Mapping neural representations of topologically non-trivial spaces

Maxim Beketov MAXBEKETOV@OUTLOOK.COM
HSE University, Moscow, Russia

Konstantin Sorokin
HSE University, Saint-Petersburg, Russia

Pavel Snopov
University of Texas Rio Grande Valley, Brownsville, Texas, USA

Anton Ayzenberg
Noeon Research, Tokyo, Japan

Konstantin Anokhin
Institute for Advanced Brain Studies, Moscow State University, Moscow, Russia

Abstract

We consider a biologically realistic artificial model of neurons, simulating place cells (PCs) found in the hippocampus of living brains, that are known to encode the physical space. We make our model encode such spaces of distinct topological structure (namely, 2D arenas with 1, 2, 3 holes in them), in which an artificial agent is moving, with its simulated PCs firing when the agent visits the receptive fields of the corresponding PCs. We analyze – with persistent homology and Isomap – such signals (artificial PCs' spike trains), showing that their topological properties reflect those of the physical space (encoded by the PCs as the agent explores the arenas), thus demonstrating that said properties can be successfully recovered from such signal alone.

Keywords: Place cells, cognitive maps, topological data analysis, neural representations.

1. Introduction

In recent years interest to the role of the hippocampus in various learning tasks, including spatial navigation, is gaining popularity. It is of no surprise, as the hippocampus is known to play a role of a "universal integrator": its cells specialize on multiple concepts, with flexible remapping mechanisms. Namely, the same neurons are known to store spatial (O'Keefe and Nadel, 1978), audial (Aronov et al., 2017) and non-spatial domain (Eichenbaum, 2015). It was also shown, modeling the hippocampus in tandem with the entorhinal cortex, as an encoder – that together the two store different kinds of information (Whittington et al., 2020). Despite the above said, the role of the hippocampus in storing more abstract, higher-order information – e.g., the topology of the ambient physical space (Dabaghian et al., 2012) – can still be considered an open question. Our work is inspired by the works (Curto and Itskov, 2008), where the topological properties of the physical space encoded by artificial place cells were investigated. In case of natural experiments with animals, like rodents or bats, one does not expect them to understand the concept of topology explicitly, but the information about loops, paths and the overall structure of the space is hypothesized to be understood by animals as a higher-order concept (Chen et al., 2014; Baas, 2016).

© 2025 M. Beketov, K. Sorokin, P. Snopov, A. Ayzenberg & K. Anokhin.

Moreover, there are studies which made successful attempts to extract more fine-grained – geometric – information from the spiking activity of the hippocampal PCs encoding the real world (Giusti et al., 2015). In present study we construct an artificial model (consensual with other studies) of hippocampal PCs and explore how it can encode the topological structure of information (of the ambient physical space) and find if standard approaches to dimensionality reduction can recover such topologies represented by PCs neural spiking activity.

2. Model

We consider the following artificial model of topologically non-trivial place cells signal. The living object is moving on the arena of circular geometry (a circle of radius $R=5$ length units centered at the origin of a Cartesian plane), but of 3 distinct topologies: i.e. with $1, 2$ or 3 holes in said geometry. For the cases of $2, 3$ holes, these are circles of radii $r_{\text{hole}}=2$ centered at points at distance of $\delta=2.5$ in directions of polar angle $2\pi k/n$ (k from 0 to $n-1$ for $n=2,3$) from the origin. For the case of 1 hole, it is a circle of radius $r_{\text{hole}}=3.5$ centered at the origin. Holes are "removed" from the geometry – thus being regions impenetrable for the object as it moves in the arena.

The arena (excluding the interior of holes) is covered with circles that are receptive fields (RFs) of artificial place cell neurons. The covering is achieved with Poisson disk sampling: on each trial, a point is chosen uniformly randomly from the bounding square of the arena circle – it either belongs to the geometry (is inside the arena circle and not inside the holes) or not; if it does – one tries to place a circle of radius r_{PoisDisc} centered at this point – if the point does not land into any of previously placed circles, its circle is added to the covering. In demonstrations shown in present paper this $r_{\text{PoisDisc}}=0.04$.

The object explores the arena, performing what's called a Levy flight random walk (RW) on it: being at position $\boldsymbol{r}$, it is proposed to make a step $\Delta\boldsymbol{r}$ – a displacement of its position in a direction given by a (uniformly in $[0, 2\pi]$) random angle and of step length r sampled from a Pareto distribution – i.e. having probability density proportional to $r^{-\alpha}$ – some (decaying) power law. In demonstrations shown in present paper this exponent was taken $\alpha = 1.5$ – as suggested by experimental evidence for certain animals of interest (namely, mice) (Shoji, 2016). If the proposed displacement $\Delta\boldsymbol{r}$ leaves the object in the geometry and does not cross (midpoint of displacement vector is not inside) a hole – such a step is added to the RW trajectory. To model the fact that animals do not move the entire time, sometimes no displacement is made (so two same consecutive points are added to the trajectory) – this is decided randomly with a Bernoulli trial, in demonstrations shown we've picked the fraction of discrete time steps where the object is moving to be 99%. On the LHS of Figure 1 one can see an arena and a such a trajectory of 170 time steps.

Given a trajectory of the object in the arena, we sample a signal of artificial place cell (PC) neurons encoding this sequence of object positions, modeling it as a stationary Poisson process (Dabaghian et al., 2012; O'Keefe and Dostrovsky, 1971): each PC has a receptive field (RF) – a Gaussian function giving the following dependence:

$$\text{firing_rate}_{\text{i_neuron}}(\boldsymbol{r}(t)) \;\propto\; \sigma_{\text{RF}}^{-1}\,\exp(-||\boldsymbol{r}-\boldsymbol{c}_{\text{RF}}||^2/2\sigma_{\text{RF}}^2) \tag{1}$$

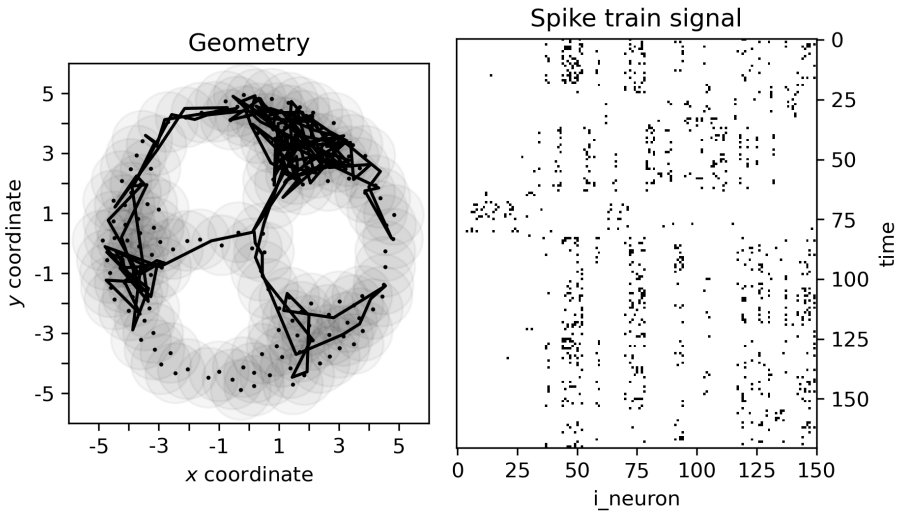

Figure 1: **Left**: arena with 3 holes, covered with RFs of 150 place cells, and a trajectory on it. **Right**: signal of 150 place cell neurons during 170 time steps of object movement along the trajectory, black dot = neuron is activated.

– each neuron is firing the more frequently, the closer the object's position $\boldsymbol{r}$ (at time t) is to the center $\boldsymbol{c}_{\mathrm{RF}}$ of the neuron's RF. In our demonstrations, $\sigma_{\mathrm{RF}}=1$, and circles covering the arena with RFs on Figure 1 are drawn with such radius for scale.

To simulate the neurons' firing process, recall that a Poisson random variable (RV) with rate λ is approximated with n Bernoulli trials with success probability p, where $\lambda=np$ – so, since Formula 1 gives the firing rate up to proportionality, we just split the time axis into discrete time steps (for each the trajectory gives $\boldsymbol{r}(t)$), and for each neuron at each time step sample a Bernoulli RV (0 = failure or 1 = success) with success probability given by Equation (1) (we experimentally pick a normalizing constant, same for all neurons, so that the spikes are "nor too frequent, neither too rare") – obtaining a sample of "spike train" signal seen on the RHS of Figure 1. It is important to note that, conditioned on $\boldsymbol{r}(t)$, constructed Bernoulli events are independent – which, from what we know, is neurobiologically realistic for living place cell neurons at the stage of initially "learning" to encode the physical environment (but not necessarily when the environment is "well known" to these). To model measurement error always present in real experiments, we corrupt the spike train signal with zero-mean Gaussian noise (in our demonstrations, of $\sigma_{\mathrm{noise}}=0.01$).

3. Analysis

Given artificial signal ($3{\cdot}10^3$ time steps) of place cells encoding the object's positions along the trajectory, we do two basic things to analyze it. First, we compute its persistent homology (PH) (Edelsbrunner et al., 2002; Zomorodian and Carlsson, 2004) (more formally, persistent homology of Vietoris-Rips (VR) filtration of Euclidean metric of the discrete space – point cloud, over $\mathbb{Z}_2$) with very effective Ripser software package (Bauer, 2021). It

is important to specify how does one define the metric: the signal S_{ij} is a matrix (seen as a spike train the RHS of Figure 1) of numbers that are close (but not equal, being corrupted by noise) to 0 or 1 (giving points close to vertices of a hypercube). Each row $S_{i\cdot}$ of this matrix can be seen as a vector of "cognitive coordinates" (neural representation in place cells' activity) of the object's position at that moment of time. We normalize these vectors (divide each row of S_{ij} by its 2-norm), so now their endpoints "live" on a hypersphere – so, as is well known, cosine distance between these points is, up to a constant, same as Euclidean distance between these. Cosine similarity of such signals is a reasonable measure of their likeliness, since it is also related to their correlation. With this normalization, "feeding" rows of S_{ij} to Ripser as Euclidean points is legit and reveals their persistent homology (as if computed from filtration in cosine similarity). From Alexandroff's Nerve theorem (Alexandroff, 1928; Mona and Hintz, 2023) it follows that the nerve of the covering of the arena with place cells' RFs is homotopy equivalent to the underlying space, hence the topology of the nerve "reflects" the topology of the arena (their homologies coincide).

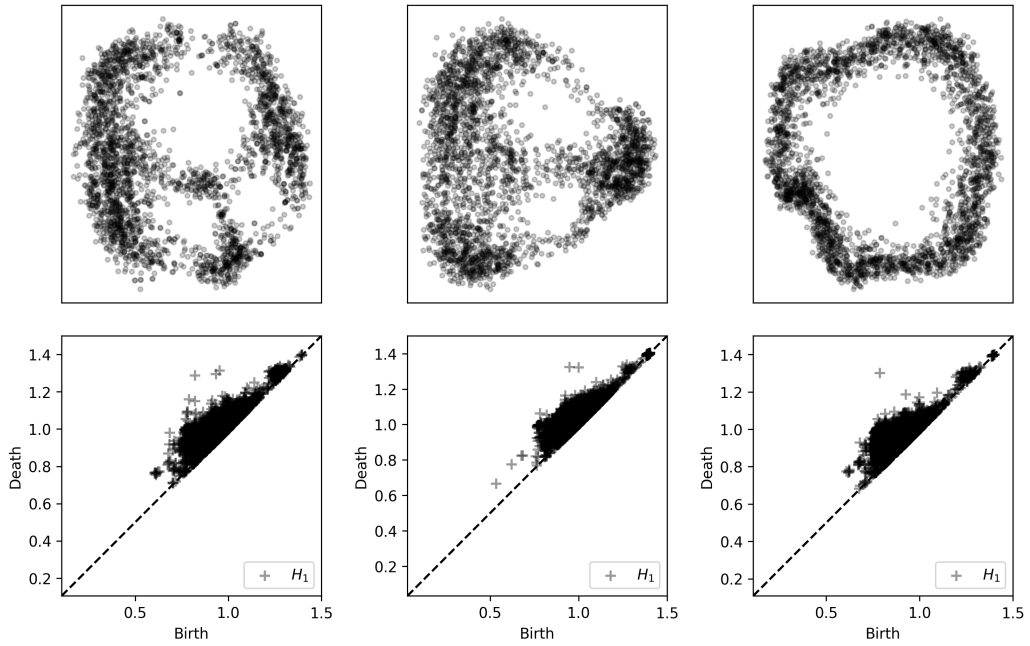

Figure 2: Arena topology reconstruction from neural signal. **Top**: Isomap embeddings of "cognitive maps" of 3,2,1-hole arenas respectively. **Bottom**: persistence diagrams of corresponding neural signals (in their original dimension $\sim$150).

On Figure 2 one can see the persistence diagrams (PDs) of the signal for arenas (with 3, 2, 1 holes respectively) in its original high-dimensional space: there are exactly as many persistent (long-living) 1-homologies in the signal, as many holes there are in the arena.

Secondly, we compute Isomap (Tenenbaum et al., 2000) embeddings of the signal – each row of S_{ij} is "projected" to a point in 2D plane. By preserving distances (with metric multidimensional scaling (MDS) part of Isomap) to nearest neighbours (after normalization,

Euclidean distance = 2-norm of difference between rows of S_{ij} = cosine distance between them, up to a constant), Isomap can be said to preserve local intrinsic geometry of the data. By doing that, since Euclidean distances between resulting 2D embeddings are forced (by MDS) to approximate distances in the nearest-neighbor graph of original data, Isomap effectively also preserves its topological structure – PHs (of VR filtration in Euclidean distance) of Isomap "projections" reveal very similar topology to that of the original signal.

4. Discussion

Considering a biologically realistic artificial model of place cells and their signal, we have shown that, if current understanding of mechanisms of neural encoding (mapping) of the physical space (O'Keefe and Dostrovsky, 1971) (as well as other spaces (O'Keefe and Nadel, 1978)) by neurons in living brains – is accurate – then the topological properties of the underlying physical space can very well be inferred from its cognitive map. Present work is, of course, not the first to explore said question – we hope that our findings contribute to this field of research at the intersection of neuroscience and topological data analysis.

We're currently exploring the applicability of machine learning models capable of autoencoding (Moor et al., 2020; Trofimov et al., 2023) high-dimensional signals to their lower-dimensional representations, preserving their topological structure – but, we should report, that we've had mixed success at that so far, if applying such models in their originally proposed form, with relatively shallow encoder and decoder neural networks. We're looking forward to share these preliminary results, hoping for a discussion of possible improvements of such topological machine learning-based approaches to our problem.

Acknowledgments

The work of M. Beketov and K. Sorokin was supported by the grant for research centers in the field of AI provided by the Ministry of Economic Development of the Russian Federation in accordance with the agreement 000000C313925P4E0002 and the agreement with HSE University № 139-15-2025-009.

5. Bibliography

References

Paul Alexandroff. Über den allgemeinen dimensionsbegriff und seine beziehungen zur elementaren geometrischen anschauung. *Mathematische Annalen*, 98(1):617–635, 1928.

Dmitriy Aronov, Rhino Nevers, and David W Tank. Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature*, 543(7647):719–722, 2017.

Nils A Baas. On the concept of space in neuroscience. *Curr. Opin. Syst. Biol.*, 1:32–37, 2016.

Ulrich Bauer. Ripser: efficient computation of Vietoris-Rips persistence barcodes. *J. Appl. Comput. Topol.*, 5(3):391–423, 2021. ISSN 2367-1726. doi: 10.1007/s41468-021-00071-5.

Zhe Chen, Stephen N Gomperts, Jun Yamamoto, and Matthew A Wilson. Neural representation of spatial topology in the rodent hippocampus. *Neural Comput.*, 26(1):1–39, 2014.

Carina Curto and Vladimir Itskov. Cell groups reveal structure of stimulus space. *PLoS computational biology*, 4(10):e1000205, 2008.

Y Dabaghian, F Mémoli, L Frank, and G Carlsson. A topological paradigm for hippocampal spatial map formation using persistent homology. *PLOS Computational Biology*, 8(8):1–14, August 2012. doi: 10.1371/journal.pcbi.1002581.

Edelsbrunner, Letscher, and Zomorodian. Topological persistence and simplification. *Discrete & computational geometry*, 28(4):511–533, 2002.

Howard Eichenbaum. The hippocampus as a cognitive map... of social space. *Neuron*, 87 (1):9–11, 2015.

Chad Giusti, Eva Pastalkova, Carina Curto, and Vladimir Itskov. Clique topology reveals intrinsic geometric structure in neural correlations. *Proceedings of the National Academy of Sciences*, 112(44):13455–13460, 2015.

Mohnhaupt Mona and S Kališnik Hintz. *The nerve theorem and its applications in topological data analysis*. PhD thesis, Bachelor's thesis, Swiss Federal Institute of Technology (ETH) Zurich, 2023.

Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt. Topological autoencoders. In *International conference on machine learning*, pages 7045–7054. PMLR, 2020.

John O'Keefe and Jonathan Dostrovsky. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 1971.

John O'Keefe and Lynn Nadel. *The hippocampus as a cognitive map*. Oxford university press, 1978.

Hiroto Shoji. Scaling law in free walking of mice in circular open fields of various diameters. *Journal of biological physics*, 42(2):259–270, 2016.

Joshua B Tenenbaum, Vin de Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

Ilya Trofimov, Daniil Cherniavskii, Eduard Tulchinskii, Nikita Balabin, Evgeny Burnaev, and Serguei Barannikov. Learning topology-preserving data representations. *arXiv preprint arXiv:2302.00136*, 2023.

James CR Whittington, Timothy H Muller, Shirley Mark, Guifen Chen, Caswell Barry, Neil Burgess, and Timothy EJ Behrens. The tolman-eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. *Cell*, 183(5): 1249–1263, 2020.

Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. In *Proceedings of the twentieth annual symposium on Computational geometry*, pages 347–356, 2004.