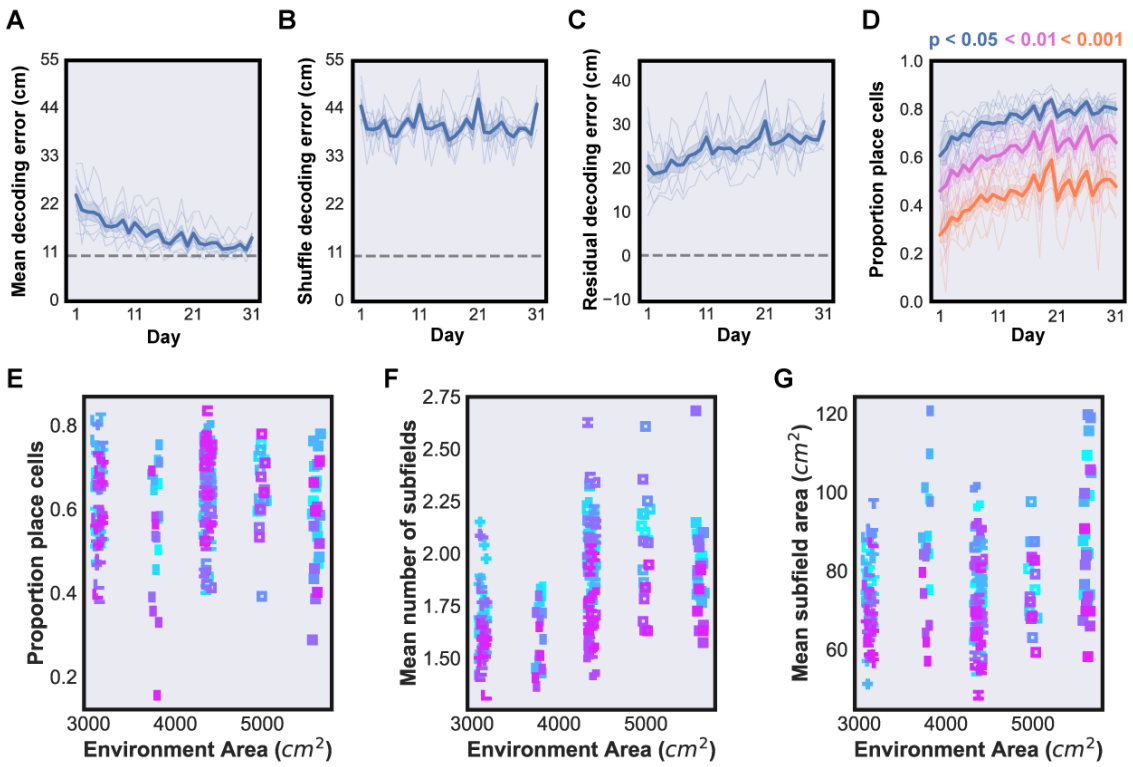

Fig. S2. Precise spatial coding of a geometrically dynamic environment in CA1 across protracted experience. (**A**) To confirm previously observed decreases in decoding error across days (Fig. 1F), we down-sampled the number of cells to match the minimum in each animal registered across all sessions, and predicted each animal's position with a Bayesian decoder (Methods). Complementing our results with the inclusion of all cells, we found a similar decrease in decoding error across sessions (ANOVA: P(Day)<0.0001, F(Day)=4.2184). The gray line indicates anticipated theoretical maximum level of decoding accuracy (10 cm) based on previous computational work [25]. (**B**) Circular shuffling procedures illustrate chance levels of decoding accuracy across all sessions. (**C**) To calculate the residual decoding error, we subtracted the mean decoding error from circularly shuffled data for each animal, and found the residual decoding error improved across sessions, and was significantly above chance levels (ANOVA: P(Day)=0.0016, F(Day)=2.1055, One-sample t-test against chance level: P<0.0001, t=59.5178). (**E-G**) Following detection of place cells across all days, we compared the proportion of place cells, mean number of detected subfields (Methods), and subfield area to varied environment area across all geometries and animals. Each datum shows the recorded geometry in the dot shape, and the animal in dot color. We found no effect of environment area on the proportion of detected place cells (Pearson correlation: P=0.4289, R=0.0640), but did observe a significant effect on the number of detected subfields (P<0.0001, R=0.4538), and subfield area for detected place cells (P<0.0001, R=0.3358).

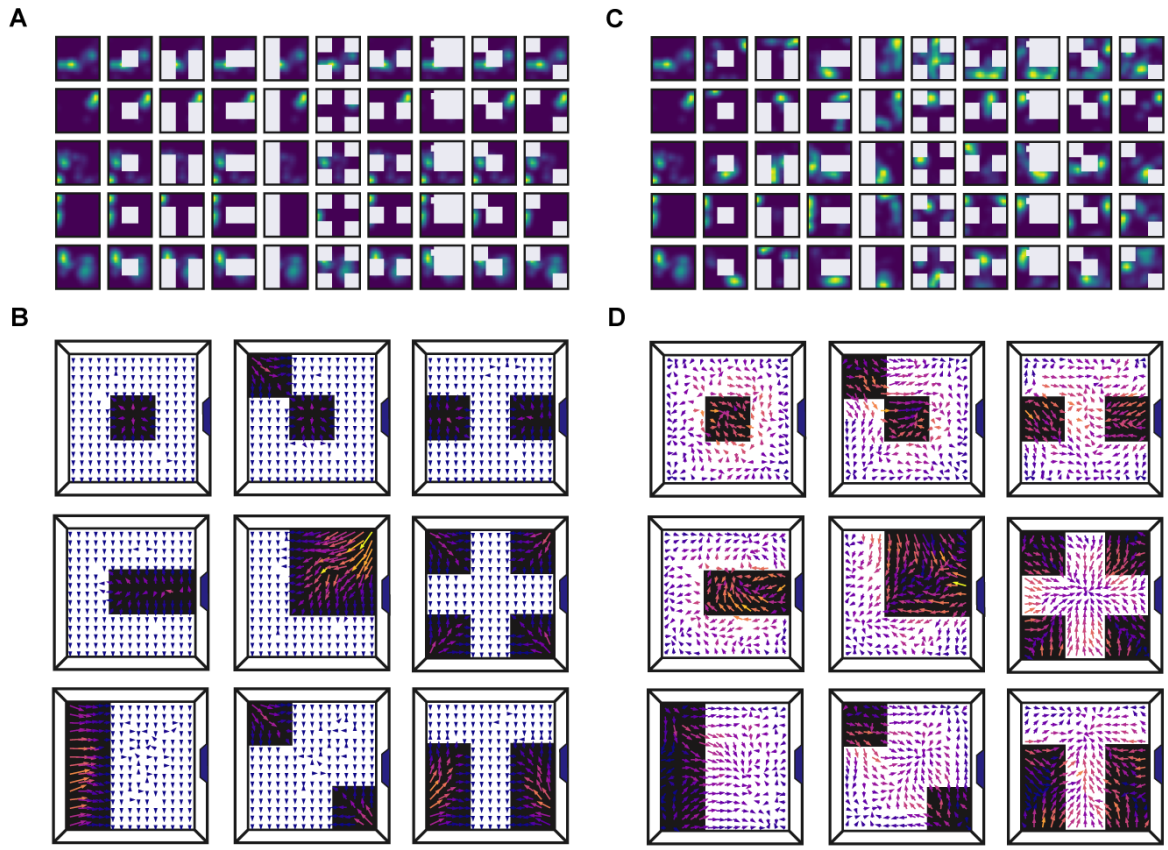

Fig. S3. Geometric remapping with simulated spatial stability and environment partitioning. (**A-B**) To determine whether similar vector fields would be observed due to partitioning of environments with allocentrically static spatial maps (Methods), we simulated stable maps using rate maps from cells active in the first square environment, and masked regions sampled by the animal in each session. We then calculated vector fields from simulated maps, which revealed anticipated movement near boundaries, but lack of vector fields in sampled partitions. (**C**) The original rate maps as those shown with stability simulations in (**A**). (**D**) To confirm the effect of geometric remapping, we calculated the residual vector fields by subtracting those calculated from stable field simulation with partitioning from vector fields calculated using actual data. With this approach, we continued to observe a strong effect of geometry on the residual vector fields, confirming the effect of geometric remapping observed in CA1.

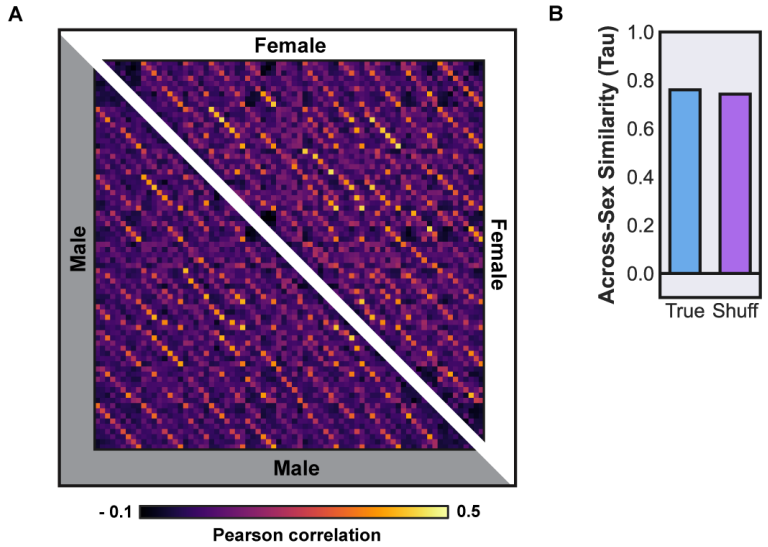

Fig. S4. Sex does not determine the representation of geometrically dynamic environments in CA1. (**A**) The RSM shows the averaged representation from male and female mice across all sequences in CA1. (**B**) To determine whether sex affects CA1 representation in this task, we calculated the rank-order correlation (Kendall's Tau) of the aggregate male and female RSM, compared to the average of 100 shuffles of sex. We found no difference between male and female RSM correlation compared to the shuffled controls (One-sample t-test: P=0.9999, t=-5.6237).

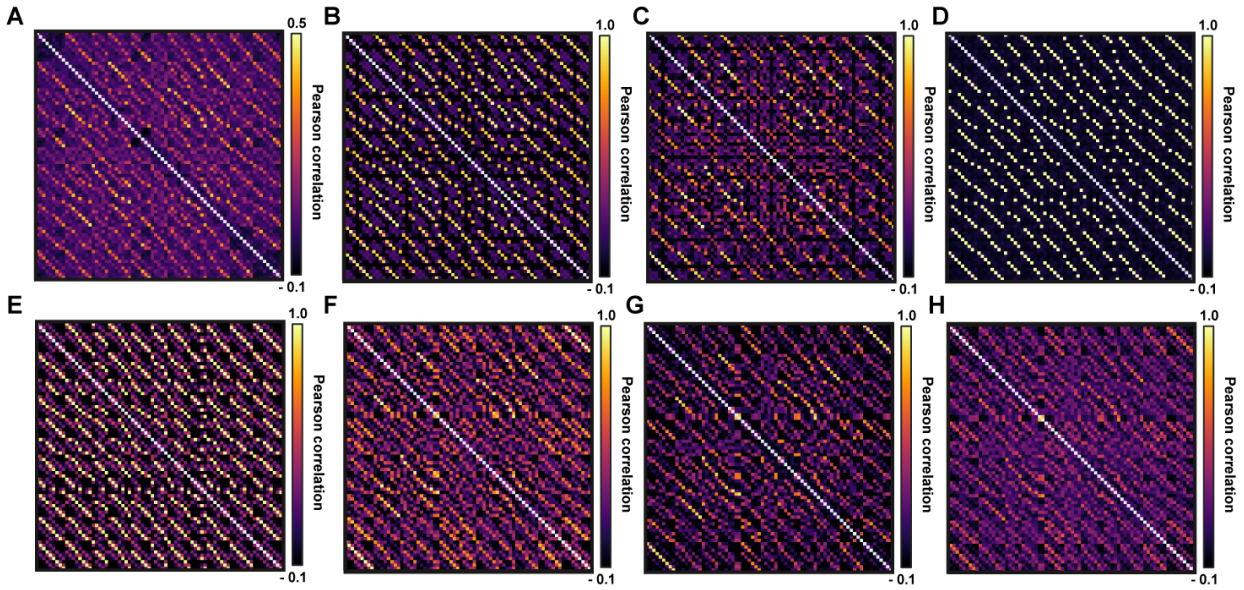

Fig. S5. RSMs for CA1 data and theoretical model predictions. The figure shows the resulting RSM for (**A**) aggregate CA1 data, (**B**) the GC2PC model without inhibition, (**C**) the boundary-tethered GC2PC model without inhibition (GC2PC-bt), (**D**) the naïve stability model, (**E**) the GC2PC model with inhibition (th-GC2PC), (**F**) the boundary-tethered GC2PC model with inhibition (bt-GC2PC-th), (**G**) the boundary-vector to place model (BVC2PC), and (**H**) the boundary-vector to successor feature model (BVC2SF).

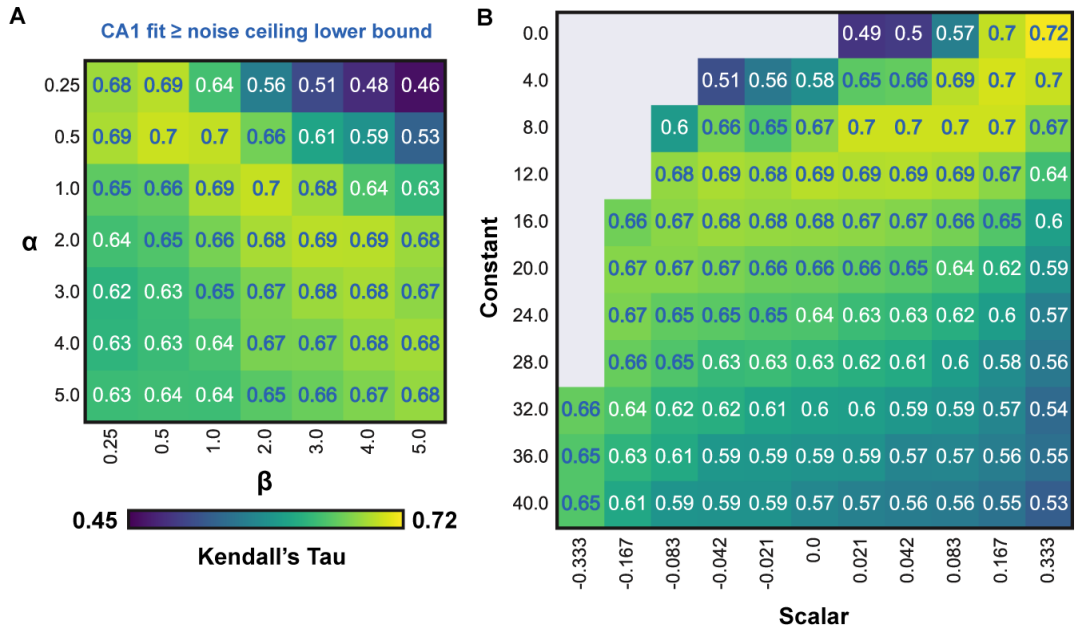

Fig. S6. BVC2PC model grid search with prediction accuracy for CA1 representation. (**A**) Resulting model fit to CA1 (Kendall's Tau) of alpha and beta distribution parameterizations that determine BVC tuning distances ($N = 49$). White labels indicate model parameterizations that fall below noise ceiling lower bound, while blue labels indicate model parameters that result in predicted representations equal to or above noise ceiling lower bound. (**B**) Model fit to CA1 (Kendall's Tau) for BVC distance constant and scalar parameters, wherein blue labels indicate model parameterizations equal to or above noise ceiling lower bound ($N = 104$).

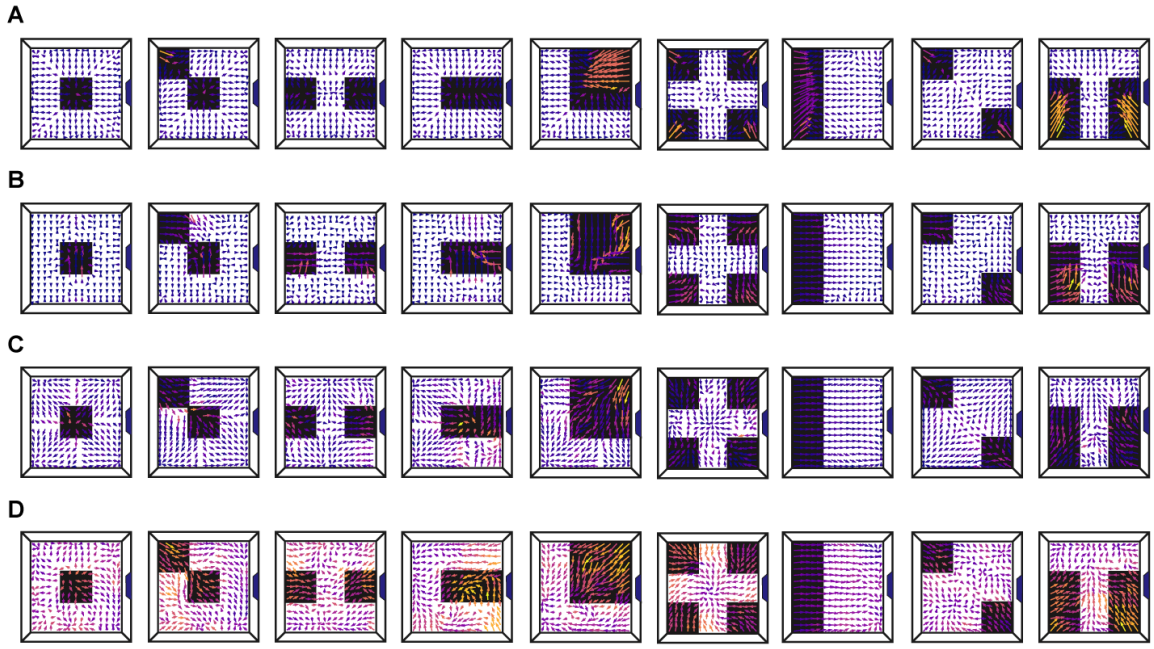

Fig. S7. Model-based vector field maps. (**A**) GC2PC model with inhibition, (**B**) boundary-tethered GC2PC model with inhibition, (**C**) BVC2PC model, and (**D**) the BVC2SF model.