Behavioral Engagement and Manifold Representation in the Hippocampus: Evidence from the Mutual Information of Population Encoding and Location

Shagesh Sridharan [1] **Anirvan Sengupta** [1,2]

[1] Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08901
[2] CCM and CCQ, Flatiron Institute, New York, NY 10010
shagesh.sridharan@rutgers.edu, anirvans.physics@gmail.com

Abstract

Although there is a significant understanding of how individual neurons in the hippocampus represent spatial location, the temporal dependence of population coding remains poorly understood. Using a novel statistical estimator and theoretical modeling, both developed in the framework of maximum entropy models, we reveal temporal changes in the fidelity of the spatial map, consistent with observed gating due to behavioral engagements.

1 Introduction

One of the central goals of neuroscience involves understanding the intricacies of the hippocampus (3). This part of the brain is involved in handling spatial tasks (38). Considering the similarity in hippocampus structure and functionality of mice and of that found in humans (27; 2; 6), results from experiments involving spatial tasks performed on mice provide a great degree of understanding of the human hippocampus.

Experiments on mammals with hippocampal lesions show that the hippocampus acts as a working memory by encoding environmental stimulus in neuron-neuron interactions (15). This includes environmental cues, such as sensory cues and location of rewards (43; 41; 1; 19), that help identify the spatial region the animal is present in, enabling the formation of an internal spatial code (36; 35; 31; 39). The encoding remaps when the animal is introduced to a different environment or set on a different task. This includes an introduction to a different positional space, usage of different spatial cues in the same space, and movement of reward locations (32; 25). Further, an environmental space can have multiple coexisting encodings in the hippocampus (44) which are switched around on separate introductions to the same space. Hippocampal codes are also based on egocentric information, including self-motion (28; 24), and often deteriorate when active motion is not present (16; 46; 10).

Current hippocampus models hypothesize that sensory and motion signals sub-consciously form an internal map of the spatial region when an animal is in active motion through an environment with sensory cues (24; 10; 20; 8; 4; 5). Consistent with this, when animals are introduced to new environments, spatial map formation happens before they realize their need, for instance prior to experiencing rewards (22; 17). In addition, even when animals are set on tasks that do not necessitate hippocampal activity, such as random foraging (9; 12), spatial codes are observed. Conjointly, this suggests that the hippocampus preserves a robust spatial map of the environment during motion, and this code is remapped when the environment changes.

In this paper, we analyze calcium indicator GCaMP3 fluorescence activity of hippocampal neurons in mice undergoing a goal-oriented activity. We outline a method to measure the fidelity of the internal spatial code and show that the spatial environment does not fully explain the hippocampal activity.

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

2 Quality of manifold representation

Hippocampal neurons manifest persistent place cell activity, producing an accurate representation of the environment in which each spatial location is mapped to a specific neural population state of activity (47; 48). Recent experiments (37) have demonstrated that the population activity profile can vary as the trial progresses, while identical visual cues and reward locations are used. This raises the question of whether the embedded position manifold has remapped or degraded.

The instantaneous activity of a sample of N neurons can be conceptualized as a point, or a neural state, in an N-dimensional Euclidean coordinate system, described as the neural state space. Within this framework, one can define neural dimensionality as the dimensionality of the entire state space (N). Neural dimensionality corresponds to the space of all possible neural states and does not indicate anything about the structure of the encoded activity. Task-relevant computations depend on correlated neural activity and can be equivalently described in an embedded space whose dimensionality is lower than that of the state space (18). This reduced dimensionality can be quantified, for instance, by applying dimensionality reduction techniques to the neural data, and determining the number of independent dimensions that account for a given fraction of variance.

The embedded dimensionality carries information about the structure, but it does not in general correspond to the number of variables that describe the environmental space. For instance, a circular virtual track can be described by a single variable embedded in a space with any number of dimensions. We call the number of independent variables the intrinsic dimension and use the physical dimension to refer to the number of variables used to describe the environmental space. Manifold learning in the neural state space can find the embedded manifold that contains the structural information about the environmental space manifold (34; 40; 26; 45; 42). Estimation of the information about the position space present in the neural state space leads to a measure of the quality of manifold representation.

2.1 Fidelity of representation via mutual information

Mutual Information (MI) (11) can be used to measure higher-order statistical dependencies between two random variables X and Y. The MI relies on the mismatch between p_{XY}, the joint probability density function (pdf) of the random variables X, Y, with the product of p_X and p_Y, representing the respective marginal pdfs. In our analysis, the expression of MI as the difference between the entropy $H(X)$ and the conditional entropy $H(X|Y)$ would be very important.

$$I(X;Y) := \mathbb{E}_{XY}\left[\log \frac{p_{XY}(X,Y)}{p_X(X)p_Y(Y)}\right] = H(X) - H(X|Y) \tag{1}$$

Hippocampus models of spatial maps have a structure conducive to analysis using a MI measure. They function as an encoder that takes environment information X as input, consisting of visual information, smell, position, and other environmental cues, and map it to a neural representation Y using $\Phi(X, \Lambda)$, where Λ refers to the set of hyper-parameters describing the tuning curves. Then, the fidelity of representation can be quantified by estimating MI between the latent code and the neural encoding.

Estimating MI in high-dimensional spaces is an infamously difficult task. As an example, consider the problem of estimating mutual information between neural activity, Y, and position, X, by binning the joint distribution space, with each individual neural activity being modeled as a binary random variable. We will need $O(2^N L/k)$ data points where N is the number of neurons, and k is the resolution of a track with length L. Instead, we estimate $H(X)$ and $H(X|Y)$. We need good estimates of place cell fields to find $H(X|Y)$, which has a lower data complexity due to their position selectivity (33; 29).

We have to compute the posterior distribution $p(x|y_t)$, where x is position at time t and y_t is the neuron activity vector with $y_{ti} = 1$ if neuron i spikes at time t and $y_{ti} = 0$ otherwise. We do this using the x-dependent activity of individual neurons and the assumption of conditional independence. We then estimate conditional entropy, $H(X|Y)$, as follows:

$$H(X|Y) \approx \frac{1}{T}\sum_t H(X|Y = y_t) = \frac{1}{T}\sum_t \int dx \left(p(x|y_t)\log\frac{1}{p(x|y_t)}\right) \tag{2}$$

2.2 Temporal dependence of MI

The Mutual Information estimate, representative of the fidelity of the spatial code in the hippocampus, depends on the formation of well-structured place cells tiling the entire track. The estimates will be lower if the track is sparsely tiled with cell fields, or if the cells have wide fields. A spatial code remapping will lead to a different but similarly high MI content representation. We evaluate the spatial code fidelity by partitioning the data into shorter temporal portions and learning the place cell fields for each partition separately using kernel density estimation. We use the estimated cell fields to compute posterior distribution and mutual information for the partition in consideration. This way we can compare mutual information across time. We correct our mutual information measures for bias and estimate the associated error (not shown here) using non-parametric bootstrap techniques (7).

3 Results

We show in Fig.1 example activity representative of CA1 neurons measured using cellular-resolution calcium imaging. This data, shared with us by collaborators in Princeton Neuroscience Institute, corresponds to a goal-oriented experiment involving water-restricted mice given water reward delivery in a virtual environment with visual cues uniquely determining every position on the track (19).

In this experiment, mice were trained to travel through a virtual reality environment in an enclosure that allowed simultaneous two-photon imaging at cellular resolution (21; 13; 14). The virtual environment was a linear track with a mixture of wall patterns and colors that provided unique visual stimuli at each point. The objective used here was a water reward given at a specific location in the environment. Mice were water constricted before introduction into the environment, and when they reached a certain place on the track, a small water reward was delivered from a tube present near the mouth. After the end of the track, the same visual stimuli and reward delivery were repeated, creating the effect of an infinite repeating corridor.

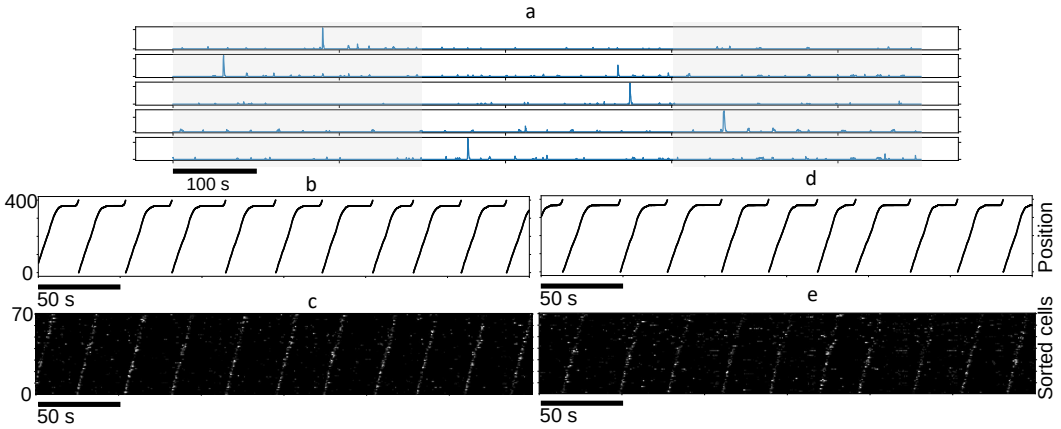

Figure 1: a, Examples of GCaMP3 $\Delta F/F$ from the activity of five CA1 neurons in a 15-min session, with the setup described in (19). b, d, Track positions for highlighted portions of example session. c, e, Plots of neural activity on each frame, with neurons sorted by location of maximum activity. The activity of each neuron is normalized. Only neurons that encode position information are shown.

As expected (47; 48), we find, Fig.1c, place cell activity in a large fraction of neurons, forming a manifold that tiles the entire track in position space. Toward the latter part of many sessions, this tiling pattern is disrupted, and the same neural population shifts to a distinctly different intrinsic code Fig.1e. This change in activity occurred while the mouse continued to traverse the track and identical visual cues were used. This is similar to the gating of hippocampal place codes observed in (37).

We validate our assumption regarding independent place cell firing by computing the correlation between neuron activities conditioned on position. We observe that our assumption is largely fulfilled. There are a few outliers that are highly correlated, neurons encoding reward, but position encoding neurons exhibit low position conditional correlation. We use kernel density estimation to model $\phi_i(x)$, the rate of firing of place cell i at position x, for its advantages with small data sets (30). We

determine the size of the RBF kernel by maximizing the posterior for each neuron i and averaging over each neuron's optimal bandwidth.

We use binarized mutual information of the mouse's spatial position and the neural activity to show that the change in neural activity is a degradation in the spatial code, instead of a remapping of the place code, Fig.2. We have binarized MI into low and high using $MI_c = 0.25$ bits. This cutoff was decided using the lowest measure among the beginning parts of the sessions since we expect well-behaved spatial maps early on.

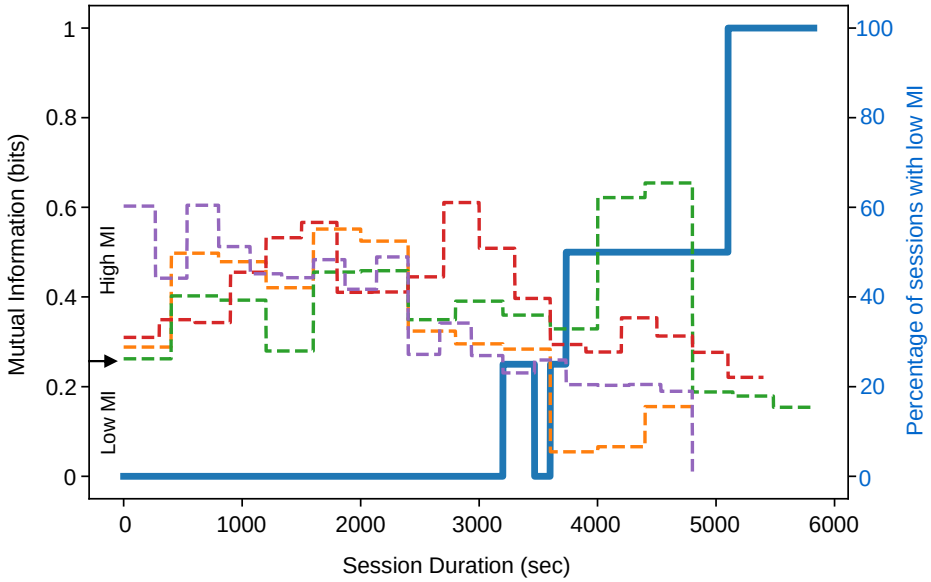

Figure 2: Dashed lines depict mutual information across session length for different sessions. Only sessions longer than 1h and exhibiting spatial code at some point during the session are shown. The solid blue line shows the percentage of sessions that have low mutual information across session length. Mutual information was binarized into low and high using $MI_c = 0.25$ bits

4 Conclusion

Our results show that the hippocampal intrinsic code of the spatial map consistently deteriorates with session duration. This breakdown of spatial code happens while the mice continued to traverse through the same virtual environment. Consequently, it is likely that sensory and motion signals are insufficient to maintain a spatial map in the hippocampus. Rather, some other internal state regulates the fidelity of the spatial maps and serves as a functional gate for the stable firing of place cells.

We speculate that an active internal mechanism controls the behavioral relevance of the environment and hence applies a top-down influence to gate spatial representations in CA1 (23). In that case, the internal state changes would modulate the fidelity of the hippocampal place code. It is also possible that the changes in place codes that we observed arise for completely different reasons, and lead to a change in the associated behavior. Future experiments are needed to test the underlying mechanisms and the internal states. Regardless of mechanisms, our observations reveal that the hippocampus does not always preserve a spatial code and that place maps degrade even as animals experience identical visual cues and spatial positions.

Our work relies on modeling responses of individual neurons, conditional upon the same input, to be independent. Our preliminary analysis (not shown here) indicates that this is mostly true of neurons coding for positions, but not so much for reward neurons. Still, further studies are needed to find out whether correlations, other than those induced by the overlap of place fields, are crucial in coding. Another limitation is that we are limited to the set of neurons that are being observed, and we have to hope to measure mutual information with their activity to provide the right trends for the overall hippocampal system. The confidence we have in our conclusion, thus, partly relies on corroborative evidence from other studies (37).

5 Acknowledgements

We thank J.L. Gauthier and D.W. Tank for the experimental setup, data collection, and fluorescence image analysis. We also thank D.B. Chklovskii, Y. Bahroun, and A. Rubin for multiple discussions. This work is supported by a grant (SF 626323) to AS from the Simons Foundation .

References

[1] Acharya, L., Aghajan, Z.M., Vuong, C., Moore, J.J., Mehta, M.R.: Causal influence of visual cues on hippocampal directional selectivity. Cell **164**(1-2), 197–207 (2016)

[2] Allen, T.A., Fortin, N.J.: The evolution of episodic memory. Proceedings of the National Academy of Sciences **110**(supplement_2), 10379–10386 (2013)

[3] Andersen, P., Morris, R., Amaral, D., Bliss, T., O'Keefe, J.: The hippocampus book. Oxford university press (2006)

[4] Arleo, A., Gerstner, W.: Spatial cognition and neuro-mimetic navigation: a model of hippocampal place cell activity. Biological cybernetics **83**(3), 287–299 (2000)

[5] Barry, C., Burgess, N.: Neural mechanisms of self-location. Current Biology **24**(8), R330–R339 (2014)

[6] Bergmann, E., Zur, G., Bershadsky, G., Kahn, I.: The organization of mouse and human cortico-hippocampal networks estimated by intrinsic functional connectivity. Cerebral Cortex pp. 1–16 (2016)

[7] Berrar, D.: Introduction to the non-parametric bootstrap. (2019)

[8] Bittner, K.C., Grienberger, C., Vaidya, S.P., Milstein, A.D., Macklin, J.J., Suh, J., Tonegawa, S., Magee, J.C.: Conjunctive input processing drives feature selectivity in hippocampal ca1 neurons. Nature neuroscience **18**(8), 1133–1142 (2015)

[9] Butler, W.N., Hardcastle, K., Giocomo, L.M.: Remembered reward locations restructure entorhinal spatial maps. Science **363**(6434), 1447–1452 (2019)

[10] Chen, G., King, J.A., Burgess, N., O'Keefe, J.: How vision and movement combine in the hippocampal place code. Proceedings of the National Academy of Sciences **110**(1), 378–383 (2013)

[11] Cover, T.M., Thomas, J.A.: Wiley series in telecommunications and signal processing. elements of information theory (2006)

[12] Danielson, N.B., Zaremba, J.D., Kaifosh, P., Bowler, J., Ladow, M., Losonczy, A.: Sublayer-specific coding dynamics during spatial navigation and learning in hippocampal area ca1. Neuron **91**(3), 652–665 (2016)

[13] Dombeck, D.A., Harvey, C.D., Tian, L., Looger, L.L., Tank, D.W.: Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. Nature neuroscience **13**(11), 1433–1440 (2010)

[14] Domnisoru, C., Kinkhabwala, A.A., Tank, D.W.: Membrane potential dynamics of grid cells. Nature **495**(7440), 199–204 (2013)

[15] Douglas, R.J.: The hippocampus and behavior. Psychological bulletin **67**(6), 416 (1967)

[16] Foster, T.C., Castro, C.A., McNaughton, B.L.: Spatial selectivity of rat hippocampal neurons: dependence on preparedness for movement. Science **244**(4912), 1580–1582 (1989)

[17] Frank, L.M., Stanley, G.B., Brown, E.N.: Hippocampal plasticity across multiple days of exposure to novel environments. Journal of Neuroscience **24**(35), 7681–7689 (2004)

[18] Gao, P., Trautmann, E., Yu, B., Santhanam, G., Ryu, S., Shenoy, K., Ganguli, S.: A theory of multineuronal dimensionality, dynamics and measurement. BioRxiv p. 214262 (2017)

[19] Gauthier, J.L., Tank, D.W.: A dedicated population for reward coding in the hippocampus. Neuron **99**(1), 179–193 (2018)

[20] Gothard, K.M., Skaggs, W.E., McNaughton, B.L.: Dynamics of mismatch correction in the hippocampal ensemble code for space: interaction between path integration and environmental cues. Journal of Neuroscience **16**(24), 8027–8040 (1996)

[21] Harvey, C.D., Collman, F., Dombeck, D.A., Tank, D.W.: Intracellular dynamics of hippocampal place cells during virtual navigation. Nature **461**(7266), 941–946 (2009)

[22] Hill, A.: First occurrence of hippocampal spatial firing in a new environment. Experimental neurology **62**(2), 282–297 (1978)

[23] Kentros, C.G., Agnihotri, N.T., Streater, S., Hawkins, R.D., Kandel, E.R.: Increased attention to spatial context increases both place field stability and spatial memory. Neuron **42**(2), 283–295 (2004)

[24] Knierim, J.J., Kudrimoti, H.S., McNaughton, B.L.: Interactions between idiothetic cues and external landmarks in the control of place cells and head direction cells. Journal of neurophysiology **80**(1), 425–446 (1998)

[25] Leutgeb, S., Leutgeb, J.K., Barnes, C.A., Moser, E.I., McNaughton, B.L., Moser, M.B.: Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. Science **309**(5734), 619–623 (2005)

[26] Low, R.J., Lewallen, S., Aronov, D., Nevers, R., Tank, D.W.: Probing variability in a cognitive map using manifold inference from neural dynamics. BioRxiv p. 418939 (2018)

[27] Manns, J.R., Eichenbaum, H.: Evolution of declarative memory. Hippocampus **16**(9), 795–808 (2006)

[28] McNaughton, B.L., Battaglia, F.P., Jensen, O., Moser, E.I., Moser, M.B.: Path integration and the neural basis of the'cognitive map'. Nature Reviews Neuroscience **7**(8), 663–678 (2006)

[29] Mizuseki, K., Royer, S., Diba, K., Buzsáki, G.: Activity dynamics and behavioral correlates of ca3 and ca1 hippocampal pyramidal neurons. Hippocampus **22**(8), 1659–1680 (2012)

[30] Moon, Y.I., Rajagopalan, B., Lall, U.: Estimation of mutual information using kernel density estimators. Physical Review E **52**(3), 2318 (1995)

[31] Morris, R.G., Garrud, P., Rawlins, J.a., O'Keefe, J.: Place navigation impaired in rats with hippocampal lesions. Nature **297**(5868), 681–683 (1982)

[32] Muller, R.U., Kubie, J.L.: The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. Journal of Neuroscience **7**(7), 1951–1968 (1987)

[33] Neunuebel, J.P., Knierim, J.J.: Spatial firing correlates of physiologically distinct cell types of the rat dentate gyrus. Journal of Neuroscience **32**(11), 3848–3858 (2012)

[34] Nieh, E.H., Schottdorf, M., Freeman, N.W., Low, R.J., Lewallen, S., Koay, S.A., Pinto, L., Gauthier, J.L., Brody, C.D., Tank, D.W.: Geometry of abstract learned knowledge in the hippocampus. Nature **595**(7865), 80–84 (2021)

[35] O'Keefe, J.: Place units in the hippocampus of the freely moving rat. Experimental neurology **51**(1), 78–109 (1976)

[36] O'Keefe, J., Dostrovsky, J.: The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. Brain research (1971)

[37] Pettit, N.L., Yuan, X.C., Harvey, C.D.: Hippocampal place codes are gated by behavioral engagement. Nature neuroscience **25**(5), 561–566 (2022)

[38] Prior, H., Schwegler, H., Dücker, G.: Dissociation of spatial reference memory, spatial working memory, and hippocampal mossy fiber distribution in two rat strains differing in emotionality. Behavioural brain research **87**(2), 183–194 (1997)

[39] Robinson, N.T., Descamps, L.A., Russell, L.E., Buchholz, M.O., Bicknell, B.A., Antonov, G.K., Lau, J.Y., Nutbrown, R., Schmidt-Hieber, C., Häusser, M.: Targeted activation of hippocampal place cells drives memory-guided spatial behavior. Cell **183**(6), 1586–1599 (2020)

[40] Rubin, A., Sheintuch, L., Brande-Eilat, N., Pinchasof, O., Rechavi, Y., Geva, N., Ziv, Y.: Revealing neural correlates of behavior without behavioral measurements. Nature communications **10**(1), 1–14 (2019)

[41] Save, E., Nerad, L., Poucet, B.: Contribution of multiple sensory information to place field stability in hippocampal place cells. Hippocampus **10**(1), 64–76 (2000)

[42] Sengupta, A., Pehlevan, C., Tepper, M., Genkin, A., Chklovskii, D.: Manifold-tiling localized receptive fields are optimal in similarity-preserving neural networks. Advances in neural information processing systems **31** (2018)

[43] Shapiro, M.L., Tanila, H., Eichenbaum, H.: Cues that hippocampal place cells encode: dynamic and hierarchical representation of local and distal stimuli. Hippocampus **7**(6), 624–642 (1997)

[44] Sheintuch, L., Geva, N., Baumer, H., Rechavi, Y., Rubin, A., Ziv, Y.: Multiple maps of the same spatial context can stably coexist in the mouse hippocampus. Current biology **30**(8), 1467–1476 (2020)

[45] Tepper, M., Sengupta, A.M., Chklovskii, D.: Clustering is semidefinitely not that hard: Nonnegative sdp for manifold disentangling. The Journal of Machine Learning Research **19**(1), 3208–3237 (2018)

[46] Terrazas, A., Krause, M., Lipa, P., Gothard, K.M., Barnes, C.A., McNaughton, B.L.: Self-motion and the hippocampal spatial metric. Journal of Neuroscience **25**(35), 8085–8096 (2005)

[47] Wilson, M.A., McNaughton, B.L.: Dynamics of the hippocampal ensemble code for space. Science **261**(5124), 1055–1058 (1993)

[48] Zhang, K., Ginzburg, I., McNaughton, B.L., Sejnowski, T.J.: Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. Journal of neurophysiology **79**(2), 1017–1044 (1998)