Ensemble Spatial Coding in the Retrosplenial Cortex

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

by

Vasudeva K. Bhat

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

Date: 20 March, 2025

Under the guidance of

Supervisor: **Dr. Hiroshi Ito**,

Max Planck Institute for Brain Research, Frankfurt am Main, Germany

From **June 2024** to **Mar 2025**

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE

Certificate

This is to certify that this dissertation entitled **"Ensemble Spatial Coding in the Retrosplenial Cortex"** towards the partial fulfillment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents work carried out by **Vasudeva K. Bhat** at **Max Planck Institute for Brain Research, Frankfurt am Main, Germany** under the supervision of **Dr. Hiroshi Ito**, Group leader, Memory and Navigation Circuits Group, Max Planck Institute for Brain Research, Frankfurt am Main, Germany during the academic year 2024-2025.

Supervisor: **Signature:**

Dr. Hiroshi Ito
Group leader,
Max Planck Institute for Brain Research,
Frankfurt am Main, Germany

Expert: **Signature:**

Dr. Collins Assisi
Assistant Professor,
Department of Biology,
Indian Institute of Science Education and Research,
Pune, India

This thesis is dedicated to my father, B. Krishnamurthy Bhat, for his endless love, support, and encouragement, which have been instrumental in my growth both as a person and as a scientist.

Declaration

I hereby declare that the matter embodied in the report entitled "***Ensemble Spatial Coding in the Retrosplenial Cortex***" is the result of the work carried out by me at **Max Planck Institute for Brain Research, Frankfurt am Main**, **Germany**, under the supervision of **Dr. Hiroshi Ito**, and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgment of collaborative research and discussions.

Signature:

Vasudeva K. Bhat
BS-MS Student
Roll no: 20201127

Table of Contents

List of Tables

List of Figures

Abstract

Spatial coding in the retrosplenial cortex (RSC) is critical for navigation and memory, yet how exactly it represents or encodes space remains to be fully understood. Additionally, the role of the medial entorhinal cortex (MEC) in driving RSC's spatial representation is not well characterized. Using electrophysiological recordings from rats performing an open-field foraging task, we systematically investigated spatial coding at both the single-cell and population levels. At the single-cell level, we first validated the existence of border and head-direction cells in the RSC, and also explored the possibility of spatially-tuned "place" cells coding for spatial information using Bayesian decoding. At the population level, we employed Linear Discriminant Analysis (LDA) decoding, revealing high spatial decodability. To further explore the geometry of RSC spatial representations, we analyzed neural manifolds, uncovering a low-dimensional structured representation underlying the population activity. We additionally carried out high-dimensional population-vector analyses of spiking activity, which revealed spatially segregated subspaces and orthogonal representations of unique locations. Finally, we applied Variational Autoencoders (VAEs) and Supervised VAEs (S-VAEs) to extract latent representations of spatial coding, mitigating the limitations of LDA. Our results demonstrated that RSC codes space/position at the level of neuronal ensembles. These findings provide new insights into how space is represented in the RSC and the role the MEC plays in driving this representation.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Hiroshi Ito, for his invaluable guidance, patience, and support throughout this project. His insights, expertise, and resources have been instrumental in shaping this work. I am also deeply grateful to my expert, Dr. Collins Assisi, for his invaluable input and support during the early part of my journey in neuroscience research, which played a significant role in shaping my current interests.

I would also like to thank Dr. Lisa Roux for introducing me to experimental neuroscience during my time in her lab last summer, where I gained crucial insights that laid the foundation for my future work in the field of systems neuroscience.

I am thankful to Dr. Marjan Mozaffrilegha for her guidance in both experimental and computational aspects, which proved instrumental in helping me progress with my project. I am also grateful to Joeri Van Wijngaarden and Dr. Susanne S. Babl, who conducted the experiments and collected the crucial data for my project. Although I have never met either of them, I am truly grateful for their tireless efforts.

I would also like to extend my gratitude to the members of the Memory and Navigation Circuits (Ito) Group for their support and for engaging in insightful discussions, which greatly enriched my understanding and experience in neuroscience.

On a personal note, I am thankful to my father, B. Krishnamurty Bhat, for his unwavering support, especially during my time away from home. I am also grateful to my friends, especially those from Room 249, who were a source of constant support and advice.

I extend my appreciation to the Max Planck Institute for Brain Research for providing the necessary resources during my stay. Finally, I would like to thank the Indian Institute of Science Education and Research, Pune, for providing me with this opportunity to carry out a master's thesis and truly experience the scientific process.

Contributions

Contributor name	Contributor role
Dr. Marjan Mozaffrilegha, Dr. Hiroshi Ito, Vasudeva K. Bhat	Conceptualization Ideas
Vasudeva K. Bhat, Dr. Marjan Mozaffrilegha	Methodology
Vasudeva K. Bhat, Dr. Marjan Mozaffrilegha	Software
Vasudeva K. Bhat, Dr. Marjan Mozaffrilegha	Validation
Vasudeva K. Bhat	Formal analysis
Dr. Joeri Van Wijngaarden, Dr. Susanne S. Babl	Investigation
Dr. Hiroshi Ito	Resources
Dr. Joeri Van Wijngaarden, Dr. Susanne S. Babl	Data Curation
Vasudeva K. Bhat	Writing - original draft preparation
Dr. Hiroshi Ito	Writing - review and editing
Vasudeva K. Bhat	Visualization
Dr. Marjan Mozaffarilegha, Dr. Hiroshi Ito	Supervision
Dr. Hiroshi Ito	Project administration
Dr. Hiroshi Ito	Funding acquisition

This contributor syntax is based on the Journal of Cell Science CRediT Taxonomy .

Chapter 1: Introduction

Stimulus-Response Learning

In the early days of neuroscience and psychology research, models of learning were rooted in stimulus-response associations. Behaviorists such as Thorndike and Skinner proposed that animals and humans navigated their environments through reinforced learning, in which actions were directly linked to rewards or punishments. Through repeated exposure, these associations shape learning and memory formation. Such a model emphasized the role played by habit formation over internal spatial representation (Thorndike EL, 1911; Skinner BF, 1938).

According to this framework, spatial learning was thought to rely on memorized sequences of motor responses (e.g., turning left or right at specific points) that were reinforced by rewards. In this paradigm, the focus was on learning the actions themselves, rather than constructing an internalized cognitive representation of space or the locations of rewards.

Over the years, a key limitation of this approach became increasingly apparent. The rigid nature of the model, which focused on memorizing fixed sequences of actions to reach a reward, proved inadequate for explaining more complex forms of spatial learning. If navigation were purely based on reinforced sequences, animals would struggle with tasks requiring novel shortcuts or detours. However, experimental findings began to suggest the presence of a more flexible, map-like spatial representation, in which animals could dynamically navigate their environment based on a cognitive understanding of spatial layout, rather than just a sequence of learned actions (Tolman EC, 1948).

Cognitive Maps

"Cognitive maps" are a fundamental concept in the study of cognitive processes such as spatial navigation and memory formation. These maps are described as detailed internal representations that animals form of their surroundings after exploring them. Key characteristics often associated with these maps include the ability of animals to flexibly traverse between any two points and the ability to optimize traversal routes by taking novel shortcuts (Tolman EC, 1948; Mackintosh NJ, 2002).

As highlighted by Mackintosh and later noted by Miller, debates surrounding "cognitive maps" are often less focused on the fundamental question of whether such maps exist in the first place, and more centered on the networks and mechanisms that could be involved in the formation and utilization of such maps (Mackintosh NJ, 2002; Miller N, 2021).

The questions often associated with "cognitive maps" include:

- Whether such a map is formed in a unitary, cohesive manner or in a distributed manner across brain brain regions.
- If they are formed in a distributed manner, how and where are these sparse representations integrated to form a coherent spatial representation?
- If these representations are indeed integrated, how are these distributed networks kept in sync with one another, and how do they interact and influence each other?

In an effort to answer these questions, neuroscientists have utilized highly specialized neurons in the brain, such as place cells in the Hippocampus, which fire in a spatially tuned manner within specific place fields, and grid cells in the Medial Entorhinal Cortex (MEC), which exhibit hexagonal grid-like firing patterns (O'Keefe J and Dostrovsky J, 1971; Hafting et al., 2005).

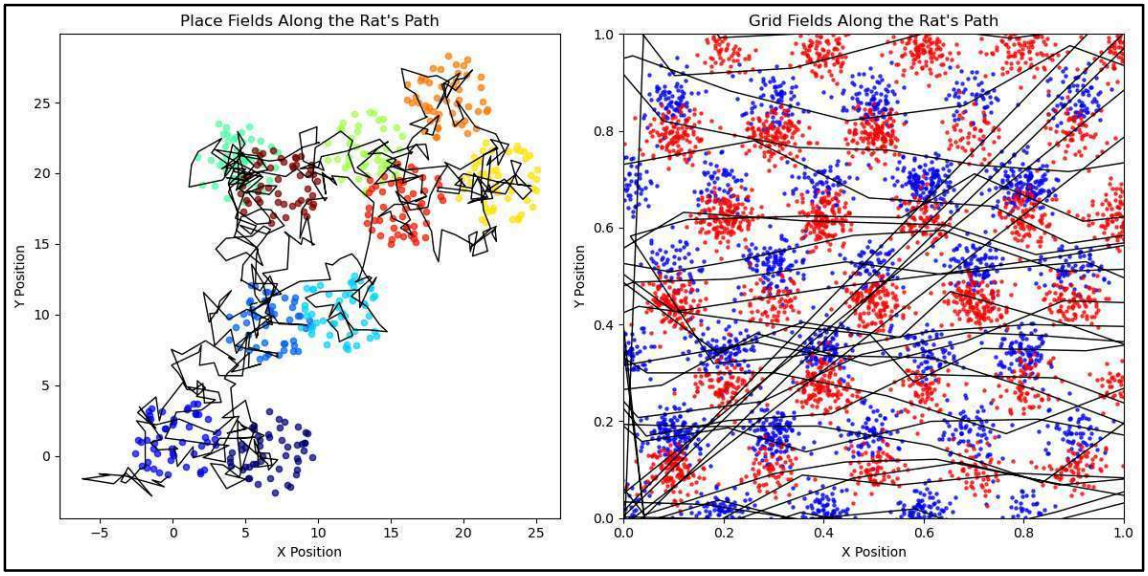

Figure 1. Place cells and Grid cells. (A) An illustration of Place cells, each represented by neural spikes of different color, showing their distinct, localized spatial firing fields as the rat traverses the arena (black line). (B) An illustration of 2 Grid cells, with blue and red circles marking the spikes within their hexagonal firing fields.

Other spatially tuned neurons contribute to cognitive mapping as well. Allocentric Border cells in the MEC spike near environmental boundaries, while Egocentric Border cells (EBCs) in the retrosplenial cortex (RSC) fire when the animal is near a boundary but only when facing a specific direction relative to it (Solstad et al., 2008; Alexander et al., 2020). In addition, a wide array of other specialized cells have been implicated in the formation of cognitive maps.

These specialized cells provide a compelling framework for understanding spatial representation. However, in recent years, there has been an overemphasis on discovering such highly specialized cells, which we believe often leads to hyperfixation on a subset of neuronal data from specific brain regions. That is, such overemphasis leads us to exclude a whole set of neurons (oftentimes a majority of neurons) that don't tick off overly specific boxes of requirements. We end up remembering the MEC for its grid cells and the CA1 region of the hippocampus for its place cells, while ignoring the vast diversity of neurons present in the population as a whole.

We argue that this microscopic approach to neuroscience can be problematic, as it shifts attention to the behavior of individual cells rather than how populations of neurons within circuits interact to form higher-level representations of the environment. This approach can also obscure the broader functional connectivity that integrates information across multiple brain regions, which is essential for generating cohesive cognitive maps.

The Retrosplenial Cortex (RSC)

The retrosplenial cortex (RSC) in both primates and rodents corresponds to Brodmann's areas 29 and 30 and functions as a key association area. Due to its central location, it maintains extensive afferent and efferent connections with various brain regions. Notably, the RSC serves as a major relayer of the hippocampal formation's output to other neocortical areas involved in spatial cognition (Wyss and van Groen, 1992). RSC neurons encode a variety of navigation-related signals, including head direction, positional information, and conjunctive representations of both allocentric (viewpoint-invariant, meaning spatial information is encoded relative to the environment, independent of the animal's position) and egocentric (from the animal's perspective, based on the animal's own position and orientation) cues (van Wijngaarden et al., 2020; Alexander et al., 2020).

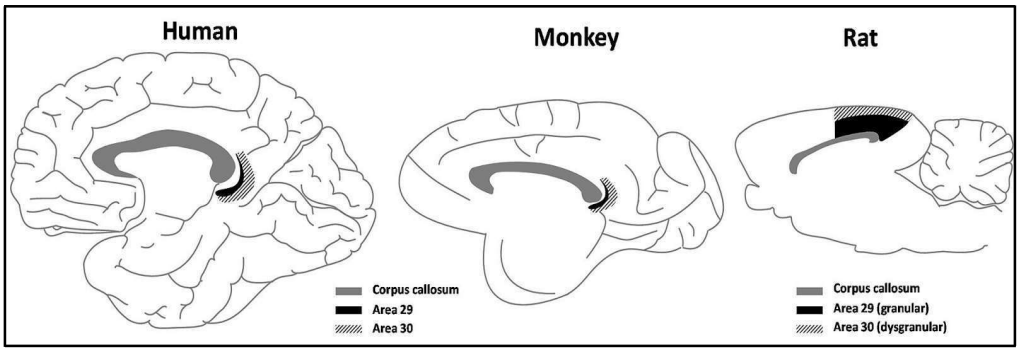

Figure 2. An Illustration of the RSC. The RSC as seen in the midsagittal-view. Source: Jeffery, 2017.

Although the RSC has emerged as a critical structure in spatial processing, its precise role when it comes to spatial coding remains underappreciated. Unlike the hippocampus (HPC) and medial entorhinal cortex (MEC), which exhibit distinct spatially tuned activity, the RSC acts as an integrative hub and hence represents space in a complex manner. Its connectivity with neocortical, hippocampal, parahippocampal, and thalamic regions (Sporns et al., 2007; Sugar et al., 2011; van Groen and Wyss, 1990, 2003; Wyss and van Groen, 1992) implicates its role in processing essential perceptions related to direction, location, landmarks, and navigation, as detailed in the schematic diagram of RSC's connectivity (**Figure 3**) (Alexander and Nitz, 2017; Chang et al., 2020; Chinzorig et al., 2020). Understanding the RSC's exact manner of spatial coding is crucial for bridging the gap between localized neural representations and large-scale cognitive mapping.

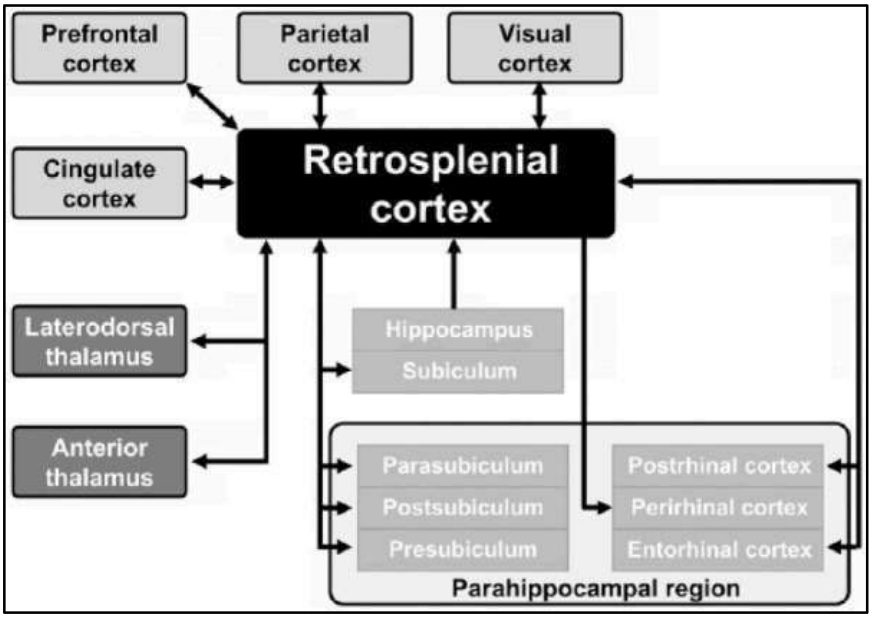

Figure 3. A schematic diagram illustrating the functional connectivity of the RSC. This image illustrates the functional role of the RSC in connecting and integrating information from various spatially relevant regions like the hippocampus, thalamus and the entorhinal cortex. Source: Mitchell, et al., 2018.

Given its role in integrating spatial information, damage to the RSC can result in significant impairments in navigation and memory. Patients with RSC lesions often experience difficulties with spatial orientation, especially with tasks that involve switching between self-centered and viewpoint-invariant representations of space (Maguire, 2001; Vann et al., 2009). Takahashi et al. (1997) described a patient with retrosplenial damage who had trouble drawing a map of a familiar street, despite being able to recognize individual landmarks. This indicates that while landmark recognition remained intact, the ability to integrate spatial relationships like distances between these landmarks to form a coherent cognitive map was impaired. In a similar vein, rodent studies have demonstrated that retrosplenial damage disrupts the ability to use landmarks for navigating (Cooper et al., 2001; Mitchell et al., 2018). These findings suggest that the RSC is not only essential for encoding spatial representations but also for transforming them into a unified cognitive map that supports flexible navigation. A deeper understanding of how the RSC encodes spatial information could offer valuable insights into these disruptions and their wider implications for spatial cognition and memory-related disorders.

The role played by the Medial Entorhinal Cortex (MEC)

Recent findings both experimental and in models, suggest that the RSC may function as a junction for transforming between viewpoint-invariant and body-centered perspectives (Bicanski and Burgess, 2018; Byrne et al., 2007; Clark et al., 2018). This is supported by the RSC's involvement in egocentric boundary/border coding (Alexander et al., 2020) and its bi-directional connections with the MEC, which encodes spatial and directional information in an allocentric manner. These findings point to the possibility that the RSC integrates allocentric position and head-direction signals and then carried out a coordinate transformation from a viewpoint-invariant (allocentric) perspective to a body-centered (egocentric) perspective.

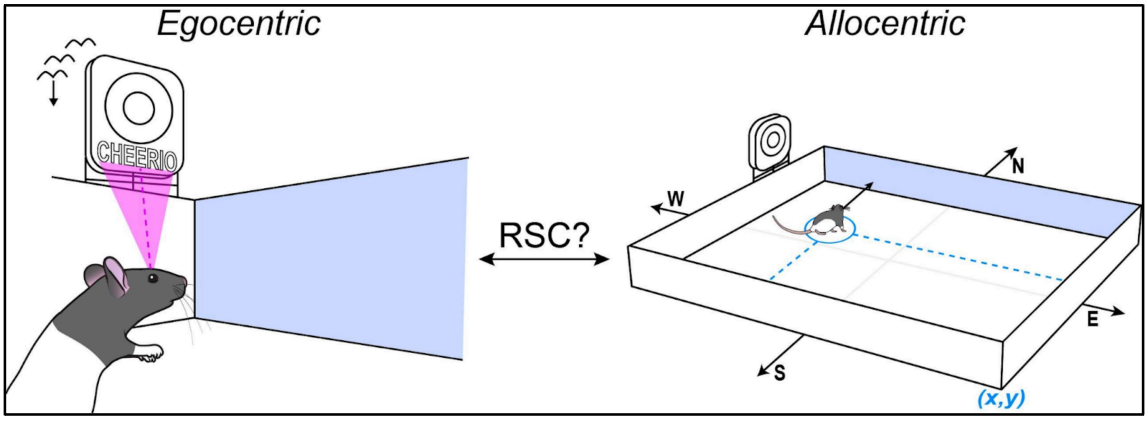

Figure 4. An illustration of RSC's potential role in coordinate-transformation (Alexander AS, 2020). The left panel depicts an egocentric (self-centered) perspective, where spatial information is encoded relative to the rat's own position and orientation (e.g., 'the cheerio sign is to my left'). The right panel represents an allocentric (world-centered) perspective, where spatial locations are encoded independently of the rat's position, based on external reference points (e.g., 'the cheerio sign is in the north-west corner'). The RSC is thought to contribute to switching between these coordinate frames. Interviewed by Sainsbury Wellcome Centre.

Previous studies on the MEC-RSC circuit hierarchy have demonstrated that disrupting MEC signals to the RSC impaired the firing activity of the egocentric border cells in the RSC, whereas inhibiting RSC inputs to the MEC does not alter border cell tuning in the MEC (van Wijngaarden et al., 2020). Based on these findings, we hypothesized that spatial representations and transformations in the RSC depend on spatial information received via the MEC pathway.

Notably, this manipulation did not impact the firing activity of the viewpoint-invariant (allocentric) Head-direction neurons present in the RSC, suggesting that head-direction signals in the RSC likely originate from a different source, such as the anterodorsal thalamic nucleus (ADTN) (Mitchell et al., 2018).

Spatial representation in the Retrosplenial Cortex (RSC)

Although the RSC contains cells tuned to spatial boundaries (i.e., egocentric-border cells), these alone are insufficient to map the entire arena an animal can explore. While border cells provide important spatial cues, indicating the arena's maximum boundaries and motion-impeding features such as walls, they do not map continuous variables like spatial location at any given moment.

Regarding the presence of Place cells, which would be ideal for accurately mapping unique locations, there is no clear consensus on whether such cells exist in the RSC, particularly in 2D arenas. A study by Mao et al. (2017) reported place cell-like activity in the RSC using cellular Ca^{2+} imaging in mice during a head-fixed 1D locomotion assay. In this setup, mice walked on treadmills and received tactile cues in the form of periodic bumps, which indicated progression towards the reward. While this study provides important insights, the setup has limitations, particularly the lower temporal resolution of calcium imaging compared to electrophysiological methods, as well as the involvement of additional tactile cues. These cues complicate the interpretation of whether the neural activity is truly reflecting spatial location, responding to the tactile cues, or representing task progression toward the goal. Since then, no follow-up studies have definitively confirmed the presence of place cells in the RSC, especially in 2D environments.

Additionally, recent studies suggest that RSC neurons may exhibit mixed-selectivity, encoding both spatial and non-spatial features, such as sensory inputs or task-related variables (Kira et al., 2023; Fischer et al., 2020). Mixed selectivity is often defined as the ability of neurons to encode various independent variables conjunctively, enabling neuronal ensembles to encode a wide range of information flexibly (Rigotti et al., 2013; Tye et al., 2024). This suggests that spatial information in the RSC may not be represented by individual place cells, but rather by a complex ensemble of cells that respond to multiple contextual factors.

These findings led us to hypothesize that the RSC could represent spatial information at the ensemble level, utilizing dynamical population coding, rather than at the single-cell level.

Neural Manifolds

The timeseries activity of an ensemble of neurons, which encodes a mixture of information at any given moment, can be thought of as a trajectory along a high-dimensional space, where each dimension represents the activity of a neuron over time. Due to the intrinsic nature of neuronal circuits, there are high levels of correlation and redundancy between the activity of individual neurons. As a result, instead of the activity following a complex path in a high-dimensional space, it often follows a simpler, lower-dimensional path. These simpler paths are called neural manifolds. In simple terms, neural manifolds can be thought of as a

low-dimensional projection of information encoded at the level of a neuronal population (Nakai et al., 2024; Langdon et al., 2023).

Neural manifolds are not just abstract mathematical constructs, they also have biological counterparts. A striking example is the elegant *Drosophila melanogaster* navigation system, where the low-dimensional ring manifold derived from the population activity of the head direction (HD) system precisely corresponds to the ellipsoid body, a neuropil structure in the fly HD system (Seelig & Jayaraman, 2015). While it is true that such simple manifold-circuit convergence has not been widely observed in other systems, neural manifolds remain valuable for uncovering general principles of population dynamics in neural ensembles and the information they encode. For instance, Nakai et al. (2024) investigated how population activity in the subiculum and hippocampus differentially encoded navigation-relevant variables, such as speed, position, and future path information, using neural manifolds. This analysis helped clarify the distinct roles of the subiculum and hippocampus in navigation processing.

Insights from these studies prompted us to explore the use of neural manifold analysis to understand how the Retrosplenial Cortex encodes spatial information at the population level.

Scope of our Study: What We Will and Won't Address

Building on the insights from previous studies and the theoretical framework outlined above, we aim to focus on specific aspects of how the Retrosplenial Cortex (RSC) encodes spatial information. Our study will address the following questions:

1. **Single-Neuron Encoding of Allocentric Variables**: As a starting point, we will explore how allocentric variables, such as view-point invariant position and head direction, are represented at the single-neuron level in the RSC to validate findings from previous studies.
2. **Ensemble Encoding of Spatial Information**: We will next investigate how well the RSC represents spatial information at the ensemble level, while also exploring the role played by the Medial Entorhinal Cortex (MEC) in this.
3. **Neural Manifold Analysis**: Finally, we will apply neural manifold analysis to the population activity in order to gain finer insights into how the dynamics of the RSC ensemble encode positional information.

While we will be addressing these specific questions, there are several areas beyond the scope of our study:

1. **Allocentric-Egocentric Transformation**: We will not explore the role of the RSC in transforming between allocentric and egocentric representations of space. We will solely focus on the representation of allocentric space.
2. **Task-Dependent Ensemble Dynamics**: We will not address how RSC ensemble dynamics may change depending on the task being performed. Our conclusions will focus solely on the representation of spatial information, excluding goal-relevant variables.
3. **Nature of the RSC Manifold**: We will not speculate on the exact nature of the RSC manifold. Our goal is to understand how the RSC encodes spatial information at the ensemble level using neural manifolds, rather than to hypothesize about the specific structure of the manifold.

While these questions are undoubtedly important and intriguing, they extend beyond the scope of the present study. Addressing them fully would require further investigation. Our goal here is to lay a foundation for future exploration in these directions. These future steps will be crucial for deepening our understanding of the RSC, not just in representing spatial information but also in transforming it between coordinate systems and integrating it into a cohesive cognitive map.

Chapter 2: Materials and Methods

All electrophysiological recordings and experimental data collection were performed by Dr. Joeri B.G. van Wijngaarden and Dr. Susanne Babl, as outlined in van Wijngaarden et al., 2020. While I engaged in certain experimental tasks for expanding the work, the analysis presented in this thesis is based exclusively on a subset of Dr. van Wijngaarden's rich dataset, specifically utilized to address the research questions explored in this work. The detailed experimental protocols are described in *van Wijngaarden et al., 2020*.

Animal Subjects

This study originally involved 19 male Long-Evans rats (3–5 months old, 400–550 g), housed individually in transparent Plexiglass cages under a reversed 12-hour light-dark cycle, with behavioral testing during the dark phase. For our analyses, we used a subset of four rats from this group, maintaining the same age and weight range. These animals were mildly food-restricted (kept at 85–90% of their free-feeding body weight) but had unrestricted access to water.

For the recording experiments, tetrodes were implanted unilaterally in the RSC of all four rats, with two rats having electrodes placed in the left hemisphere and two in the right hemisphere. Additionally, one of the four rats received bilateral injections of an AAV encoding inhibitory DREADDs in the MEC, along with the aforementioned tetrode drive in the right hemisphere of the RSC.

Surgical Drive Implantation and Viral Vector Injections

For tetrode recordings, all four rats were unilaterally implanted with tetrodes in the retrosplenial cortex (RSC), with two rats receiving implants in the left hemisphere and two in the right. One of these four rats also received bilateral injections of an AAV encoding inhibitory DREADDs in the medial entorhinal cortex (MEC) alongside a tetrode drive in the right RSC.

Surgeries were conducted under isoflurane anesthesia (5% induction, 0.5–2% maintenance), with analgesia provided via subcutaneous Buprenorphine and local application of

Ropivacaine. Rats were placed in a stereotaxic frame, and after aligning the skull, craniotomies were made for microdrive implantation. The microdrive, containing 28 independently adjustable tetrodes, was secured with dental cement, and ground screws were placed above the cerebellum. Tetrodes were positioned along the anteroposterior axis of the RSC at a 25° coronal angle to avoid the superior sagittal sinus.

For DREADDs experiments, an AAV8-hSyn-hM4Di-mCherry virus was injected bilaterally into the MEC using a NanoFil syringe at two depths per site (2.5 mm and 3.5 mm from the cortical surface). Injections were controlled via a Micro4 microsyringe pump, with the needle left in place for 10 minutes post-injection. A small microdrive was implanted near the injection site to assess the effects of the manipulation. Virus injection and electrode implantation occurred in the same surgery, with recordings beginning at least three weeks later to allow for viral expression.

Post-surgical care included analgesics (Metacam) and antibiotics (Baytril) for at least five days.

Behavior and Movement Tracking

Data was collected for 30–120 minutes per day as rats foraged for chocolate cereal in a 100 × 100 cm open-field arena. Each session included 10–15 minutes of free exploration, followed by 5-minute rest periods on a pedestal. The arena, elevated 50 cm above the ground, had three black walls and one white wall, consistently oriented for all animals. Except for darkness experiments, no curtains enclosed the arena. The setup was cleaned with 70% ethanol between sessions to eliminate odors.

For DREADDs experiments, animals received subcutaneous injections of agonist-21 (3.52 mg/mL) after the first session. Recordings resumed after a 30-minute waiting period to allow the drug to take effect.

Animal position and head direction were tracked at 25 Hz using two LEDs under dim light. During darkness sessions, an infra-red OptiTrack system with six ceiling-mounted cameras (850 nm IR) tracked reflective markers on the headstage. All visible light sources were covered, and the arena was enclosed with a black curtain. A dim room lamp was used between trials but turned off 10 seconds before recording. The experimenter remained stationary and silent, scattering food pellets throughout the session.

Spike sorting and Cell Classification

Initial data processing and primary analyses were conducted in MATLAB (MathWorks). Neural signals were recorded and amplified using two 64-channel RHD2164 headstages (Intan Technologies) with an OpenEphys system, sampling at 15 kHz. Spikes were detected by band-pass filtering the local field potential (0.6–6 kHz) and processed with the 'Kilosort' algorithm (Pachitariu et al., 2016) to classify them into clusters based on waveform properties. Clusters were manually refined using autocorrelograms and principal component analysis to isolate single units, excluding multi-unit activity or noise. Tetrodes were adjusted by at least 80 μm between sessions to record new neurons.

RSC Border Cell Classification

We used a template-matching approach to identify RSC neurons as border cells using Earth Mover's Distance (EMD) (Hitchcock, 1941; Rubner et al., 1998). The animal's spatial occupancy was divided into 4 × 4 cm bins, with the firing rate for each bin calculated by dividing the total number of spikes that occurred in each bin by the time spent in said bin. The rate map was smoothed with a 2D Gaussian filter (1-bin width) and normalized to form a probability distribution with unit weight. We then computed the EMD relative to a ‘boundary template’ using the fastEMD algorithm in MATLAB (Koslicki, 2015; Pele et al., 2008; Pele and Werman, 2009). The boundary template was a 25 × 25 matrix with 0 assigned to all bins except the outermost ring, which was assigned a value equal to 1, smoothed with the same Gaussian kernel, and normalized to unit weight. Additional templates were generated to incorporate behavioral manipulations, with extra weight at object/wall locations. The EMD distance between the actual rate map and the template represents the minimum cost it would take to transform one distribution into the other, with a score normalized from 0 (a perfect border cell) to 1 (maximum dissimilarity possible) (Grossberger et al., 2018).

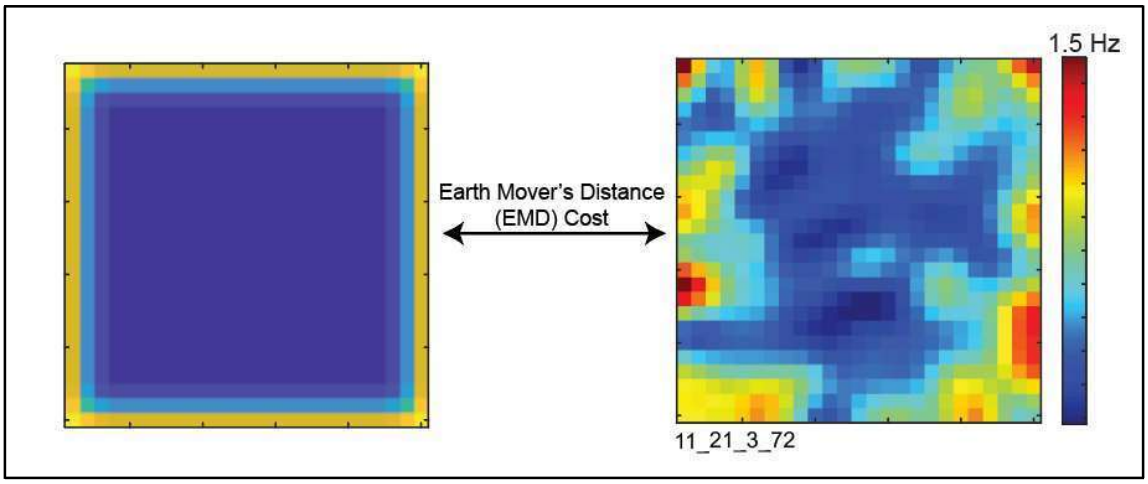

Figure 5. An Illustration of the Earth Mover's Distance (EMD) Score. It quantifies the minimum cost required to transform the actual rate map into the template (or vice versa).

To check if a neuron's rate map resembled the boundary template, we created a null distribution for comparison. This involved 200 iterations of a shuffling procedure, where spike trains were randomly time-shifted along the animal's trajectory (by at least 4 seconds and less than the total trial length) and wrapped around if they exceeded the session duration. EMD values were recalculated for these shuffled rate maps. A neuron was classified as a border cell if its EMD score was below the 5th percentile of the null distribution in all regular sessions and had an average firing rate of at least 0.5 Hz. This classification was applied only to regular sessions in manipulation experiments.

Head-direction Cell Classification

The rat's head direction was determined using the positions of two LEDs fixed to the acquisition Intan headstage attached to the retrodrive, with an offset correction for the LEDs' placement relative to the animal's true head direction. For each cell, the mean vector length (MVL) was calculated by computing the circular mean from a vector containing the animal's head direction at spike times. A cell was classified as a head-direction cell if its MVL exceeded the 95th percentile of a null distribution, generated through 200 iterations of temporal shuffling with randomly time-shifted spike trains.

Bayesian Decoding Analysis

For decoding the animal's position from neural activity, we first assigned spike trains to discrete time bins (250 ms) and the animal's position to corresponding spatial bins (5 cm), creating a time series of spiking activity and positional occupancy. The spatial environment was divided into a grid of equally spaced bins, each 5 cm in size. For each neuron, the mean firing rate within each bin was computed and a gaussian filter (250ms) was applied to estimate the smoothed spatial tuning curve.

To implement Bayesian position (and later, head-direction) decoding, we employed a standard probabilistic approach where the probability of the animal being in a given position at time 't', given the observed neural activity, was computed using Bayes' theorem. The likelihood function was estimated based on the tuning curves of the recorded neurons, under the assumption of independent Poisson firing. A uniform prior was used, meaning the probability of the rat being in any location was assumed to be equal, or that the rat explored the arena uniformly without any preferred areas. For each time bin, the most probable position was inferred by computing the posterior probability across spatial bins and selecting the maximum a posteriori (MAP) estimate (Davidson et al., 2009).

To evaluate decoding accuracy, this process was repeated for all time bins. Decoding performance was quantified by computing the accuracy of the decoded positions relative to the actual positions. The coefficient of determination (R2) was used to assess the goodness of fit between the decoded and true positions.

Linear Discriminant Analysis

For decoding the animal's position from neural activity, we assigned spike trains to discrete time bins (250 ms) and the animal's position to corresponding spatial bins (5 cm), creating a time series of spiking activity and positional occupancy. The spatial environment was divided into a grid of equally spaced bins, each 5 cm in size. For each neuron, the mean firing rate within each bin was computed and a gaussian filter of 250ms was applied to estimate the smoothed spatial tuning curve.
To implement position decoding using Linear Discriminant Analysis (LDA), we applied LDA to the spike rate data to infer the animal's position. The spike rates for each time bin were used as features, with the corresponding spatial bin labels as targets. LDA computed the

weights for each neuron's firing rate, which were used to predict the most likely position at each time bin.

To evaluate decoding accuracy, this process was repeated for all time bins. Decoding performance was quantified by computing the accuracy of the decoded positions relative to the actual positions. The coefficient of determination (R^2) was used to assess the goodness of fit between the decoded and true positions.

Neural Manifold Analysis

Neural manifolds provide a lower-dimensional representation of the dynamic activity of a neural ensemble. After performing LDA, we applied Principal Component Analysis (PCA) to the resulting Linear Discriminants (LDs) to obtain the 2D neural manifold. To visualize spatial information encoded within the ensemble, we colored the manifold based on the decoded position of the animal at each time point, assigning a unique color to each (x, y) location. This provided an intuitive representation of how neural activity relates to spatial occupancy.

Population Vector Analysis

To analyze neural population activity, we first binned spike trains into 250 ms time bins and computed the firing rates of neurons. The firing rates were then normalized using Z-scoring. This resulted in an N-dimensional time series, where each time bin contained the firing rates of N neurons.

We sorted these time series based on the rat’s spatial location and computed a mean population vector for each unique location visited. To quantify the consistency of neural representations, we calculated the Euclidean distance between each population vector and the mean vector of its corresponding location. We then compared these distances to those between each population vector and the mean vectors of other locations.

To investigate whether spatial locations are represented in orthogonal subspaces, we examined the dot products between the local variability of population vectors (i.e., deviations from the mean vector within a location) and the differences between mean population vectors of distinct locations. A distribution of these dot products was computed to assess whether

representations of different locations were aligned or orthogonal in population activity space.

Variational Autoencoder (VAE) Analysis

To address the limitations of LDA, we adopted an unsupervised deep learning approach. LDA assumes that data from different classes (e.g., spatial positions) can be separated using linear decision boundaries, which may not hold for neural representations of space. Additionally, LDA assumes homoscedasticity (equal variance across classes), which is often violated in neural activity patterns. To overcome these constraints, we implemented a Variational Autoencoder (VAE), which learns a continuous latent space representation of neural activity without explicit class labels.

The neural time series data, structured as (Neurons × Time bins), was processed using a 10-fold cross-validation approach to ensure model robustness and prevent overfitting. The VAE consists of an encoder-decoder architecture: the encoder maps high-dimensional neural activity into a lower-dimensional latent space, capturing essential structure while enforcing a probabilistic distribution over the latent variables. The decoder then reconstructs the original neural activity from this latent representation, ensuring that the latent space meaningfully captures underlying neural dynamics.

By training the VAE on neural population activity, we obtained a smooth, continuous representation of neural manifolds in the latent space. This allowed us to compare the structure of these manifolds with those derived using PCA on LDA-obtained discriminants. Importantly, the VAE-based approach provided a more flexible and data-driven way to explore neural representations, avoiding constraints imposed by linear separability assumptions. Additionally, because VAE is an unsupervised method, it does not impose constraints on the learning process. As a result, the latent representation (neural manifold) captures a composite structure of the neural activity, whereas the LDA approach explicitly focused on spatial information.

Supervised Variational Autoencoder (S-VAE)

To extend VAE into a supervised framework, we added a supervised loss alongside the Kruskal-Wallis and Mean Squared Error (MSE) losses (Le, L. et al., 2018). The supervised

loss was computed as the MSE between the actual and predicted (x, y) position of the rat. Additionally, we introduced a separate output head where the latent representation projected to a layer predicting the rat's position. All other steps remained unchanged from the original VAE protocol.

Statistical procedures

All statistical tests were two-sided and non-parametric unless stated otherwise. Error bars in all figures represent the standard error of the mean (SEM) unless indicated otherwise.

Chapter 3: Results

A part of this work is published as a paper in eLife.
Full reference:
van Wijngaarden JB, Babl SS, Ito HT. 2020. Entorhinal-retrosplenial circuits for allocentric-egocentric transformation of boundary coding. *eLife*. 9:e57278.

Electrophysiological recordings of neuronal activity in RSC of rats were carried out as they explored a squared open field arena of size 1m-by-1m and foraged for scattered chocolate pellets. All sessions considered here, were from recordings done in darkness, without any objects obstructing the free movement of the rats. The rats were properly habituated to the arena, before carrying out any recordings. The following are the results of our manipulations and analysis from the recorded data.

RSC cells show Border & Head-direction tuning

The retrosplenial cortex (RSC) is known to contain two types of navigation-relevant cells: head direction (HD) cells and border cells (Solstad et al., 2008; Alexander et al., 2020). We began our analysis by examining the composition of our dataset and classifying cells based on whether they encoded boundary information, head direction, or both conjunctively.

Border cells (**Figure 6A**) were classified using the Earth Mover's Distance (EMD) score (Hitchcock, 1941; Rubner et al., 1998), a normalized measure quantifying the cost of transforming a neuron's rate map (heatmap of firing rate across the explored area) into an idealized border cell template, which fires exclusively at the borders (Koslicki, 2015; Pele et al., 2008; Pele and Werman, 2009; Grossberger et al., 2018). To ensure significance, we compared the observed EMD scores to a distribution generated by shuffling spike times randomly, confirming that border coding was not due to chance.

Similarly, HD cells (**Figure 6B**) were identified using the Mean Vector Length (MVL), which quantifies directional bias in spiking activity. A true HD cell should exhibit a strong preference for a specific directional range, with an MVL significantly above chance. We applied the same spike-time shuffling approach to validate significance.

Cells exhibiting both significant border and head-direction tuning were classified as conjunctive cells (**Figure 6C**).

Across nine recording sessions from three animals (**n = 866 neurons**), we identified:

- **114 border cells (~13.2%)**
- **185 head-direction cells (~21.4%)**
- **39 conjunctive cells (~4.5%)**
- **528 non-coding neurons (~61%) (Figure 6D)**

We observe that a majority of the cells (~61%) do not code for either of the allocentric variables.

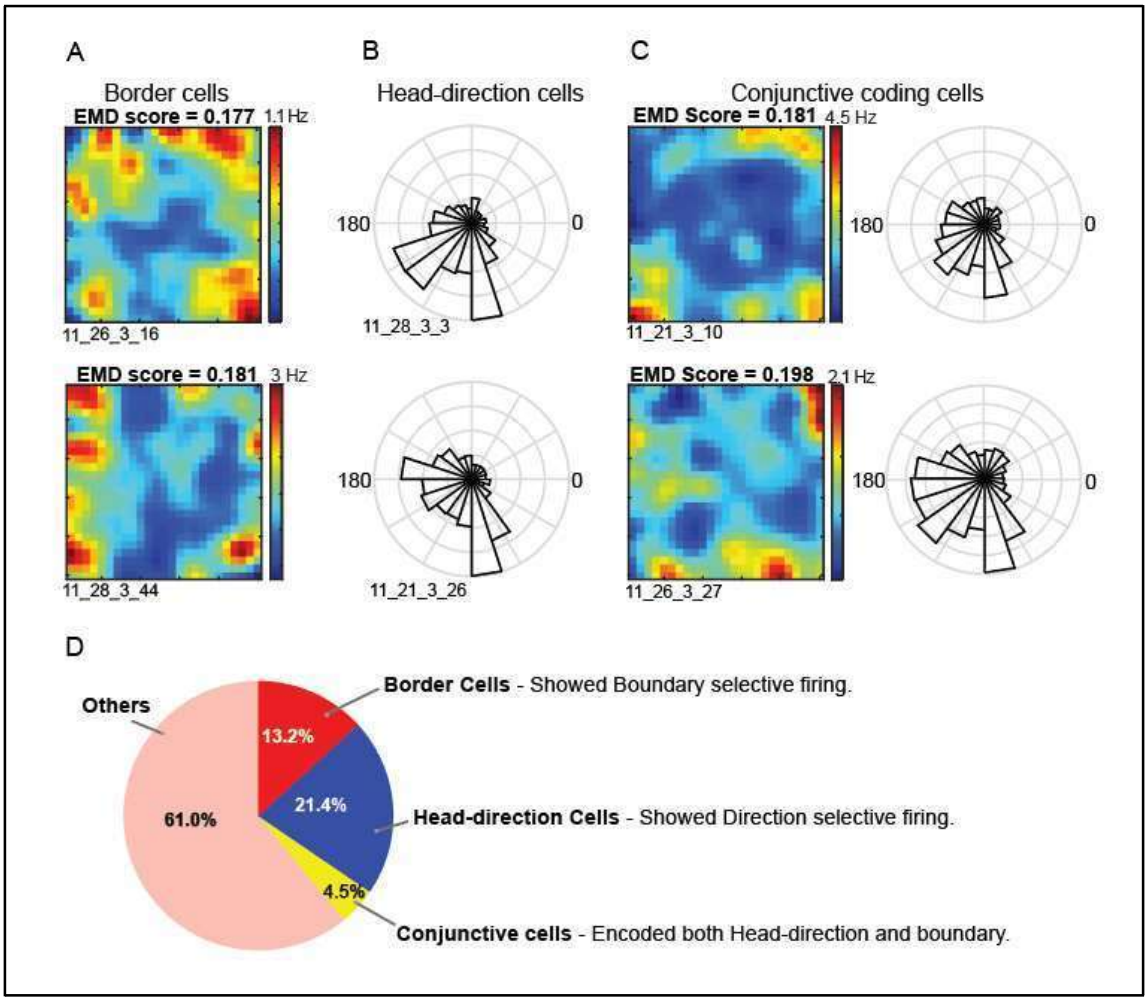

Figure 6. RSC cells exhibit Border and Head-Direction tuning. (A) Example Border cell rate maps with corresponding EMD scores. (B) Example Head-Direction cell polar histograms. (C) Example Conjunctive-coding cells, shown with both rate maps and polar histograms. (D) Distribution of cell types in the RSC.

RSC does not contain spatially-tuned cells

To explore whether positional information could be decoded from the spiking of individual RSC neurons, we applied Bayesian decoding analysis. This approach leverages each neuron's ratemap as a probability distribution of spiking activity given the location.

Using Bayes' theorem, we computed the probability of the animal being in a given position at each timepoint, based on neural activity. The likelihood was estimated from the neurons' tuning curves, assuming Poisson firing, with a uniform prior over spatial locations. The most probable position was then inferred by selecting the maximum a posteriori (MAP) estimate (Davidson et al., 2009). We similarly decoded head-direction from the neural time series.

Across **9 sessions** recorded from **3 rats**, the decoding yielded **poor coefficients of determination (R^2) values**, indicating weak predictive power (**Figure 7A**). The R^2 values were mostly **negative**, suggesting that the decoded positions were, on average, less accurate than simply selecting **the median location** across time. We obtained the following R^2 (Mean ± S.D.):

- **R^2_HD** = **-**0.3657 ± 0.2702
- **R^2_X** = **-**0.3428 ± 0.3922
- **R^2_Y** = **-**0.6547 ± 0.4776

Given the low R^2 values obtained from single-cell Bayesian decoding, we next explored whether spatial information might be encoded at the ensemble level.

RSC codes space at an ensemble level

To explore whether positional information could be decoded from RSC neural activity, we applied Linear Discriminant Analysis (LDA) to classify the animal's position based on its spiking patterns.

LDA was trained on the dataset, using binned spike rates as input features and spatial bin labels as targets. The decoder computed discriminant weights for each neuron's firing rate, allowing for the classification of the most likely position at each time bin. We similarly decoded head-direction from the neural time series.

Decoding performance was assessed by comparing predicted positions to actual positions across time, quantified using the coefficient of determination (R^2). Across **9 sessions** recorded from **3 rats**, the decoding yielded **good coefficients of determination (R^2) values**, indicating great predictive power (**Figure 7A**). We obtained the following R^2 (Mean ± S.D.):

- **R^2_HD** = 0.3835 ± 0.1546
- **R^2_X** = 0.8386 ± 0.0578
- **R^2_Y** = 0.8433 ± 0.0729

We also examined whether the apparent high decodability of position was uniform across the environment or if it was influenced by the rat's location—such as near the borders—due to the presence of border cells in the ensemble. To test this, we analyzed decoding accuracy across different spatial regions of the arena. We found no systematic bias toward boundaries or other arena features, as decoding accuracy appeared randomly distributed across space (**Figure 7B**).

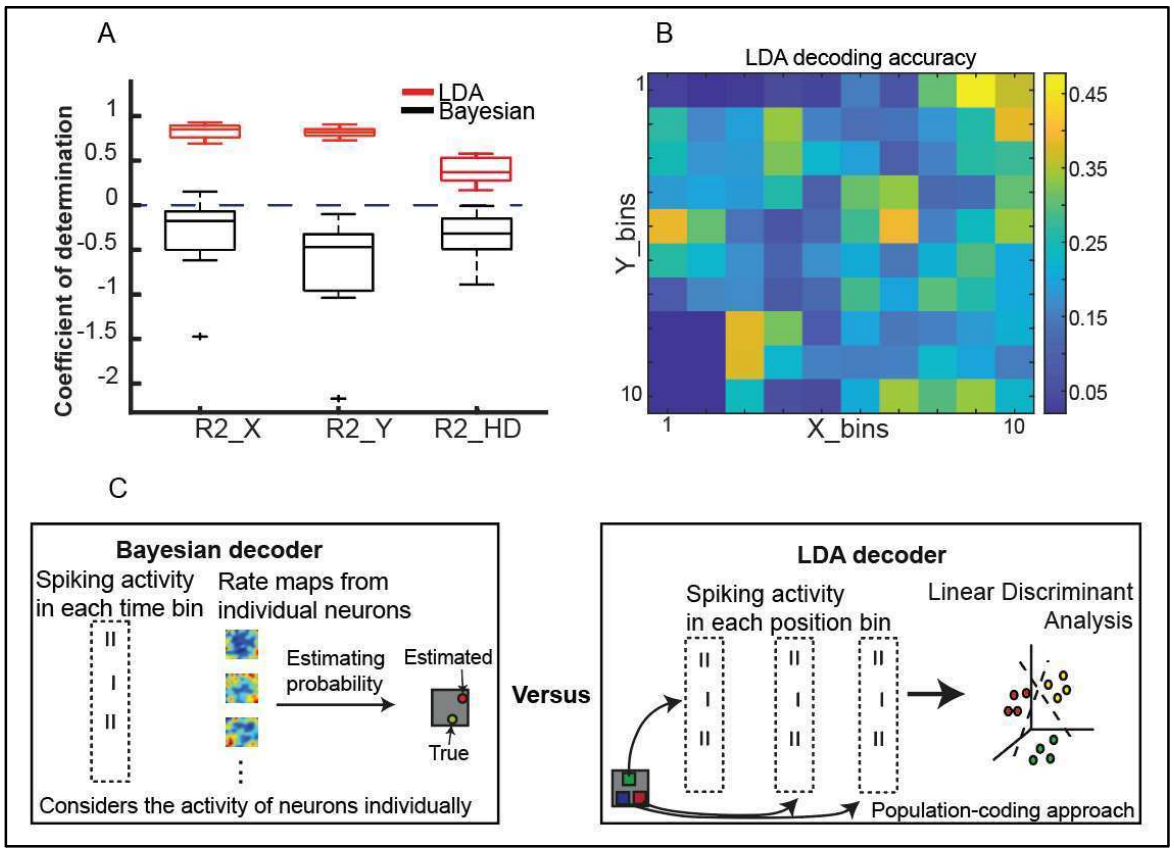

Figure 7. High-accuracy position decoding using LDA. (A) Boxplots comparing R^2 values obtained from Bayesian decoding and LDA-based decoding. (B) Example of spatial-dependence of decoding accuracy for a given session. (C) Schematic comparison of the underlying principles of Bayesian and LDA decoders.

Impact of MEC silencing on spatial coding in RSC

We investigated the impact of AAV-mediated silencing of the Medial Entorhinal Cortex (MEC) on the RSC's ability to encode spatial information, including head-directional coding. The aforementioned LDA pipeline was used for decoding positional information. The decoding performance was assessed by comparing predicted positions to actual positions over time, quantified by the coefficient of determination (R^2).

Across **eight** sessions recorded from **one rat**, **four** of which involved AAV-mediated silencing of the MEC, there was a directional trend toward a reduction in R^2, but this was **not statistically significant** (**Figure 8**). The following R^2 values (Mean ± SEM) were observed:

	R2_HD	R2_X	R2_Y
Pre-AAV	0.7503 ± 0.0341	0.8861 ± 0.0152	0.8829 ± 0.0224
Post-AAV	0.7482 ± 0.0311	0.8521 ± 0.0348	0.8730 ± 0.0127

Table 1. R^2 values (Mean ± SEM) Pre- and Post-AAV Injections.

Given the limited number of data points, we performed a Wilcoxon Signed-Rank test for the three paired comparisons (pre-AAV vs. post-AAV) for each coding type (head-direction, X-position, and Y-position):

- p_value(HD) = 0.875
- p_value(X) = 0.625
- p_value(Y) = 1

In all cases, there was no significant reduction in R^2 values post-AAV silencing.

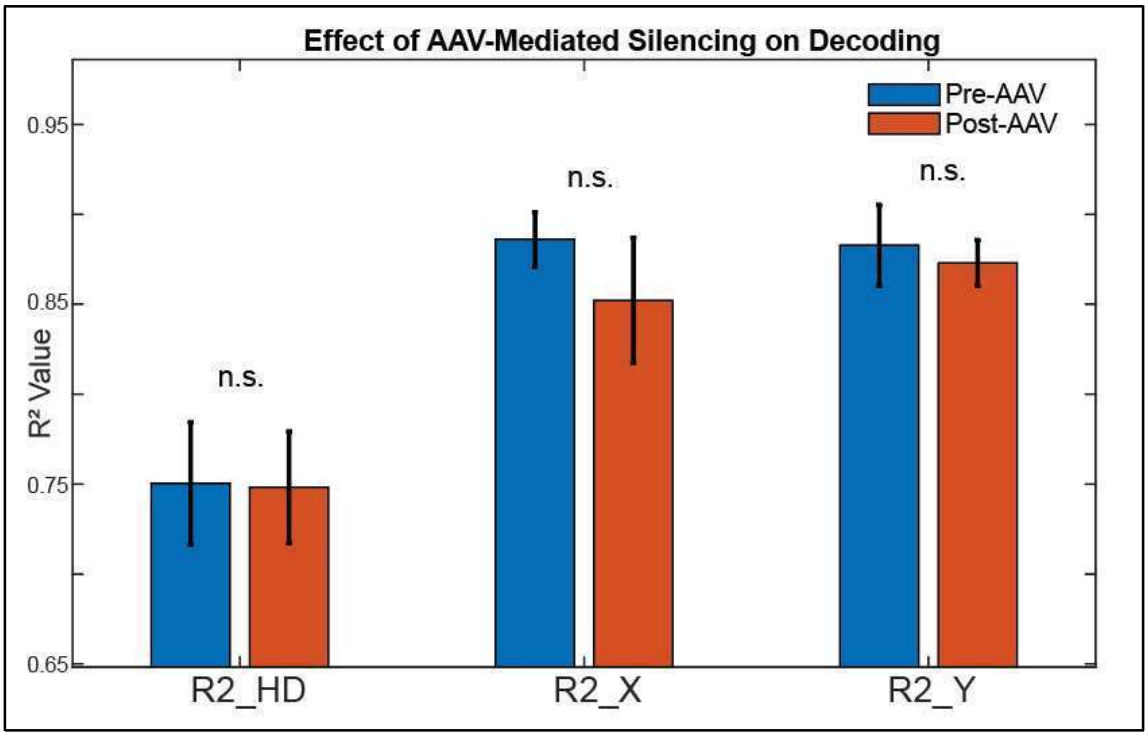

Figure 8. Effects of silencing MEC. Silencing of MEC leads to a reduction in spatial and head-directional coding, but this reduction is not statistically significant.

RSC Neural manifolds reflect Euclidean space

We observed that the neural manifolds represented the neural population activity in a structured manner, with smooth transitions and distinct boundaries (**Figure 9**). To further refine this representation, we calculated the center of mass (COM) of individual sub-manifolds. Each sub-manifold consisted of all neural activity points corresponding to the same real-world location—i.e., whenever the rat visited a specific location multiple times, the resulting manifold points were grouped together. By computing the COM for each sub-manifold, we condensed these multiple points into a single representative point per location. This reduced variability while preserving the overall structure of the neural representation, allowing for a clearer analysis of how spatial information is encoded within the manifold.

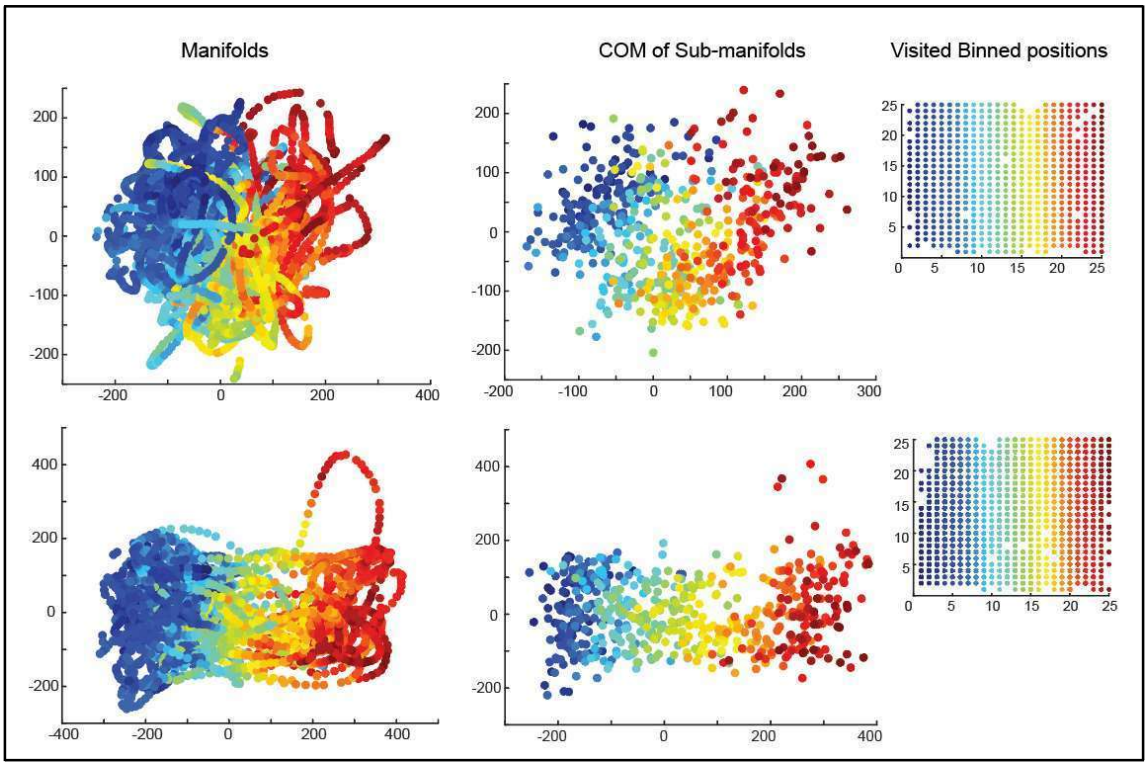

Figure 9. LDA Neural Manifolds. Example neural manifolds (left) illustrating the low-dimensional representation of neural activity. The corresponding simplified manifolds (center) show the center of mass (COM) of each sub-manifold, where each real-world location is represented by a single point. The right panel displays the real-world position bins visited by the rat, with each position bin colored using a unique color, to serve as a reference for the previous figures.

To further assess how well real-world distances are preserved in the neural manifolds, we calculated the pairwise distances between the COMs and real-world positions in the arena (**Figure 10A**). This analysis was based on data from eight recording sessions, recorded from a **single rat**, **four** of which involved AAV-mediated silencing of the MEC.

Real-world distances were strongly correlated with manifold distances, with locations farther apart in the real world also being represented farther apart in the manifold, and vice versa for closer locations. However, in sessions with AAV-mediated MEC silencing, we observed a clear reduction in resolution, as real-world locations were represented closer together in the manifold compared to control sessions (**Figure 10B**).

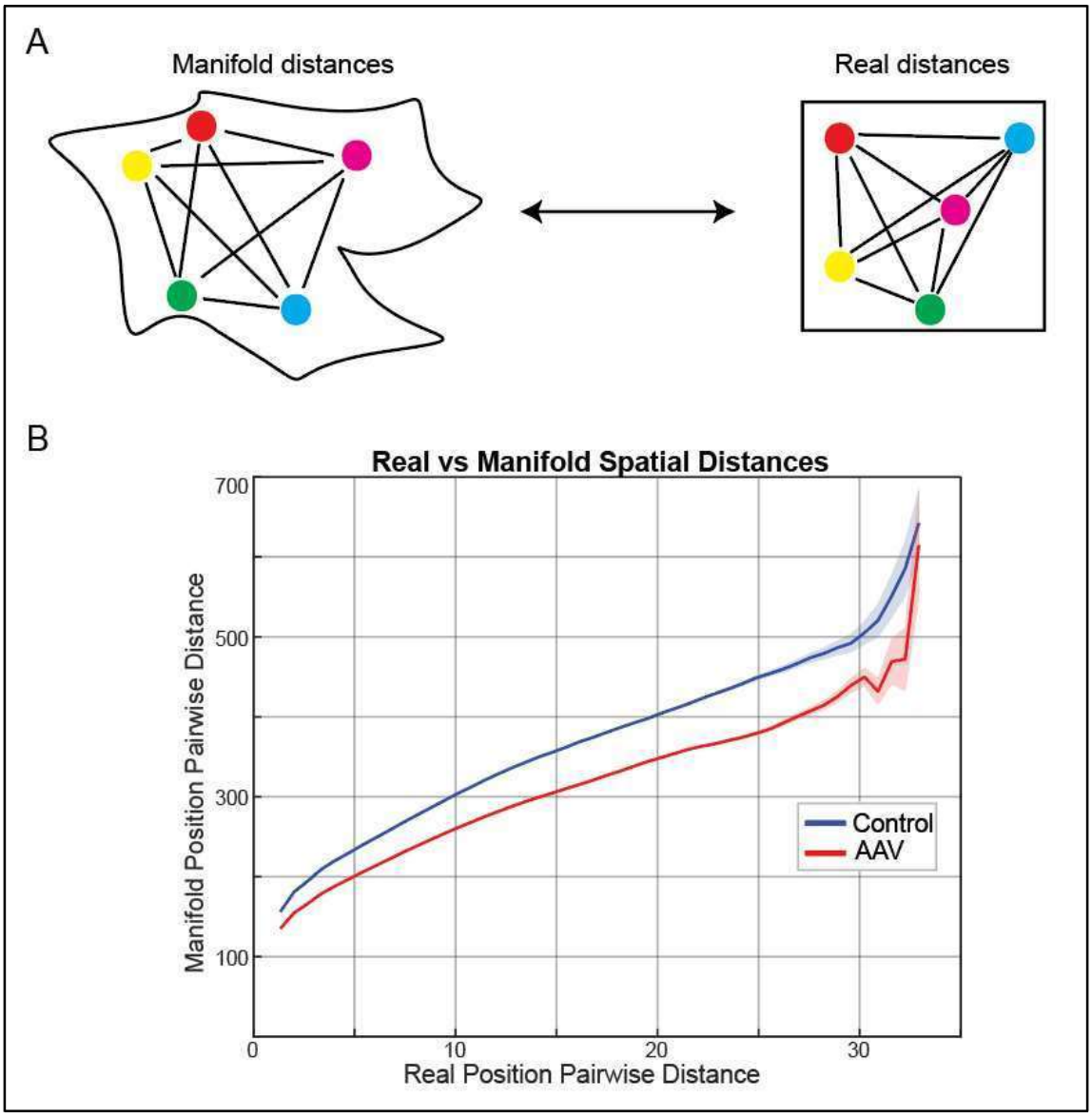

Figure 10. Manifold distances vs real-world distances. (A) Schematic representation of the analysis: Comparing the neural manifold representations of locations with their real-world counterparts, where unique colors represent distinct real-world locations. (B) Pairwise distance comparison between real-world and manifold representations under control and AAV conditions (Median ± SEM).

To assess the significance of this reduction in resolution, we carried out a Wilcoxon signed-rank test on the control vs AAV conditions. We found the reduction of manifold distances to be highly significant ($p = 0.0000$, $h = 1.00$).

RSC Population Activity Forms Distinct and Orthogonal Spatial Representations

To further explore the population coding of spatial representation in the RSC, we employed a population vector approach, using the N-dimensional firing rate activity (for N neurons) to examine the high-dimensional representation of space.

We found that the spiking activity corresponding to different locations clustered into separate subspaces, meaning that the distance between spiking activity and the mean activity for a given location was smaller than the distance between mean activity vectors of different locations (**Figure 11 A, B**).

Furthermore, these subspaces exhibited orthogonality, meaning that the variability in population activity within one location was largely perpendicular to the variability in activity representing another location. This suggests that spatial coding in the RSC maintains a structured and independent representation of different locations (**Figure 11C**).

The clustering of spatial subspaces and their orthogonality were preserved **even after excluding Border cells and Head Direction cells** from the dataset. Notably, these characteristics remained consistent across all **17 sessions**, spanning **four rats**, even in the sessions involving AAV-mediated MEC silencing.

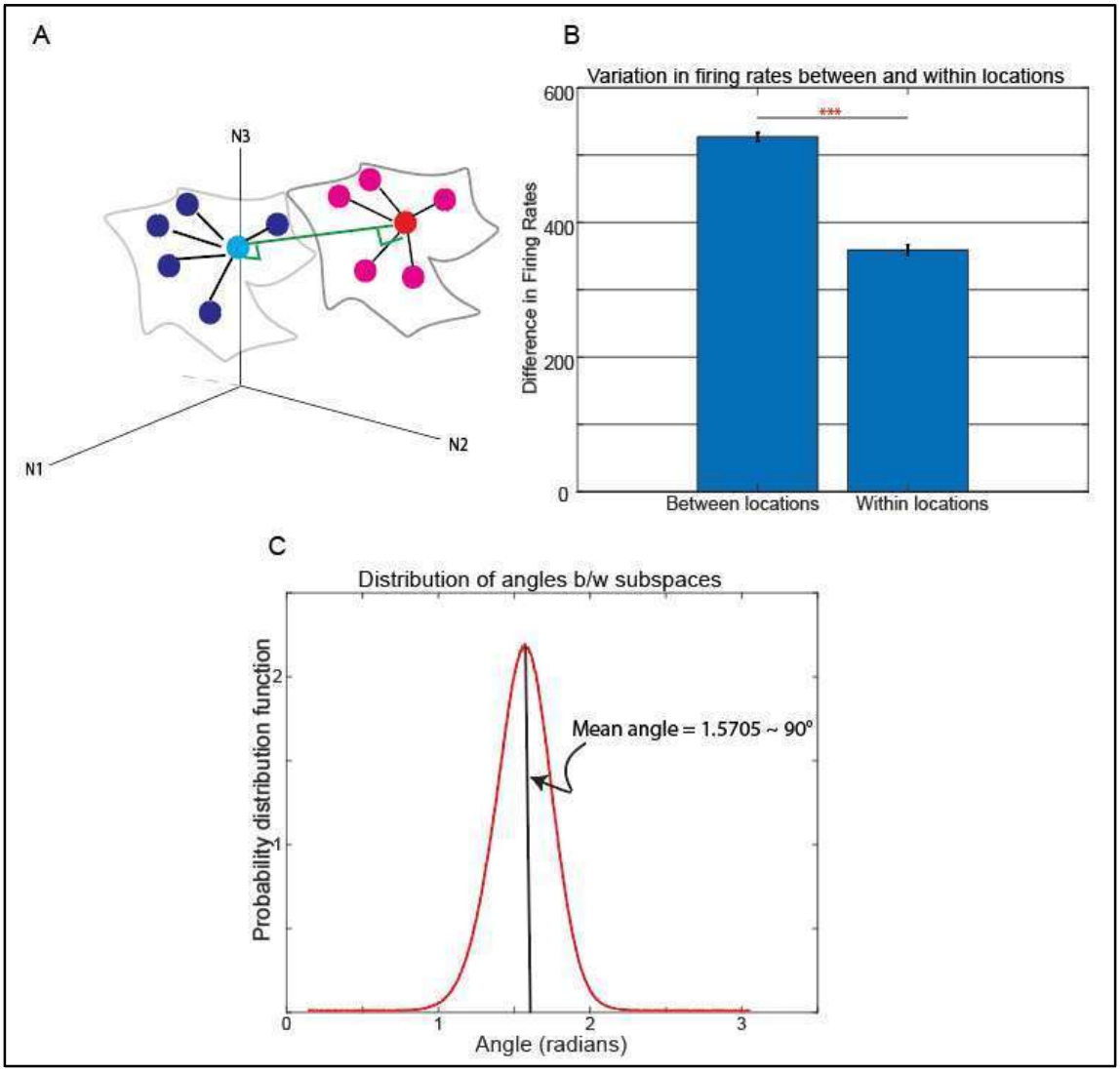

Figure 11. Population-vector Analysis. (A) Illustration of two subspaces representing two distinct spatial locations, with mean activity for each location shown in light blue and red, respectively. The between-location distance (in Hz) is shown in green, while within-location distances (in Hz) are shown in black. (B) Comparison of firing rate variations within subspaces and between subspaces, where each subspace corresponds to the firing rates observed for a specific spatial location. (C) Distribution of angles between subspaces, showing a clear peak around 90 degrees, indicating an orthogonal representation of spatial information.

RSC spiking activity can be learned using a VAE framework

Due to the limitations of Linear Discriminant Analysis (LDA)—such as its assumption that neural activity can be directly mapped onto spatial targets, its reliance on homoscedasticity, and the risk of overfitting, which we observed in our approach—we opted to use a Variational Autoencoder (VAE). This neural network-based method employs an encoder-decoder

framework, compressing the data into a latent space (neural manifold) and later reconstructing spike-time data from this compressed representation (**Figure 12**). This approach not only mitigated the limitations of LDA but also allowed us to extract neural manifolds, providing deeper insights into how spatial information is represented in spiking activity.

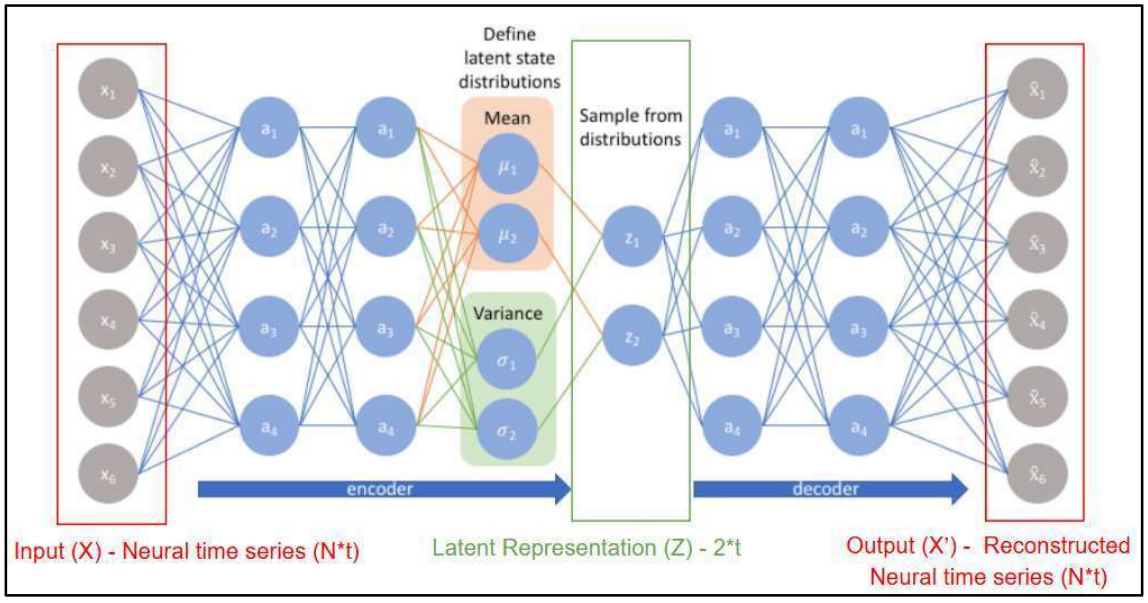

Figure 12. An illustration of the Variational Autoencoder (VAE) framework. The encoder maps high-dimensional neural activity into a lower-dimensional latent space, enforcing a probabilistic distribution. The decoder then reconstructs the original neural activity from the latent representation. Source: Jordan J, 2018.

We successfully implemented a Variational Autoencoder (VAE) to encode and decode the spiking activity of RSC neurons, achieving strong generalization without overtraining (Figure 13). This was done using data from 17 sessions across 4 rats, including those with AAV-mediated silencing of MEC inputs. The training R^2 for spiking activity was 0.8676 ± 0.0082, and the reconstruction test R^2 for spiking activity was 0.8671 ± 0.0083.

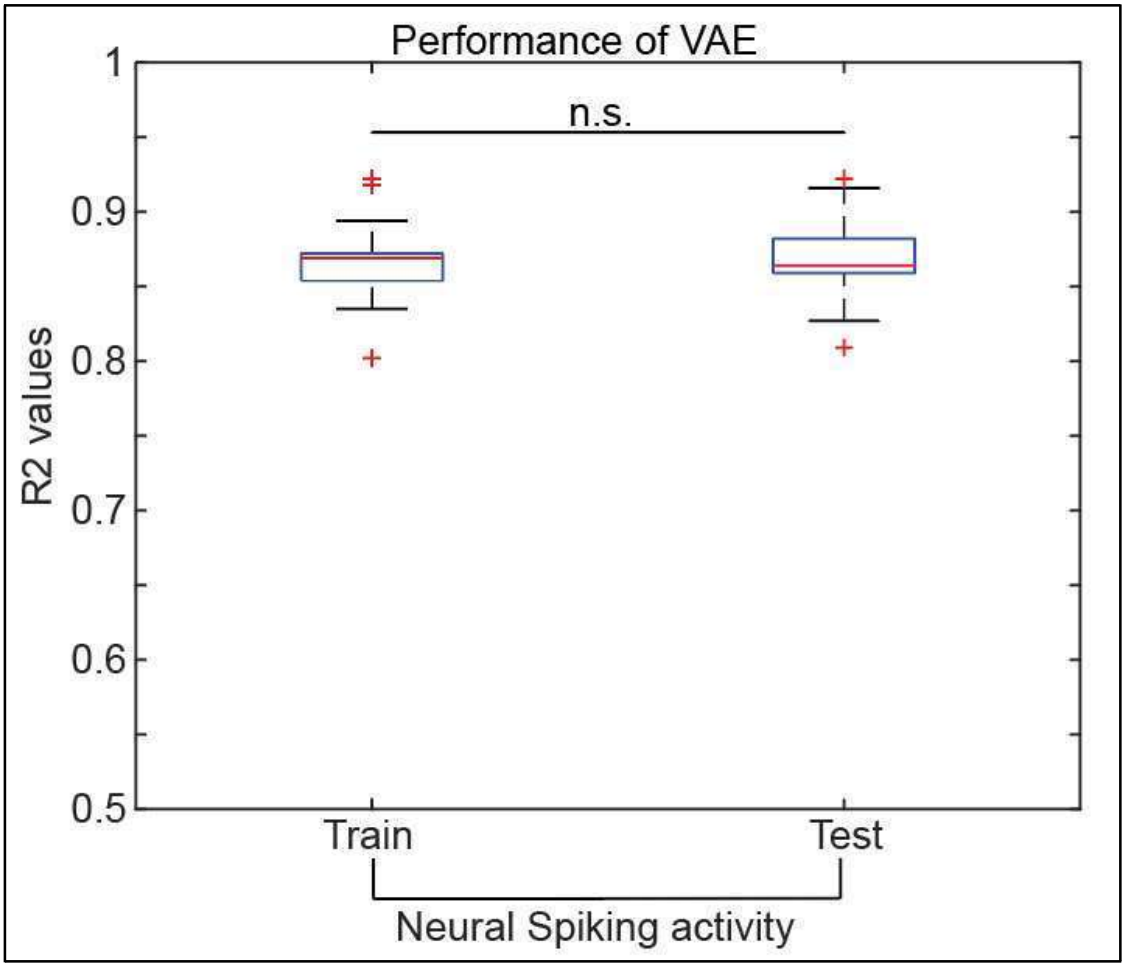

Figure 13. Performance of the Variational Autoencoder (VAE). The VAE algorithm successfully learns and reconstructs neural activity without overfitting, as evidenced by the lack of a significant difference between training and test performance.

S-VAE can learn the spatial information from the spiking activity of the RSC

As a Variational Autoencoder (VAE) is an unsupervised machine learning algorithm, it does not explicitly learn the spatial information encoded within the neural activity of the RSC.

To address this, we extended the model to a Supervised Variational Autoencoder (S-VAE) by incorporating a supervised loss function that predicts position from the latent space (neural manifold) and adding a mapping from the latent space to spatial (x, y) coordinates. This approach is advantageous because it not only learns to extract spatial information but also structures the latent space in a way that reflects the underlying spatial structure in the neural activity.

We observe that the S-VAE algorithm effectively learns and reconstructs spiking activity, but there is some evidence of overtraining, as the train and test R^2 values are significantly different (Wilcoxon signed-rank test, $p = 0.00024$), and there is instability in how cross-validation splits have been organized (**Figure 14**).

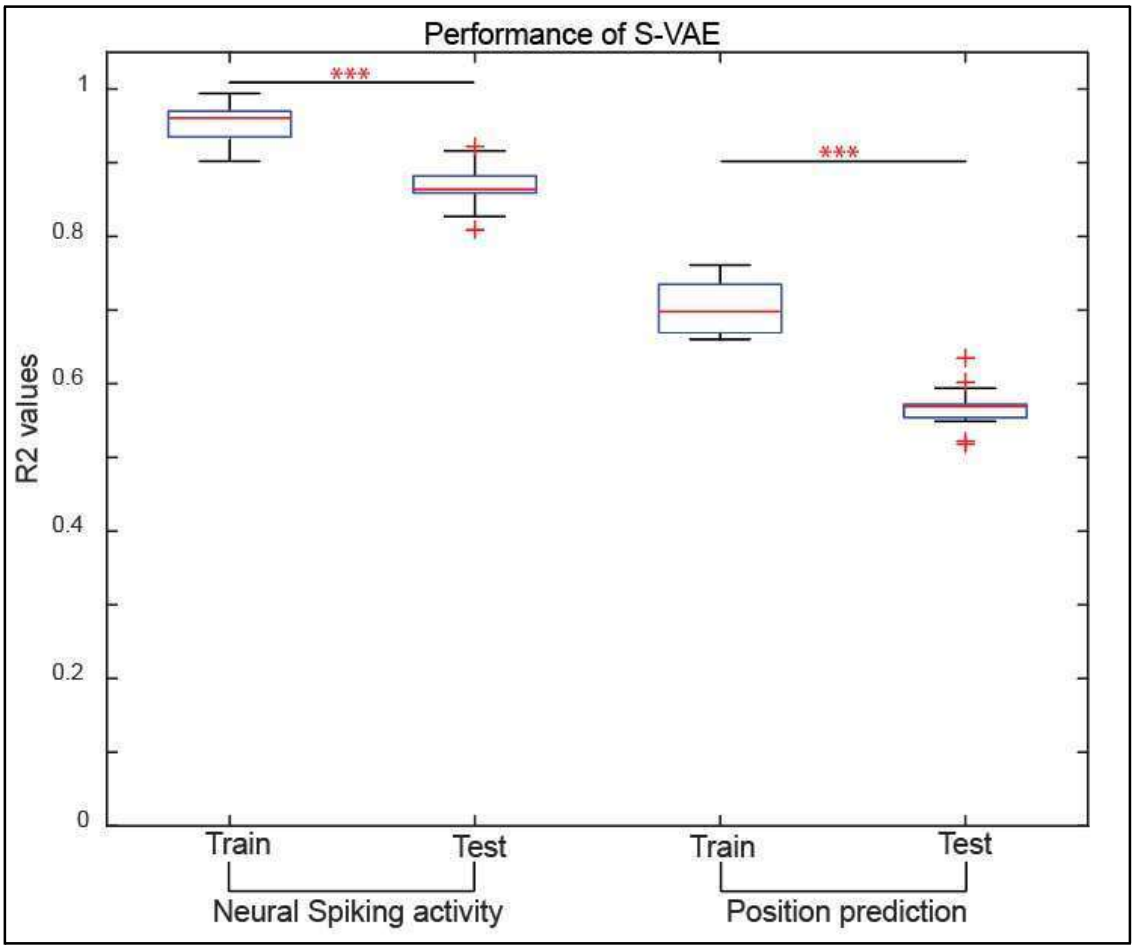

Figure 14. Performance of the Supervised-Variational Autoencoder (S-VAE). The first two bar plots show how well the algorithm encodes and decodes (reconstructs) neural activity. The difference between training and test performance suggests some overfitting, with the model performing well on training data but less effectively on test data. The next two bar plots show how well the learned latent space predicts spatial information (position). The significant difference between the training and test position R^2 values indicates that the latent space captures the (x, y) position but with reduced generalization on unseen data.

Next, we visualized the latent space learned by the algorithm by plotting pairs of latent space dimensions. We observed that the manifolds (**Figure 15**) were well-structured across different dimensions, displaying a gradual representation of 2D space.

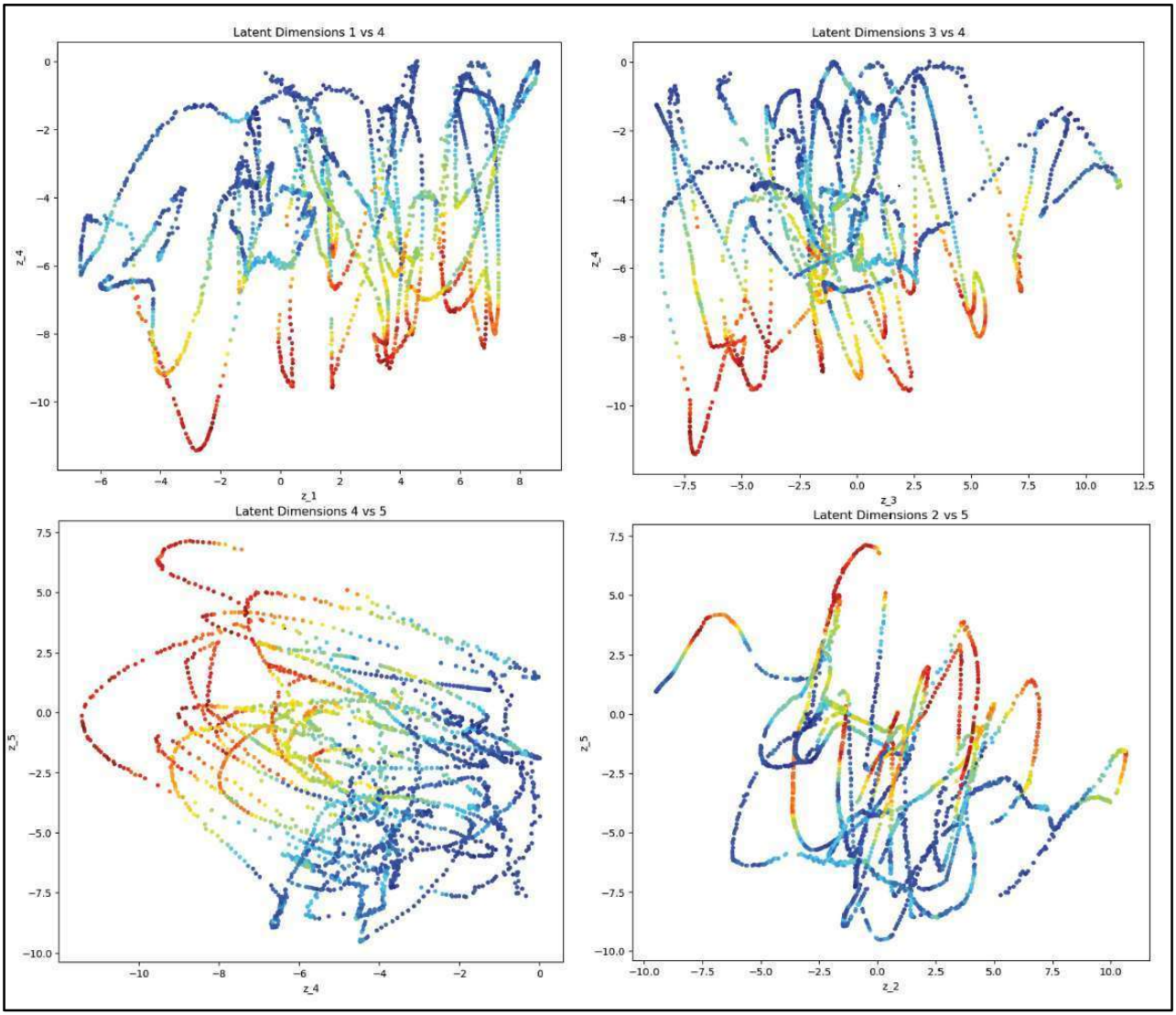

Figure 15. S-VAE Neural Manifolds. Example neural manifolds (latent space) plotted pairwise for a session, with individual points representing neural spikes and colors indicating the unique locations visited by the rat at the time of these spikes. The closer the colors, the closer the corresponding real-world locations. The coloring scheme follows the same convention as in Figure 9.

Chapter 4: Discussions

The retrosplenial cortex (RSC) is thought to integrate navigation-related signals and mediate allocentric-egocentric transformations.

However, the exact manner in which the RSC represents space has not been well established. While strides have been made in this direction leading to the discovery of egocentric border cells (EBCs) and head direction cells, the precise nature of spatial and positional representation in the RSC remains unclear.

In this study, we investigated how spatial information is represented in the RSC using electrophysiological recordings conducted in a 2D open field set up, in combination with decoding analyses, neural manifold analysis and population-vector approaches. Our results demonstrate that neural ensembles in the RSC encode spatial locations in a structured manner via neural manifolds, preserving real-world distances and maintaining clear spatial boundaries. Additionally, the high-dimensional ensemble spiking activity was well parsed into subspaces that were orthogonal to each other.

RSC Contains Border and Head-Direction Cells

We first estimated the proportion of previously identified border and head-direction cells (Alexander et al., 2020). Border cells made up ~13.2% of the population, head-direction cells ~21.4%, and conjunctive cells ~4.5%. Notably, the majority of neurons (~61%) did not code for either of these cues. This prompted us to further investigate the role of these remaining neurons within the ensemble.

We explored the possibility of the RSC containing spatially tuned (place) cells. Our Bayesian analysis revealed that the RSC does not exhibit strong single-cell spatial tuning.

A previous study by Mao et al. (2017) provided evidence for place cell-like activity in a one-dimensional setup using in vivo calcium imaging. In this setup, a mouse walked on a treadmill and periodically received tactile cues in the form of bumps. These cues indicated task progression, and a reward was delivered after each lap. We believe this setup is not directly comparable to our 2D open-field setup, as it includes additional tactile and visual cues. These extra cues make it more difficult to determine whether a neuron is truly coding for spatial location, responding to the tactile cue, or representing task progression toward the

goal.

RSC Ensembles Encode Spatial Information

We then examined whether the RSC represents space at an ensemble level rather than a single-cell level. Our approach using linear discriminant analysis (LDA) revealed that RSC activity supports ensemble-level spatial coding, which could be visualized as neural manifolds.

Our LDA-based spatial decoding suggested a directional influence of the medial entorhinal cortex (MEC) on RSC spatial coding, but the effect was not statistically significant. Nonetheless, we found that the neural manifolds derived from RSC activity faithfully represented real-world spatial relationships, with pairwise distances in the manifold space closely mirroring pairwise distances in the environment. This aligns with prior studies suggesting that higher-order cortical regions like the RSC play a role in integrating spatial and contextual information (Alexander & Nitz, 2015).

We also observed that although AAV-mediated MEC silencing did not affect positional decodability, it altered the resolution of distance representation in the manifold. Specifically, MEC inactivation caused real-world locations to appear closer in the manifold space than they otherwise would, indicating that MEC may contribute to the fidelity of spatial representation in the RSC.

An important caveat to note is that our LDA decoder did end up overfitting the data, and as a result, the observed effects could be slightly exaggerated. Nonetheless, the analysis provided valuable directions for further investigation.

Population-Level Organization of Spatial Codes

Using a population-vector approach, we observed that spatial locations were encoded in distinct subspaces of the neural spiking activity. Spiking activity corresponding to a given location was more tightly clustered around its mean activity than around the mean activity of other locations, indicating that unique locations are represented in separable neural activity patterns. Additionally, the subspaces corresponding to different locations were largely orthogonal, suggesting an efficient, high-dimensional encoding strategy that minimizes

interference between spatial representations. This finding is in line with previous reports of high-dimensional neural representations in sensory and cognitive systems, which may enhance the separability of information and support flexible computations (Fusi et al., 2016).

Impact of MEC Silencing on RSC Spatial Representations

In sessions where medial entorhinal cortex (MEC) input was silenced using AAV-mediated manipulation, we observed a systematic reduction in the resolution of spatial representation within the RSC. Locations that were well-separated in control conditions were represented closer together in the manifold under MEC silencing, indicating a loss of spatial precision. This suggests that while RSC is capable of generating structured spatial representations, its ability to maintain fine-grained spatial distinctions is influenced by MEC input. The reduction in spatial resolution aligns with previous findings that MEC provides critical spatial information to downstream regions, including RSC (Mao et al., 2017).

Advancing RSC Modeling with Variational Autoencoders

Given the limitations of LDA in capturing the full complexity of neural dynamics, we employed Variational Autoencoders (VAEs) to model RSC activity. The VAE effectively learned structured latent representations of neural activity, preserving spatial structure without overfitting. However, as an unsupervised model, it did not explicitly learn spatial information and instead captured the composite activity of the RSC, which is involved in multiple cognitive processes.

To address this, we implemented a Supervised-VAE (S-VAE) framework that incorporated a supervised loss function, mapping latent space representations to spatial (x, y) coordinates. This approach not only enabled the extraction of spatially meaningful neural manifolds but also provided insight into how hidden spatial structure is embedded in neural population activity.

The R^2 values from our VAE reconstructions indicated that the model captured a substantial fraction of the variance in neural activity, reinforcing the idea of a structured organization in spatial coding within the RSC. This approach helped mitigate the overfitting observed with LDA. Unlike traditional decoding methods, VAE-based models extract underlying structure rather than enforcing rigid mappings, making them particularly well-suited for studying

flexible neural representations.

Conclusion and Future Directions

Our findings highlight the role of the RSC in forming structured neural representations of space, integrating information from the MEC while maintaining its own population-level organization. The observed orthogonality of spatial subspaces suggests an efficient coding strategy that minimizes interference between spatial locations. The use of VAEs further supports the idea that RSC representations can be modeled as structured latent manifolds.

Future work should explore how these representations evolve in goal-directed tasks, extending beyond the purely allocentric representation observed in open-field exploration. Additionally, investigating how the RSC transforms between allocentric and egocentric representations could provide deeper insight into its function.

Finally, understanding how the RSC interacts with regions like the MEC and thalamus could shed light on its broader role in spatial cognition and memory. By combining experimental and a wide-array of computational approaches, this study lays the groundwork for future research into the circuit mechanisms underlying spatial representation.

References

1. Alexander AS, Carstensen LC, Hinman JR, Raudies F, Chapman GW, Hasselmo ME. 2020. Egocentric boundary vector tuning of the retrosplenial cortex. *Sci Adv*. 6:eaaz2322.
2. Alexander AS, Nitz DA. 2017. Spatially periodic activation patterns of retrosplenial cortex encode route sub-spaces and distance traveled. *Curr Biol*. 27:1551-1560.e4.
3. Alexander AS, Place R, Starrett MJ, Chrastil ER, Nitz DA. 2023. Rethinking retrosplenial cortex: perspectives and predictions. *Neuron*. 111:150-175.
4. Bicanski A, Burgess N. 2018. A neural-level model of spatial memory and imagery. *eLife*. 7:e33752.
5. Byrne P, Becker S, Burgess N. 2007. Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychol Rev*. 114:340–375.
6. Chang H, Esteves IM, Neumann AR, Sun J, Mohajerani MH, McNaughton BL. 2020. Coordinated activities of retrosplenial ensembles during resting-state encode spatial landmarks. *Philos Trans R Soc Lond B Biol Sci*. 375:20190228.
7. Chinzorig C, Nishimaru H, Matsumoto J, Takamura Y, Berthoz A, Ono T, Nishijo H. 2020. Rat retrosplenial cortical involvement in wayfinding using visual and locomotor cues. *Cereb Cortex*. 30:1985-2004.
8. Clark BJ, Simmons CM, Berkowitz LE, Wilber AA. 2018. The retrosplenial-parietal network and reference frame coordination for spatial navigation. *Behav Neurosci*. 132:416–429.
9. Cooper BG, Manka TF, van der Meer MA, McNaughton BL. 2001. Retrosplenial cortex maps the conjunction of internal and external spaces. *Nat Neurosci*. 4:580–585.
10. Davidson TJ, Kloosterman F, Wilson MA. 2009. Hippocampal replay of extended experience. *Neuron*. 63(4):497–507.
11. Fischer LF, Soto-Albors RM, Buck F, Harnett MT. 2020. Representation of visual landmarks in retrosplenial cortex. *eLife*. 9:e51458.
12. Fusi S, Miller EK, Rigotti M. 2016. Why neurons mix: high dimensionality for higher cognition. *Curr Opin Neurobiol*. 37:66-74.
13. Grossberger L, Battaglia FP, Vinck M. 2018. Unsupervised clustering of temporal patterns in high-dimensional neuronal ensembles using a novel dissimilarity measure. *PLoS Comput Biol*. 14:e1006283.

14. Hafting T, Fyhn M, Molden S, Moser MB, Moser EI. 2005. Microstructure of a spatial map in the entorhinal cortex. *Nature*. 436:801–806.
15. Hitchcock FL. 1941. The distribution of a product from several sources to numerous localities. *J Math Phys*. 20:224–230.
16. Interview: Alexander AS. 2020. Switching between egocentric and allocentric coordinates. *Sainsbury Wellcome Centre*. Available at: https://www.sainsburywellcome.org/web/qa/switching-between-egocentric-and-allocentric-coordinates.
17. Jacobs GH, Fenwick JA, Williams GA. 2001. Cone-based vision of rats for ultraviolet and visible lights. *J Exp Biol*. 204:2439–2446.
18. Jeffery K. 2017. Schematic of the RSC as seen in the midsagittal section and located just posterior to the corpus callosum in humans, rhesus monkeys, and rats. Available at: https://doi.org/10.6084/m9.figshare.5414179.v1 under a CC-BY 4.0 license.
19. Jordan J. 2018. Variational autoencoders. Available at: https://www.jeremyjordan.me/variational-autoencoders/.
20. Kira S, Safaai H, Morcos AS, Panzeri S, Harvey CD. 2023. A distributed and efficient population code of mixed selectivity neurons for flexible navigation decisions. *Nat Commun*. 14:2121.
21. Koslicki D. 2015. EMDeBruijn [GitHub repository]. Retrieved from https://github.com/dkoslicki/EMDeBruijn.
22. Langdon C, Genkin M, Engel TA. 2023. A unifying perspective on neural manifolds and circuits for cognition. *Nat Rev Neurosci*. 24:363–377.
23. Le L, Patterson A, White M. 2018. Supervised autoencoders: Improving generalization performance with unsupervised regularizers. *Department of Computer Science, Indiana University*.
24. Mackintosh NJ. 2002. Do not ask whether they have a cognitive map, but how they find their way about. *Psicológica*. 23:165–185.
25. Maguire EA. 2001. The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. *Scand J Psychol*. 42:225–238.
26. Mao D, Kandler S, McNaughton BL, Bonin V. 2017. Sparse orthogonal population representation of spatial context in the retrosplenial cortex. *Nat Commun*. 8:243.
27. Miller N. 2021. Taking shortcuts in the study of cognitive maps. *Learn Behav*. 49:261–262.
28. Mitchell AS, Czajkowski R, Zhang N, Jeffery K, Nelson AJD. 2018. Retrosplenial cortex and its role in spatial cognition. *Brain Neurosci Adv*. 2:2398212818757098.

29. Nakai S, Kitanishi T, Mizuseki K. 2024. Distinct manifold encoding of navigational information in the subiculum and hippocampus. *Sci Adv*. 10:eadi4471.
30. O'Keefe J, Dostrovsky J. 1971. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely moving rat. *Brain Res*. 34:171–175.
31. Pachitariu M, Steinmetz N, Kadir S, Carandini M, Harris KD. 2016. Kilosort: real-time spike sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv*. 061481.
32. Pele O, Werman M, Zisserman A. 2008. A linear time histogram metric for improved SIFT matching. In: Forsyth D, Torr P, editors. *Computer Vision - ECCV*. Berlin, Heidelberg: Springer. p. 495–508.
33. Pele O, Werman M. 2009. Fast and robust Earth mover's distances. *2009 IEEE 12th International Conference on Computer Vision*. p. 460–467. IEEE.
34. Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S. 2013. The importance of mixed selectivity in complex cognitive tasks. *Nature*. 497:585–590.
35. Rubner Y, Tomasi C, Guibas LJ. 1998. A metric for distributions with applications to image databases. *Proceedings of the Sixth International Conference on Computer Vision*. p. 59–66. IEEE.
36. Seelig JD, Jayaraman V. 2015. Neural dynamics for landmark orientation and angular path integration. *Nature*. 521:186–191.
37. Skinner BF. 1938. *The behavior of organisms: an experimental analysis*. Appleton-Century.
38. Solstad T, Boccara CN, Kropff E, Moser MB, Moser EI. 2008. Representation of geometric borders in the entorhinal cortex. *Science*. 322:1865–1868.
39. Sporns O, Honey CJ, Kötter R. 2007. Identification and classification of hubs in brain networks. *PLoS One*. 2:e1049.
40. Sugar J, Witter MP, van Strien NM, Cappaert NLM. 2011. The retrosplenial cortex: intrinsic connectivity and connections with the (para)hippocampal region in the rat. An interactive connectome. *Front Neuroinform*. 5:7.
41. Takahashi N, Kawamura M, Meguro K, Fukuzawa K, Yamaguchi S. 1997. Disproportionate retrosplenial MRI atrophy in a case of mild amnestic Alzheimer's disease. *J Neurol Neurosurg Psychiatry*. 63:624–626.
42. Thorndike EL. 1911. *Animal intelligence: experimental studies*. Macmillan.
43. Tolman EC. 1948. Cognitive maps in rats and men. *Psychol Rev*. 55:189–208.
44. Tye KM, Miller EK, Taschbach FH, Benna MK, Rigotti M, Fusi S. 2024. Mixed selectivity: cellular computations for complexity. *Neuron*. 112:2289–2303.

45. van Groen T, Wyss JM. 1990. Connections of the retrosplenial granular a cortex in the rat. *J Comp Neurol*. 300:593–606.
46. van Groen T, Wyss JM. 2003. Connections of the retrosplenial granular b cortex in the rat. *J Comp Neurol*. 463:249–263.
47. van Wijngaarden JB, Babl SS, Ito HT. 2020. Entorhinal-retrosplenial circuits for allocentric-egocentric transformation of boundary coding. *eLife*. 9:e57278.
48. Vann SD, Aggleton JP, Maguire EA. 2009. What does the retrosplenial cortex do? *Nat Rev Neurosci*. 10:792–802.
49. Wyss JM, van Groen T. 1992. Connections between the retrosplenial cortex and the hippocampal formation in the rat: a review. *Hippocampus*. 2:1–11.