
Flexible navigation with neuromodulated cognitive maps

Anonymous Author(s)

Affiliation

Address

email

Abstract

Animals naturally form personalized cognitive maps to support efficient navigation and goal-directed behavior. In the brain, the CA1 subregion of the hippocampus plays a key role in this process, hosting spatially tuned neurons that adapt based on the behavioral context and internal states. Computational models of this ability include labeled graphs with locally specified spatial information, which avoid global metric structure, and deep neural networks trained on spatial tasks that exhibit emergent spatial tuning. However, these approaches often struggle to model one-shot adaptive mapping and typically rely on plasticity rules that lack biological plausibility.

We propose a neural architecture inspired by place-cell dynamics that enables rapid on-the-fly construction of cognitive maps during exploration of novel environments. The model relies on velocity inputs and grid cell modules to generate spatial representations and integrates neuromodulatory signals responsive to boundaries and rewards. Learning combines synaptic plasticity, lateral inhibition, and modulatory gating of place-cell activity. For reward-driven navigation, the agent uses a modified Dijkstra algorithm to plan paths on the emergent cognitive map, treating place cells as nodes in a locally structured graph.

We compare our model with standard reinforcement learning (RL) agents and find that it achieves significantly higher sample efficiency, solving tasks in a single episode that RL agents require thousands of training steps to master. This performance advantage arises from biologically inspired inductive biases embedded in the model architecture. In simulation, the agent adapts to dynamic reward locations and changes in the environment layout. Analysis of neuromodulated place cells reveals task-dependent changes in tuning field size and spatial density, aligning with experimental findings from hippocampal recordings. These results highlight the promise of biologically grounded computation and locally structured graph representations for flexible and data-efficient cognitive mapping.

1 Introduction

Survival in complex environments demands efficient navigational strategies. From desert ants to humans, successful wayfinding—navigating toward goals out of sight—relies on internal spatial representations, or *cognitive maps* [1, 2]. These maps enable flexible planning and decision-making beyond simple stimulus-response associations.

Cognitive map theories propose multiple strategies for spatial navigation, ranging from simple route learning to survey-based and graph-based models. Route learning stores paths as action-position pairs but struggles with scalability and generalization at intersections [3, 4, 5]. Survey maps, grounded in Euclidean geometry, offer greater flexibility [4, 6] but sometimes contradict neural and behavioral data showing geometric distortions and topological biases [7, 8, 9, 10]. Labeled graphs strike a balance—encoding landmarks and transitions in a topological network that supports vector-like operations, planning, and prediction [11, 12, 13].

Neural substrates of spatial representations reside in the hippocampus (HP) and entorhinal cortex (EC), where specialized cells—including grid, border, speed, and place cells—encode geometric and contextual variables [14, 15, 16]. Place cells, particularly in CA1, anchor cognitive maps [17], integrating converging inputs from grid cells, CA3, and lateral EC [18, 19, 20, 21].

Neuromodulation plays a critical role in shaping this circuitry. Dopamine and other modulators adjust synaptic strength, tune place fields, and encode novelty and reward-prediction errors [22, 23, 24]. These projections reshape spatial tuning [25, 26, 27, 28], support novelty detection [23], and transmit prediction errors, particularly via lateral EC inputs [29, 30]. These mechanisms echo principles of reinforcement learning (RL), highlighting the role of neuromodulation in adapting spatial representations to behavioral relevance [31, 24].

Computational models have captured individual components of this system. Early work proposed that the hippocampus encodes spatial position and direction [32], while topological models based on route learning highlight scalability challenges [5]. More recent approaches draw from predictive coding and reinforcement learning, including successor representations and the Tolman-Eichenbaum Machine, which generalize across spatial and relational tasks while mimicking biological activity patterns [33, 34, 35]. Path integration models trained on motion cues give rise to grid- and place-like tunings [36, 37, 38]. Others incorporate reward-driven Hebbian plasticity modulated by neuromodulators [39]. Yet, few architectures unify these ingredients into a biologically grounded system that learns online and adapts flexibly to novelty.

In this work, we introduce a biologically inspired model of cognitive map formation that integrates place cell representations, neuromodulatory signals, and graph-based spatial reasoning. The model constructs a topological map online, linking place cells along experienced trajectories and enriching them with scalar-valued modulatory signals. These modulators form analog fields across the map [40], drive local Hebbian plasticity in response to sensory prediction errors [41, 31, 42], and help maintain and adapt reward-based neural representations [43, 44, 45, 46, 34].

This architecture supports efficient goal-directed navigation without extensive offline training, leveraging spatial priors, online plasticity, and modulatory feedback. We demonstrate how this system adapts to environmental changes and how neuromodulation influences place field allocation and remapping [47, 48], linking cognitive flexibility to underlying physiological mechanisms.

The remainder of the paper is organized as follows: Section 2 details the model and experimental setup; Section 3 presents results; Section 4 discusses broader implications and future directions.

2 Methods

Our model focuses on cognitive map formation through an agent’s experiences within a closed environment, as illustrated in Figure 1a..

The architecture operates with minimal external information—just binary reward and collision signals. Instead, spatial representation emerges primarily from idiothetic information: the agent’s perception of self-motion [49]. Specifically, we use the factual velocity vector (the agent’s actual environmental displacement) as the primary navigational cue, which in biological systems emerges from integrated inertial and relative motion signals [50, 51].

Place Cell Formation Spatial representation begins with a stack of grid cell modules, each encoding a periodic tiling of space over a 2D torus \mathbf{T}^2 (Fig. 1b,c). Grid cell activity vectors \mathbf{u}^{GC} project to two place cell layers with initially zeroed weights. When no place cells activate for the current input, a silent neuron is randomly selected and imprinted with the current activity. To ensure selectivity, imprinting is aborted if the new weight vector overlaps too strongly (via cosine similarity) with existing ones.

Each place cell's activation is given by a bounded activation function based on cosine similarity. Details of lateral inhibition and recurrent connectivity are given in Appendix A.1.

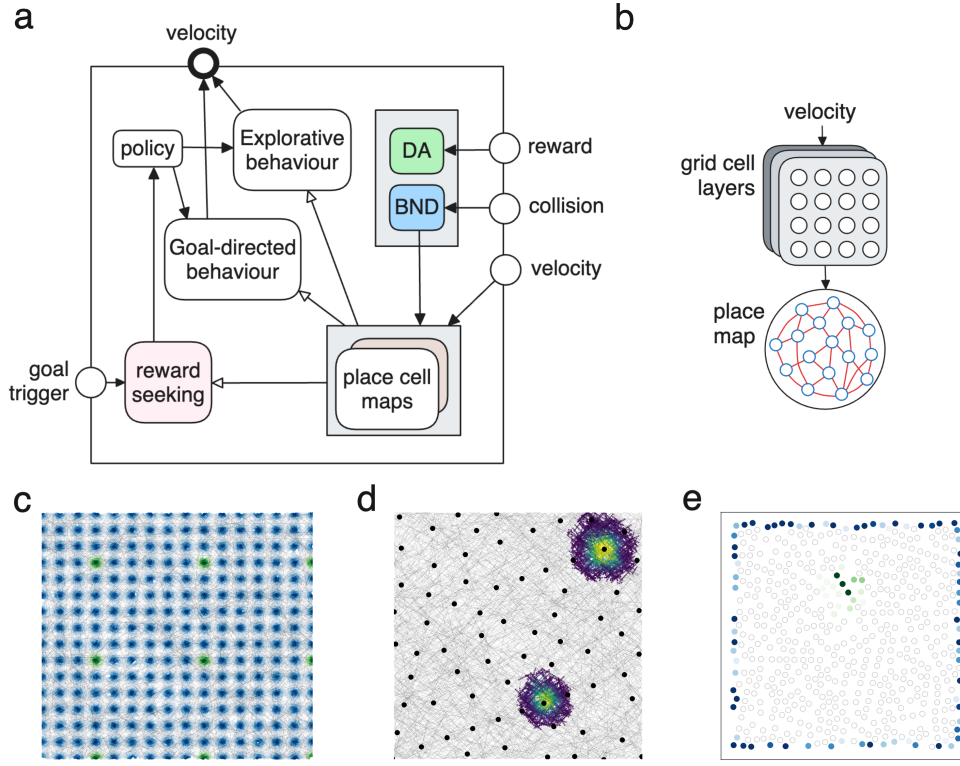


Figure 1: Model layout and spatial representations - **a**: the full architecture of the model, consisting of three main sensory input, targeting the two modulators and the cognitive map module, and the executive components, represented by a policy module, two behavioral programs and a reward receiver. **b**: the cognitive map component, organized with a stack of grid cell modules receiving the velocity input and projecting to the layer of place cells. **c**: the neural activity of a grid cell module from a random trajectory; in blue the repeating activity of all cell, while in green the activity of only one, highlighting the periodicity in space. **d**: the distribution in space of the place cells centers, together with the activity of two cells showing the size of their place field. **e**: neuromodulation activity over the place cells map, with in blue the cells tagged by the collision modulation, and in green the ones targeted by reward modulation.

Neuromodulation Neuromodulators encode environmental salience. Two scalar modulators, one for reward (DA) and one for boundary collisions (BND), are driven by binary inputs I and accumulate over time via exponential decay.

Each modulator k updates synaptic weights to place cells through Hebbian plasticity:

$$\Delta \mathbf{W}^k = \eta^k \mathbf{v}^k \mathbf{u}^{\text{PC}}$$

Weight vectors are constrained to remain non-negative. Reward modulation tags cells near rewarded locations, while boundary modulation builds a representation of environmental edges. These scalar fields form the core of the cognitive map (Fig. 1e). See Appendix A.2 for full learning rules and parameter settings.

Online Adaptation To remain resilient to environmental changes (e.g., moving rewards), the model uses a predictive mechanism to correct outdated internal representations. Before executing a movement toward position \mathbf{x}_{t+1} , the system predicts the expected value \hat{v}^k of neuromodulator k at

103 that location. Then, a prediction error is computed and used to adjust the weights:

$$\Delta \mathbf{W}^k \leftarrow \Delta \mathbf{W}^k + \hat{\eta}^k (\hat{v}^k - v^k) \mathbf{u}^{\text{PC}}$$

104 The prediction learning rate $\hat{\eta}^k < 1$ ensures that small errors are corrected gradually, while repeated
 105 mismatches lead to weight depression. This mechanism implements a simple form of predictive cod-
 106 ing and is inspired by temporal-difference learning [52], aligning with evidence that neuromodulatory
 107 systems signal prediction errors and update beliefs [53, 34, 54].

108 **Modulation of Place Fields** We further tested whether neuromodulators could directly alter spatial
 109 tuning. Place fields were dynamically shifted and resized based on recent salience signals.

110 Following a salient event (reward or collision), place field centers were displaced in grid cell space,
 111 with magnitude scaled by the neuromodulator v^k and proximity to the event:

$$\Delta \mathbf{W}_i^{\text{GC,PC}} = c^k v^k \varphi_{\sigma^k}(\mathbf{u}^{\text{GC}} - \mathbf{W}_i^{\text{GC,PC}})$$

112 Here, φ_{σ} is a Gaussian function, and c^k a scaling factor. This rule is inspired by BTSP plasticity
 113 [27, 47], which shifts CA1 place fields following salient experiences. Lateral inhibition prevents field
 114 overlap during remapping.

115 In addition to dislocation, field size was modulated by scaling the gain of recently active neurons.
 116 The gain β_i of each cell was adjusted via a trace variable m_i . This provides a mechanism for
 117 neuromodulators to transiently enhance or suppress spatial sensitivity.

118 **Policy and Behavior** To evaluate the model’s utility in navigation, we implemented a simple policy
 119 toggling between exploration and goal-seeking behavior, depending on an external goal flag and the
 120 internal map.

121 Exploration consisted of a stochastic walk with persistence, plus periodic plans to visit random known
 122 locations to avoid stagnation. Goal-directed navigation was triggered when a reward representation
 123 was present and executed via shortest-path planning on the place cell graph.

124 The graph was defined by place cells as nodes and synaptic links as edges. The agent selected targets
 125 either randomly or based on cells with high dopaminergic weight, echoing hippocampal replay and
 126 value-based navigation [55, 56, 57]. Planning was achieved via Dijkstra’s algorithm with a cost
 127 function that penalized proximity to boundaries.

Table 1: Comparison of neural network models for spatial navigation and representation

Model	Architecture	Training method	Ext. C.
Banino et al. [36]	LSTM + linear layers + CNN	BPTT and deep RL, supervised	Yes
Cueva et al. [38]	RNN + linear layers	Hessian-free algorithm with regularization	Yes
Sorcher et al. [37]	RNN + linear layers	Backpropagation with regularization	Yes
Whittington et al. [35]	Attractor network and deep networks	Backpropagation and Hebbian learning	No
de-Cothi et al. [34]	Successor representation	TD-learning + eligibility traces	Yes
Brozsko et al. [58]	Spike Response Model	Online modulated Hebbian plasticity	Yes
Ours	Rate layers	Online neuromodulated plasticity	No

Model	Task	Input	Output
Banino et al.	Path integration, goal navigation	Velocity, visual input, reward	PC, HDC
Cueva et al.	Path integration	Velocity	Position
Sorcher et al.	Path integration	Velocity	PC
Whittington et al.	Relational graph knowledge	Observation and action	Observation
de-Cothi et al.	Planned navigation	Observation	—
Brozsko et al.	Goal navigation	Position, reward	Action
Ours	Goal navigation	Velocity, reward, collision	Action

Note: PC = Place Cells, HDC = Head Direction Cells, Ext. C. = External Spatial Coordinates

Comparison with previous architectures Several previous computational models show structural and conceptual similarities with the present work. A prominent category among them employs deep neural networks—often with recurrent components—and relies on gradient-based learning strategies such as backpropagation through time. These models typically require multiple training episodes or large datasets for convergence. In contrast, our model adopts biologically inspired, synaptically local plasticity rules, and requires only a single training episode for adaptation.

Other models utilize spiking neurons [58] or explicit neural representations [34], and incorporate online learning rules more closely aligned with ours. These models also focus more directly on goal-directed navigation, in contrast to purely path integration tasks. However, both of these rely on external spatial coordinates to represent current position. Our model instead constructs an internal coordinate system by integrating its own velocity output, enabling endogenous spatial tracking.

Naturalistic task The model was evaluated on a biologically inspired navigation benchmark involving exploration and goal-seeking behavior in closed environments. Performance was measured as the total number of rewards collected over multiple trials.

Optimization of the model parameters was carried out using the evolutionary Covariance-Matrix Adaptation strategy (CMA-ES) [59] with a population of 128 individuals for 100 generations.

3 Results

Performance in wayfinding Our primary aim was to evaluate the formation of the cognitive map through neuromodulation in terms of the performance of the goal navigation in different environments.

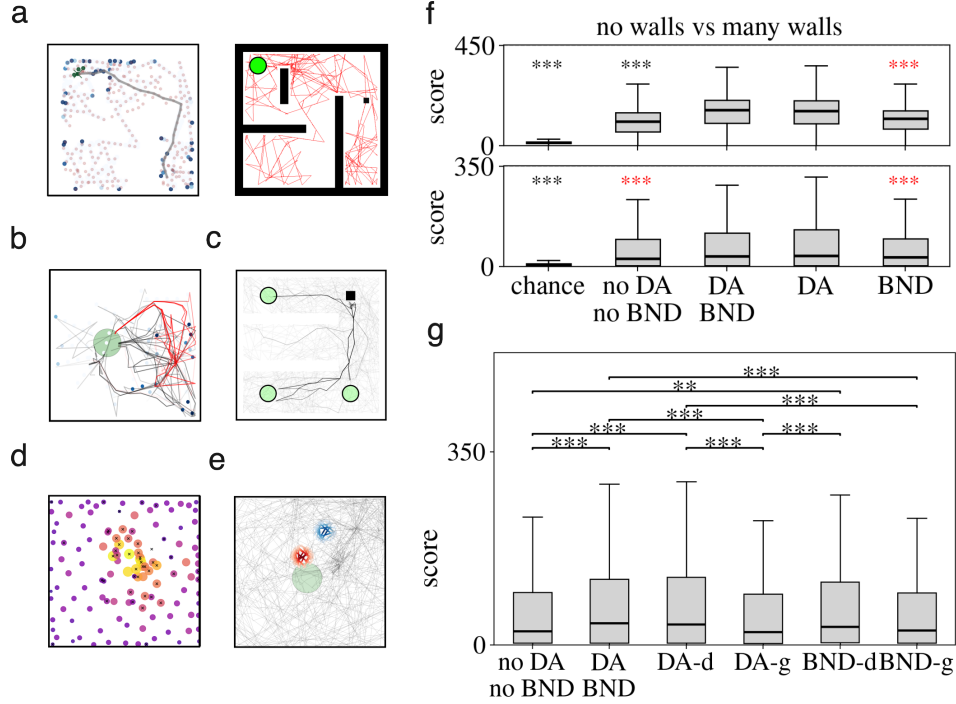


Figure 2: Cognitive maps and performance results - **a:** the plot on the left represent a cognitive map over a space, together with the plan (grey line) to reach a target location from a starting position. The plot on the right is a view of the same environment but with highlighted walls (black thick lines) the reward (green circle), trajectory (red line), agent position (black square). - **b:** plot of trajectories before (black) and after (red) the insertion of a wall between the starting and goal positions, the wall can be spotted from the boundary cells in blue - **c:** trajectories for multiple trials with the agent starting at the same position (black square) but with the reward location (green circles) periodically moving - **d:** place cells centers with size and color proportional to its node degree (number of neighbors); further, it notable is the correlation with the distance from the reward (in the center) - **e:** place fields of the same cell before and after several relocation of its center following reward events - **f:** performance comparison with same environmental conditions for five different models: one baseline as chance level (plasticity disabled, except for place cells generation), and four variants with different ablations of DA and BND modulations of place fields density and size. Results in terms of reward count and Bonferroni-corrected pairwise t-test, black stars stand for statistical difference with respect to all other groups, red stars only to the DA and DA+BND. - **g:** ablation performance comparison similar to c, and over the same environment for model variants with different active modulatory mechanisms: no DA and BND place field modulation, full DA and BND, DA-d for PC density, DA-g activation gain, BND-d for PC density, BND-g for activation gain.

147 The best model resulting from evolution reached solid navigation and adaptation skills. The agent
 148 was able to visit a significant portion of the environment during exploration and use neuromodulation
 149 to produce useful spatial representations.

150 The left panel of plot 2a displays place cells associated with collisions and reward events, signaling
 151 boundaries (in blue) and reward (in green) locations. The overlap of these two representations and
 152 the place cells (in pink) is what we refer to as a cognitive map, since these are the main sources of
 153 spatial and contextual information used during planned navigation, whose path is depicted as a gray
 154 line. The right panel instead portrays the actual environment with walls (black), reward location
 155 (green), and multiple trajectories (red). During exploration, the main areas were visited until the
 156 reward position was located and the goal-directed navigation dominated, as highlighted by the density
 157 of the path lines. Considering the position of the walls and corners, the layout of this environment
 158 does not always make the target locations visible, as it is a non-convex area and therefore can be
 159 classified as wayfinding [60]. The challenge of not being able to use straight lines is overcome by the
 160 graph approach using local data and the consideration of boundary place cells, allowing the agent to

161 plan accordingly. In addition, by construction the path also minimized the length of the path, within
162 the part of the space covered by the cognitive map.

163 In general, this result confirms the ability of the model to focus on navigation and obstacle avoidance.
164 However, it is worth noting that not all simulations resulted in a reward being found in the first place,
165 due to the randomness of the exploratory process; this was more pronounced in complex environment.

166 **Detour task** The planning ability and the plastic nature of the cognitive map should provide re-
167 siliience against unexpected changes in the environment layout. In order to verify this we implemented
168 a detour experiment. Initially, the agent was familiarized with a square environment with the reward
169 in the middle and starting always from the same position. Then, a wall was placed in between the
170 starting position and the reward, therefore forcing new trajectory for reaching it. As expected, the
171 agent was able to form a representation of the new obstacle and calculating new paths around it,
172 succeeding the task. In plot 2b they are shown the trajectories before and after the wall placement,
173 and it is manifested the ability of detour in the new layout.

174 **Adaptive goal representation through prediction error** Then, we tested the adaptability to
175 environmental changes. In this scenario, the reward object was moved after being fetched a fixed
176 number of times. Here, the difficulty was to unlearning previous locations and discovering new ones,
177 in a protocol similar to [39]. In plot 2c is reported the set of trajectories over many trials with the
178 reward displaced in three possible locations. The agent was capable of planning behavior, as earlier,
179 but also exploring and finding the new rewards, as shown by the density of lines. Whenever a goal
180 path resulted in a failed prediction, the DA-based sensory error weakened the association between the
181 place cells and the reward signal, leading to an extinction of its representation at that location.

182 This result validates the resilience of the model to changing sensory expectations, in this case the
183 reward position.

184 **Modulation of spatial resolution affects performance** Lastly, we investigated the effect of
185 modulating the density of place cells and the size of the field. The goal position was fixed, but the
186 agent was randomly relocated after fetching; performance was defined as the total number of reward
187 counts within a time window.

188 Our working hypothesis is that these experience-driven neuronal changes would improve the quality
189 of the cognitive map and be reflected in navigational abilities. The assessment of this claim was
190 conducted by comparing variants of the model, obtained by progressively ablating the reward (DA)
191 and collision (BND) modulation of density and field size, but a cognitive map was still possible as
192 their representations of goals and boundaries were preserved. We also defined a chance level by
193 instead blocking all modulation-based plasticity and allowing only place cells to form. All models
194 were run in two different environments differing by number of internal walls, in total 2048 simulations
195 were done for each case.

196 Plot 2d showcases the distribution of place cells with the circle size and color represent the node
197 degree, which aligns discretely with the reward position and the density. Further, in plot 2e the
198 place field of one cell is shown, before and after several reward occurrences and consequent center
199 relocation.

200 The statistical results are shown in plot 2f. The top reports the scores in the setting without internal
201 walls. All models performed above chance, but the main finding is the affirmation of the importance
202 of neuronal modulation, as revealed by the statistical difference between those endowed with it and
203 the one not. Furthermore, possession of DA modulation resulted in a significantly higher score than
204 BND modulation alone (red stars). In fact, in a situation with a convex region such as this, once the
205 reward has been located, the boundary information has limited utility.

206 A similar pattern emerged when more internal walls were introduced, as shown by the results at the
207 bottom. In this environment, scores were overall lower as navigation became more difficult, and
208 the difference among groups thinner, although noticeable. Here, the sole presence of the boundary
209 modulation of place cells did not result in better performance than the baseline. This outcome
210 highlights how the main improvements are brought by reward-driven neuronal modulation (DA).

211 Then, in order to enquire which particular modulatory action was mostly relevant, we conducted a
212 similar test on the same environment but ablating not only specific neuromodulators, but also whether

they affected the place cells density or the activation gain parameter. In plot 2g are reported the comparisons among different variants. The general trend matched the previous results 2c, in that DA is the best neuromodulator in terms of correlation with performance. The additional finding was that density modulation is significantly the primary action behind the behavioural improvements, and alone does not perform worse than the model with all actions together.

Taken together, these findings support the hypothesis of practical utility of direct modulation of place-field structure for active navigation, even in these simple settings.

4 Discussion

Exploration and planning in known and past environment are essential behaviors of animals, directly affecting their success in world understanding and goal reaching.

An important element behind these abilities is the formation of a map of their surroundings as they make new experiences, known as a cognitive map. Numerous speculations have been made about the shape and neural foundations of such an object, varying in the types of modeling assumptions and experimental support.

The contribution of the present work was to propose a rate network model, inspired by the CA1 hippocampal region [17]. We used grid cells together with synaptic plasticity as a mechanism to develop information-rich representations based on place cells updated through experience, grouped with common perspectives on cognitive maps [61]. In the spirit of minimizing the geometric assumptions in the neural space, we treated the generated place network as a topological graph, with sensory information added locally through the action of neuromodulators. This idea aligned with the concept of a *labeled graph* [62, 8], however, it is also true that no metric violations were actually possible in these settings.

The tasks we applied the agent consisted of an exploratory and exploratory phase, in which it was prompted to plan and reach reward positions. For simplicity, the first stage relied on a random walk process, as it was outside the scope of this work. This choice had the side effect that the reward was not always discovered, leading to the formation of incomplete maps, and thus impairing performance. However, this issue was limited in frequency.

The simulation results validated the model, showing the expected emergence of cognitive maps and their encoding of information collected during the experience. The online nature of the formation of the locations on the map aligns with the idea of using only idiothetic velocity input, as in path integration [6, 63, 64]. Previous work followed a similar direction using recurrent networks, but required extensive gradient-based training [65, 38, 51]. Another important difference is that our resulting neural network was composed solely of place cells, although neuromodulated, and no other types of neuron were present. This distinction is justified by the partially different task structure, which did not involve supervised learning and did not receive visual information as in [36]. Furthermore, our model relied on predefined grid cells layers, which constituted a strong and sufficient inductive bias, and did not have to be learned from scratch.

An additional relevant aspect is also the consideration of the place cell layer as an explicit graph data structure, on which the path-planning and decision-making algorithm was applied. The adoption of this level of description lead to robustness and flexibility, enabling effective navigation in all tested environments, which varying in layout complexity. Nevertheless, this approach did act as another clear inductive bias, which lifted the need to learn an approximation of it through network dynamics and even more differently tuned neurons.

Adaptability was tested by occasionally moving the reward position, leading to the generation of an internal prediction error that was used to update its representation on the map. The agent was proved capable of unlearning previous associations, returning to exploration, and memorizing new reward locations. This behavioral protocol is similar to previous work [58], in which dopaminergic and cholinergic activity was utilized within a Hebbian plasticity rule to strengthen or weaken reward-associated spatial representations. However, alternatively to exploiting neuromodulators with opposite valence, we followed a predictive coding framework, a direction linked to hippocampal representations [34, 66] and explored various computational approaches [67, 68, 69]. This choice departed from our focus on using operations on the cognitive map itself by simulating future sensory experiences and

learning from feedback. In fact, neuromodulation has been long associated with this functionality [40], especially dopamine [25, 43, 46, 31].

Lastly, the hypothesis of relevance of the active modulation of the neuronal properties of place cells was corroborated by simulating ablation experiments. These tests reported a significant impact of altering the place cells density on the total count of collected rewards. In general, these results are consistent with the experimental observations of alteration of place cells following reward events [27, 70], in particular in terms of increased clustering of cells [71, 72], reminiscent of changes in firing rate after contextual changes [73, 74].

Concerning the modulation of place fields, there is significant experimental evidence of their alteration during reward events [75, 76, 77], some reporting shrinkage near reward objects [78], and boundaries [79]. The coupling with higher local density could be explained by better optimization of the cell distribution for goal representation and planning [80]. However, in our settings, the fields become enlarged, especially in the direction of the target, although the performance improvements were not tested significant. A possible explanation can be the simplicity of our reward, which was solely defined as an area of space. The lack of rich non-spatial features thus did not require the place cells to code for smaller spatial variation. Therefore, enlargement might have improved the stability of the representation, marking the nodes associated with rewards more solidly, given the stochasticity of its delivery. Further, the graph-path algorithm utilized the strength of the DA-modulated connections for determining the goal representation, stronger fields inherently developed stronger weights, making planning more reliable planning. Although these findings are limited within the limit of our simulation protocol, there have been experimental observations of elongation of place fields along trajectories over meaningful experiences [81, 82].

In conclusion, this work showed a possible architecture for coupling emergent spatial representations with neuromodulated plasticity to achieve an experience-driven cognitive map. The reliance of few spatial and algorithmic inductive biases, grid cells and planning algorithm, supports the idea of label graph for goal navigation. Future work can investigate the application to other spatial domains, such as motor control and three-dimensional navigation. In addition, a richer input feature can be added, such as visual information [83], as well as new neuromodulators that encode different sensory dimensions or internally generated signals.

Acknowledgements & Statements

The authors declare no competing interests.

The code is publicly available and can be found at <https://github.com/iKiru-hub/PCNN>.

This research was funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N° 945371 and the University of Oslo.

The research presented in this paper has benefited from the Experimental Infrastructure for Exploration of Exascale Computing (eX3), which is financially supported by the Research Council of Norway under contract 270053.

References

- [1] Reginald Golledge, Dan Jacobson, Rob Kitchin, and Mark Blades. Cognitive Maps, Spatial Abilities, and Human Wayfinding. *GEOGRAPHICAL REVIEW OF JAPAN SERIES B*, 73:93–104, December 2000.
- [2] Russell A. Epstein and Lindsay K. Vass. Neural systems for landmark-based wayfinding in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1635):20120533, February 2014.
- [3] Michael Peer, Iva K. Brunec, Nora S. Newcombe, and Russell A. Epstein. Structuring Knowledge with Cognitive Maps and Cognitive Graphs. *Trends in cognitive sciences*, 25(1):37–54, January 2021.
- [4] Elizabeth R. Chrastil and William H. Warren. From Cognitive Maps to Cognitive Graphs. *PLoS ONE*, 9(11):e112544, November 2014.

- [5] Steffen Werner, Bernd Krieg-Brückner, and Theo Herrmann. Modelling Navigational Knowledge by Route Graphs. In Christian Freksa, Christopher Habel, Wilfried Brauer, and Karl F. Wender, editors, *Spatial Cognition II: Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications*, pages 295–316. Springer, Berlin, Heidelberg, 2000.
- [6] C. R. Gallistel and Audrey E. Cramer. Computations on Metric Maps in Mammals: Getting Oriented and Choosing a Multi-Destination Route. *Journal of Experimental Biology*, 199(1):211–217, January 1996.
- [7] Michael Peer, Catherine Nadar, and Russell A. Epstein. The format of the cognitive map depends on the structure of the environment. *Journal of Experimental Psychology: General*, 153(1):224–240, January 2024.
- [8] William H. Warren. Non-Euclidean navigation. *Journal of Experimental Biology*, 222(Suppl_1):jeb187971, February 2019.
- [9] Mark Wagner. Comparing the psychophysical and geometric characteristics of spatial perception and cognitive maps. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, 15(1):6–21, 2008.
- [10] Rainer Rothkegel, Karl F. Wender, and Sabine Schumacher. Judging Spatial Relations from Memory. In Christian Freksa, Christopher Habel, and Karl F. Wender, editors, *Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, pages 79–105. Springer, Berlin, Heidelberg, 1998.
- [11] Tobias Meilinger. The Network of Reference Frames Theory: A Synthesis of Graphs and Cognitive Maps. In Christian Freksa, Nora S. Newcombe, Peter Gärdenfors, and Stefan Wölfl, editors, *Spatial Cognition VI. Learning, Reasoning, and Talking about Space*, pages 344–360. Berlin, Heidelberg, 2008. Springer.
- [12] Jane X. Wang, Zeb Kurth-Nelson, Dhruva Tirumala, Hubert Soyer, Joel Z. Leibo, Remi Munos, Charles Blundell, Dharshan Kumaran, and Matt Botvinick. Learning to reinforcement learn, January 2017.
- [13] Victor R. Schinazi, Daniele Nardi, Nora S. Newcombe, Thomas F. Shipley, and Russell A. Epstein. Hippocampal size predicts rapid learning of a cognitive map in humans. *Hippocampus*, 23(6):515–528, 2013.
- [14] Francesca Sargolini, Marianne Fyhn, Torkel Hafting, Bruce L. McNaughton, Menno P. Witter, May-Britt Moser, and Edvard I. Moser. Conjunctive Representation of Position, Direction, and Velocity in Entorhinal Cortex. *Science*, 312(5774):758–762, May 2006.
- [15] Emilio Kropff, James E. Carmichael, May-Britt Moser, and Edvard I. Moser. Speed cells in the medial entorhinal cortex. *Nature*, 523(7561):419–424, July 2015.
- [16] Trygve Solstad, Edvard I. Moser, and Gaute T. Einevoll. From grid cells to place cells: A mathematical model. *Hippocampus*, 16(12):1026–1031, 2006.
- [17] Flavio Donato, Anja Xu Schwartzlose, and Renan Augusto Viana Mendes. How Do You Build a Cognitive Map? The Development of Circuits and Computations for the Representation of Space in the Brain. *Annual Review of Neuroscience*, 46(Volume 46, 2023):281–299, July 2023.
- [18] Daniel Bush, Caswell Barry, and Neil Burgess. What do grid cells contribute to place cell firing? *Trends in Neurosciences*, 37(3):136–145, March 2014.
- [19] Torsten Neher, Amir Hossein Azizi, and Sen Cheng. From grid cells to place cells with realistic field sizes. *PLOS ONE*, 12(7):e0181618, July 2017.
- [20] Tianyi Li, Angelo Arleo, and Denis Sheynikhovich. *Modeling Place Cells and Grid Cells in Multi-Compartment Environments: Hippocampal-Entorhinal Loop as a Multisensory Integration Circuit*. April 2019.
- [21] Olesia M. Bilash, Spyridon Chavlis, Cara D. Johnson, Panayiota Poirazi, and Jayeeta Basu. Lateral entorhinal cortex inputs modulate hippocampal dendritic excitability by recruiting a local disinhibitory microcircuit. *Cell Reports*, 42(1):111962, January 2023.

- [22] John E. Lisman and Anthony A. Grace. The Hippocampal-VTA Loop: Controlling the Entry of Information into Long-Term Memory. *Neuron*, 46(5):703–713, June 2005.
- [23] Adrian J. Duzsikiewicz, Colin G. McNamara, Tomonori Takeuchi, and Lisa Genzel. Novelty and Dopaminergic Modulation of Memory Persistence: A Tale of Two Systems. *Trends in Neurosciences*, 42(2):102–114, February 2019.
- [24] Wolfram Schultz, Peter Dayan, and P. Read Montague. A Neural Substrate of Prediction and Reward. *Science*, 275(5306):1593–1599, March 1997.
- [25] Kimberly A. Kempadoo, Eugene V. Mosharov, Se Joon Choi, David Sulzer, and Eric R. Kandel. Dopamine release from the locus coeruleus to the dorsal hippocampus promotes spatial learning and memory. *Proceedings of the National Academy of Sciences*, 113(51):14835–14840, December 2016.
- [26] Aude Retailleau and Thomas Boraud. The Michelin red guide of the brain: Role of dopamine in goal-oriented navigation. *Frontiers in Systems Neuroscience*, 8, March 2014.
- [27] Katie C. Bittner, Aaron D. Milstein, Christine Grienberger, Sandro Romani, and Jeffrey C. Magee. Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*, 357(6355):1033–1036, September 2017.
- [28] Alexandra Mansell Kaufman, Tristan Geiller, and Attila Losonczy. A Role for the Locus Coeruleus in Hippocampal CA1 Place Cell Reorganization during Spatial Reward Learning. *Neuron*, 105(6):1018–1026.e4, March 2020.
- [29] Kei M. Igarashi, Hiroshi T. Ito, Edvard I. Moser, and May-Britt Moser. Functional diversity along the transverse axis of hippocampal area CA1. *FEBS Letters*, 588(15):2470–2476, August 2014.
- [30] Hiroshi T. Ito and Erin M. Schuman. Functional division of hippocampal area CA1 via modulatory gating of entorhinal cortical inputs. *Hippocampus*, 22(2):372–387, 2012.
- [31] Denis Sheynikhovich, Satoru Otani, Jing Bai, and Angelo Arleo. Long-term memory, synaptic plasticity and dopamine in rodent medial prefrontal cortex: Role in executive functions. *Frontiers in Behavioral Neuroscience*, 16, January 2023.
- [32] Bruno Poucet. Spatial cognitive maps in animals: New hypotheses on their structure and neural mechanisms. *Psychological Review*, 100(2):163–182, 1993.
- [33] Paul Stoewer, Achim Schilling, Andreas Maier, and Patrick Krauss. Neural network based formation of cognitive maps of semantic spaces and the putative emergence of abstract concepts. *Scientific Reports*, 13(1):3644, March 2023.
- [34] William de Cothi, Nils Nyberg, Eva-Maria Griesbauer, Carole Ghanamé, Fiona Zisch, Julie M. Lefort, Lydia Fletcher, Coco Newton, Sophie Renaudineau, Daniel Bendor, Roddy Grieves, Éléonore Duvelle, Caswell Barry, and Hugo J. Spiers. Predictive maps in rats and humans for spatial navigation. *Current Biology*, 32(17):3676–3689.e5, September 2022.
- [35] James C. R. Whittington, Timothy H. Muller, Shirley Mark, Guifen Chen, Caswell Barry, Neil Burgess, and Timothy E. J. Behrens. The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. *Cell*, 183(5):1249–1263.e23, November 2020.
- [36] Andrea Banino, Caswell Barry, Benigno Uribe, Charles Blundell, Timothy Lillicrap, Piotr Mirowski, Alexander Pritzel, Martin J. Chadwick, Thomas Degris, Joseph Modayil, Greg Wayne, Hubert Soyer, Fabio Viola, Brian Zhang, Ross Goroshin, Neil Rabinowitz, Razvan Pascanu, Charlie Beattie, Stig Petersen, Amir Sadik, Stephen Gaffney, Helen King, Koray Kavukcuoglu, Demis Hassabis, Raia Hadsell, and Dhharshan Kumaran. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, May 2018.
- [37] Ben Sorscher, Gabriel Mel, Surya Ganguli, and Samuel Ocko. A unified theory for the origin of grid cells through the lens of pattern formation. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

- [38] Christopher J. Cueva and Xue-Xin Wei. Emergence of grid-like representations by training recurrent neural networks to perform spatial localization, March 2018.
- [39] Zuzanna Brzosko, Susanna B. Mierau, and Ole Paulsen. Neuromodulation of Spike-Timing-Dependent Plasticity: Past, Present, and Future. *Neuron*, 103(4):563–581, August 2019.
- [40] Marielena Sosa, Mark H. Plitt, and Lisa M. Giocomo. Hippocampal sequences span experience relative to rewards. *bioRxiv*, page 2023.12.27.573490, February 2024.
- [41] Abdullahi Ali, Nasir Ahmad, Elgar de Groot, Marcel A. J. van Gerven, and Tim C. Kietzmann. Predictive coding is a consequence of energy efficiency in recurrent neural networks, November 2021.
- [42] Jacopo Bono, Sara Zannone, Victor Pedrosa, and Claudia Clopath. Learning predictive cognitive maps with spiking neurons during behavior and replays. *eLife*, 12:e80671, March 2023.
- [43] Wolfram Schultz. Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1):23–32, March 2016.
- [44] Jeffrey B. Inglis, Vivian V. Valentin, and F. Gregory Ashby. Modulation of Dopamine for Adaptive Learning: A Neurocomputational Model. *Computational brain & behavior*, 4(1):34–52, March 2021.
- [45] Philippe N. Tobler, Christopher D. Fiorillo, and Wolfram Schultz. Adaptive Coding of Reward Value by Dopamine Neurons. *Science*, 307(5715):1642–1645, March 2005.
- [46] Roshan Cools. Chemistry of the Adaptive Mind: Lessons from Dopamine. *Neuron*, 104(1):113–131, October 2019.
- [47] Aaron D Milstein, Yiding Li, Katie C Bittner, Christine Grienberger, Ivan Soltesz, Jeffrey C Magee, and Sandro Romani. Bidirectional synaptic plasticity rapidly modifies hippocampal representations. *eLife*, 10:e73046, December 2021.
- [48] André A. Fenton. Remapping revisited: How the hippocampus represents different spaces. *Nature Reviews Neuroscience*, 25(6):428–448, June 2024.
- [49] Luxin Zhou and Yong Gu. Cortical Mechanisms of Multisensory Linear Self-motion Perception. *Neuroscience Bulletin*, 39(1):125–137, July 2022.
- [50] Steven J. Jerjian, Devin R. Harsch, and Christopher R. Fetsch. Self-motion perception and sequential decision-making: Where are we heading? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 378(1886):20220333, August 2023.
- [51] Ian Q. Whishaw and Brian L. Brooks. Calibrating space: Exploration is important for allothetic and idiothetic navigation. *Hippocampus*, 9(6):659–667, 1999.
- [52] Richard S Sutton and Andrew G Barto. The Reinforcement Learning Problem.
- [53] P. R. Montague, P. Dayan, and T. J. Sejnowski. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, 16(5):1936–1947, March 1996.
- [54] Seetha Krishnan, Chad Heer, Chery Cherian, and Mark E. J. Sheffield. Reward expectation extinction restructures and degrades CA1 spatial maps through loss of a dopaminergic reward proximity signal. *Nature Communications*, 13(1):6662, November 2022.
- [55] Colin G. McNamara, Álvaro Tejero-Cantero, Stéphanie Trouche, Natalia Campo-Urriza, and David Dupret. Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. *Nature Neuroscience*, 17(12):1658–1660, December 2014.
- [56] Frédéric Michon, Esther Krul, Jyh-Jang Sun, and Fabian Kloosterman. Single-trial dynamics of hippocampal spatial representations are modulated by reward value. *Current biology: CB*, 31(20):4423–4435.e5, October 2021.

- 459 [57] Philip Shamash and Tiago Branco. Mice identify subgoal locations through an action-driven
460 mapping process, December 2021.
- 461 [58] Zuzanna Brzosko, Sara Zannone, Wolfram Schultz, Claudia Clopath, and Ole Paulsen. Se-
462 quential neuromodulation of Hebbian plasticity offers mechanism for effective reward-based
463 navigation. *eLife*, 6:e27756, 2017.
- 464 [59] Christian Igel, Nikolaus Hansen, and Stefan Roth. Covariance Matrix Adaptation for Multi-
465 objective Optimization. *Evolutionary Computation*, 15(1):1–28, March 2007.
- 466 [60] Tobias Meilinger, Marianne Strickrodt, and Heinrich H. Bülthoff. Qualitative differences in
467 memory for vista and environmental spaces are caused by opaque borders, not movement or
468 successive presentation. *Cognition*, 155:77–95, October 2016.
- 469 [61] Vincent Hok, Pierre-Pascal Lenck-Santini, Sébastien Roux, Etienne Save, Robert U. Muller,
470 and Bruno Poucet. Goal-Related Activity in Hippocampal Place Cells. *Journal of Neuroscience*,
471 27(3):472–482, January 2007.
- 472 [62] Toru Ishikawa and Daniel R. Montello. Spatial knowledge acquisition from direct experience
473 in the environment: Individual differences in the development of metric knowledge and the
474 integration of separately learned places. *Cognitive Psychology*, 52(2):93–129, March 2006.
- 475 [63] Sabine Gillner and Hanspeter A. Mallot. Navigation and Acquisition of Spatial Knowledge in a
476 Virtual Maze. *Journal of Cognitive Neuroscience*, 10(4):445–463, July 1998.
- 477 [64] Bruce L. McNaughton, Francesco P. Battaglia, Ole Jensen, Edvard I. Moser, and May-Britt
478 Moser. Path integration and the neural basis of the ‘cognitive map’. *Nature Reviews Neuro-*
479 *science*, 7(8):663–678, August 2006.
- 480 [65] Ben Sorscher, Gabriel C. Mel, Samuel A. Ocko, Lisa M. Giocomo, and Surya Ganguli. A unified
481 theory for the computational and mechanistic origins of grid cells. *Neuron*, 111(1):121–137.e13,
482 January 2023.
- 483 [66] Fraser Aitken and Peter Kok. Hippocampal representations switch from errors to predictions
484 during acquisition of predictive associations. *Nature Communications*, 13(1):3294, June 2022.
- 485 [67] Manu Srinath Halvagal and Friedemann Zenke. The combination of Hebbian and predictive
486 plasticity learns invariant object representations in deep sensory networks. *Nature Neuroscience*,
487 pages 1–10, October 2023.
- 488 [68] Alexander Ororbia. Spiking neural predictive coding for continually learning from data streams.
489 *Neurocomputing*, 544:126292, August 2023.
- 490 [69] Kimberly L Stachenfeld, Matthew M Botvinick, and Samuel J Gershman. The hippocampus as
491 a predictive map. *Nature Neuroscience*, 20(11):1643–1653, November 2017.
- 492 [70] Indrajith R. Nair, Guncha Bhasin, and Dipanjan Roy. Hippocampus Maintains a Coherent Map
493 Under Reward Feature–Landmark Cue Conflict. *Frontiers in Neural Circuits*, 16, April 2022.
- 494 [71] Valerie L. Tryon, Marsha R. Penner, Shawn W. Heide, Hunter O. King, Joshua Larkin, and
495 Sheri J. Y. Mizumori. Hippocampal neural activity reflects the economy of choices during
496 goal-directed navigation. *Hippocampus*, 27(7):743–758, July 2017.
- 497 [72] Hannah S Wirtshafter and Matthew A Wilson. Differences in reward biased spatial representa-
498 tions in the lateral septum and hippocampus. *eLife*, 9:e55252, May 2020.
- 499 [73] Michael I. Anderson and Kathryn J. Jeffery. Heterogeneous Modulation of Place Cell Firing by
500 Changes in Context. *Journal of Neuroscience*, 23(26):8827–8835, October 2003.
- 501 [74] Inah Lee, Amy L. Griffin, Eric A. Zilli, Howard Eichenbaum, and Michael E. Hasselmo. Gradual
502 Translocation of Spatial Correlates of Neuronal Firing in the Hippocampus toward Prospective
503 Reward Locations. *Neuron*, 51(5):639–650, September 2006.

- [75] Marianne Fyhn, Sturla Molden, Stig Hollup, May-Britt Moser, and Edvard I. Moser. Hippocampal Neurons Responding to First-Time Dislocation of a Target Object. *Neuron*, 35(3):555–566, August 2002.
- [76] P.-P. Lenck-Santini, B. Rivard, R.u. Muller, and B. Poucet. Study of CA1 place cell activity and exploratory behavior following spatial and nonspatial changes in the environment. *Hippocampus*, 15(3):356–369, 2005.
- [77] David Dupret, Joseph O’Neill, Barty Pleydell-Bouverie, and Jozsef Csicsvari. The reorganization and reactivation of hippocampal maps predict spatial memory performance. *Nature Neuroscience*, 13(8):995–1002, August 2010.
- [78] S.N. Burke, A.P. Maurer, S. Nematollahi, A.R. Uprety, J.L. Wallace, and C.A. Barnes. The Influence of Objects on Place Field Expression and Size in Distal Hippocampal CA1. *Hippocampus*, 21(7):783–801, July 2011.
- [79] Sander Tanni, William De Cothi, and Caswell Barry. State transitions in the statistically stable place cell population correspond to rate of perceptual change. *Current Biology*, 32(16):3505–3514.e7, August 2022.
- [80] Pablo Scleidorovich, Jean-Marc Fellous, and Alfredo Weitzenfeld. Adapting hippocampus multi-scale place field distributions in cluttered environments optimizes spatial navigation and learning. *Frontiers in Computational Neuroscience*, 16:1039822, December 2022.
- [81] Mayank R. Mehta, Carol A. Barnes, and Bruce L. McNaughton. Experience-dependent, asymmetric expansion of hippocampal place fields. *Proceedings of the National Academy of Sciences*, 94(16):8918–8921, August 1997.
- [82] Jangho Lee, Jeonghee Jo, Byounghwa Lee, Jung-Hoon Lee, and Sungroh Yoon. Brain-inspired Predictive Coding Improves the Performance of Machine Challenging Tasks. *Frontiers in Computational Neuroscience*, 16:1062678, 2022.
- [83] John H. Wen, Ben Sorscher, Emily A. Aery Jones, Surya Ganguli, and Lisa M. Giocomo. One-shot entorhinal maps enable flexible navigation in novel environments. *Nature*, 635(8040):943–950, November 2024.
- [84] Yuri Dabaghian. Grid Cells, Border Cells and Discrete Complex Analysis.
- [85] Vemund Sigmundson Schøyen, Kosio Beshkov, Markus Borud Pettersen, Erik Hermansen, Konstantin Holzhausen, Anders Malthe-Sørensen, Marianne Fyhn, and Mikkel Elle Løpperød. Hexagons all the way down: Grid cells as a conformal isometric map of space. *PLOS Computational Biology*, 21(2):e1012804, February 2025.

5 Appendix

5.1 Grid cell module

Unlike other approaches to generate grid fields [84, 85], we defined a correspondence between the global environment in which the agent moves, a two-dimensional Euclidean space \mathbf{R}^2 , and a bounded local space of a grid module, corresponding to the torus.

The global velocity $\mathbf{v} = \{x, y\}$ is then mapped to a local velocity, scaled by a speed scalar s_l^{gc} specific to the grid cell module l , which determines its periodicity in space. This approach has been used in previous works [20].

The choice of a toroidal space is motivated by consolidated experimental evidence of the neural space of grid cells, which are organized in modules of different sizes spanning the animal’s environment. However, the shape of their firing pattern is known to be hexagonal, which corresponds to the optimal tiling of a two-dimensional plane, giving rise to a neural space lying on a twisted torus. In this work, for simplicity, we consider a square tiling and thus a square torus, without much loss of generality except for the slight increase of grid cells required for a sufficiently cover.

550 A grid cell module l of size N^{gc} is identified by a set of positions defined over a square centered
 551 on the origin and size of 2, such that $\{(x_i, y_i) \mid i \in N^{gc} \wedge x_i, y_i \in (-1, 1)\}$. This local square
 552 space has boundary conditions for each dimension, such that, for instance, when $x_t + s_l^{gc} \cdot v_x > 2$
 553 the position update is taken to the other side $x_{t+1} = x_t + s_l^{gc} \cdot v_x - 2$, where s_l^{gc} is the scale of the
 554 velocity in the local space of the module l with respect to the real global agent velocity $v = \{v_x, v_y\}$.
 555 When the module is initialized, the starting positions of its cells are uniformly distributed over the
 556 square forming a lattice. When the agent is reset in a new position at the beginning of new trial, a
 557 displacement vector is applied to the last cells positions such that the mapping between the module
 558 local space and the global environment is preserved.

559 The firing rate vector of each cell is determined with respect to a 2D Gaussian tuning curve centered
 560 at the origin at $(0, 0)$, and it is calculated as

$$561 \quad r_i = \exp\left(-\frac{x_i^2 + y_i^2}{\sigma_l^{gc}}\right), \text{ where } \sigma_l^{gc} \text{ is the width of the tuning curve for module } l.$$

562 The final population vector of the grid cell network GC is the concatenated and flattened firing rate
 563 vector of all modules \mathbf{u}^{GC} .

564

565 In our model, each grid cell had a tuning width of 0.04. They were defined as 8 modules of size 36,
 566 and the relative speed scales were
 567 $\{1., 0.8, 0.7, 0.5, 0.4, 0.3, 0.2, 0.1, 0.07\}$.

568 5.2 Place cells

569 **Tuning formation** The tuning of a new place cell is simply defined as the current GC population
 570 vector \mathbf{u}_t^{GC} , and its index is that of the first silent cell, which is added to the forward weight matrix
 571 $\mathbf{W}_i^{GCtoPC} \leftarrow \mathbf{u}^{GC}$.

572 In order to avoid overlapping of place fields, the tuning process is aborted in case the cosine similarity
 573 of the new pattern and the old ones is greater than a threshold θ_{rep}^{PC} .

574 Each cell represents a position in the GC activity space, which can be considered a node within a
 575 graph of place cells (PC). Although it is totally possible to only use the N^{GC} -dimensional tuning
 576 patterns and be agnostic about the dimensionality of the space in which the agent lives, to simplify
 577 the calculations, we mapped each pattern to 2D positions in a vector space. Then, the PC recurrent
 578 connectivity matrix is calculated with a nearest neighbors algorithm, which instead of a fixed number
 579 K of neighbors uses a distance threshold θ_{rec}^{PC} .

580 **Activity** The current firing rate of the PC population is determined by the cosine similarity between
 581 the GC input and the forward weight matrix, then passed through a generalized sigmoid $\phi(z) =$
 582 $[1 + \exp(-\beta(z - \alpha))]^{-1}$. The parameter α represents the activation threshold, or horizontal offset,
 583 while β the gain, or steepness.

$$u_i^{PC} = \phi\left(\cos\left(\mathbf{u}^{GC}, \mathbf{W}_i^{GC,PC}\right)\right)$$

584 It is also defined as an activity trace, which has an upper value of 1 and decays exponentially:

$$m_i = -m_i/\tau^{PC} + u_i$$

585 It is used as a proxy for a memory trace.

586

587 In the model, a PC population is defined by its average place field size, determining the granularity of
 588 the representation of the place.

589 Below, a table of the parameters specific to the two PC populations is reported.

	β	α	$\theta_{\text{rep}}^{\text{PC}}$	$\theta_{\text{rec}}^{\text{PC}}$	τ^{PC}
pc	33.0	1.0	0.86	43	140

Table 2: **PC layer parameters**

5.3 Modulation

Neuromodulation is defined in terms of a leaky variable v whose state is perturbed by an external input x , whose qualitative meaning differs for each neuromodulator k .

$$\begin{aligned} v_k &= -v_k/\tau_k + x_k \\ v_k &= \max(v_k, 0) \end{aligned} \quad (1)$$

Learning rule The connection weights \mathbf{W}^k are updated according to a plasticity rule composed of an Hebbian term, involving the leaky variable and the place cells above a certain threshold θ^k , and a prediction error.

A prediction \mathbf{p} is calculated before each time-step and signifies the expected neuromodulation activation for a given place cell $\hat{v}_i^k = \mathbf{W}_i^k \hat{\mathbf{u}}_i^{\text{PC}}$, where $\hat{\mathbf{u}}_i^{\text{PC}}$ is the PC population vector obtained by simulating the planned action. A prediction error is computed as the difference between the prediction and the current modulated place activation.

The full connection update then becomes:

$$\Delta \mathbf{W}_i^k = \eta^k v^k u_i^{\text{PC}} - \eta_{\text{pred}}^k (\hat{v}_i^k - v_i^k) u_i^{\text{PC}}$$

where η^k , η_{pred}^k are the learning rates, interpretable as the weight contribution of the Hebbian coupling and prediction accuracy respectively. Additionally, connections values are kept non-negative and, in order to speed up the extinction of past place-reward associations, if a non-zero prediction error is below a 0.5 then it is set to 0.5.

Active neuronal modulation Neuromodulation acts on the neuronal profile of the place cells by affecting the value of the activation gain and relocate the center of their tuning.

Gain modulation is implemented using the activity traces and a constant reference gain value $\bar{\beta}$:

$$\beta_i = c_a^k m_i \bar{\beta} + (1 - m_i) \bar{\beta}$$

where c_a^k is a scaling gain parameter, and if it is 1 then no modulation takes place.

Concerning center relocation, it is applied to recently active neurons with non-zero trace m_i . For a place cell i with position \mathbf{x}_i (in the vector space), it is calculated a displacement vector \mathbf{q}_i with respect to the current position \mathbf{x}_j , identified as the most active place cell j .

$$\mathbf{q}_i = c_b^k v^k \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{\sigma^k}\right)$$

where c_b^k is a scaling relocation parameter, while σ^k the width of the Gaussian distance. This displacement is used to move in GC activity space and get the new GC population vector to use as tuning pattern.

Also in this case, it is ensured that the new place field center is at a minimum distance $\theta_{\text{min}}^{\text{PC}}$ from the others; here Euclidean distance is used.

5.4 Decision making

Behaviour selection logic

The possible behaviours are *exploration* and *exploitation*, and an action is defined as a 2D velocity

vector. For exploration, an action can be generated either as random navigation, using a polar vector of fixed magnitude (the speed) and angle sampled from a uniform distribution, or as a step within a goal-directed navigation plan to reach a random destination, which corresponds to a randomly sampled existing place cells. In the goal-directed navigation the magnitude of the velocity vector is less or equal than a fixed speed value, depending on the distance from the next target position in the plan. Instead, for the exploitation behaviour, the action is a step within a goal-directed navigation towards the reward location. The behaviour selection process depends on the experience of collision, the presence of a plan, and the success in the navigation planning. A diagram of this logic is reported in Figure 3.

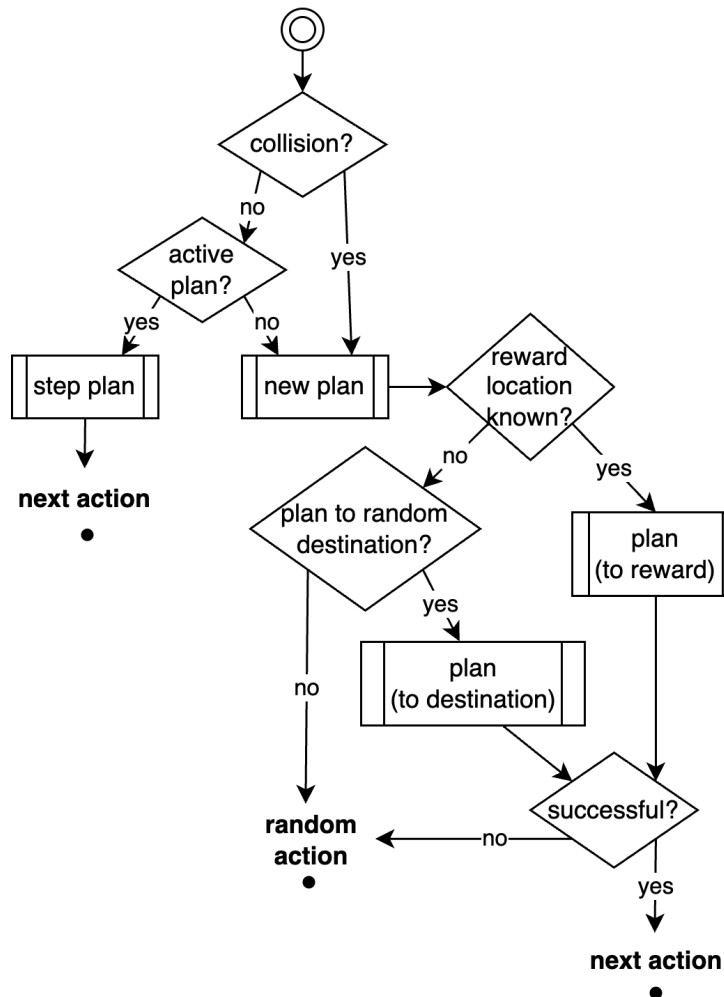


Figure 3: Diagram of the behaviour selection process

Algorithm 1 Modified Dijkstra algorithm

Require: Connectivity matrix $\mathbf{C} \in \{0, 1\}^{N^{PC} \times N^{PC}}$, Node coordinates $\mathbf{x} \in \mathbb{R}^{N^{PC} \times 2}$, Node weights $\tilde{\mathbf{W}} \in \mathbb{R}^{N^{PC}}$, Start node s , End node t

Ensure: Shortest path from t_0 to T

```
1: distances  $\leftarrow [\infty, \infty, \dots, \infty]$  ▷ Initialize distances
2: distances[ $t_0$ ]  $\leftarrow 0$ 
3: parent  $\leftarrow [-1, -1, \dots, -1]$  ▷ Parent pointers
4: finalized  $\leftarrow [\text{false}, \text{false}, \dots, \text{false}]$  ▷ Set of finalized nodes
5: PQ  $\leftarrow \emptyset$  ▷ Priority queue
6: PQ.push((0,  $t_0$ )) ▷ Insert start node with priority 0
7: while PQ  $\neq \emptyset$  do
8:   ( $\text{dist}, j$ )  $\leftarrow$  PQ.extractMin()
9:   if finalized[ $j$ ] or  $\text{dist} > \text{distances}[j]$  then
10:    continue
11:   end if
12:   finalized[ $j$ ]  $\leftarrow$  true
13:   if  $j = T$  then
14:    break ▷ Destination reached
15:   end if
16:   for each node  $i$  where  $\mathbf{C}_{i,j} = 1$  and not finalized[ $i$ ] do
17:     if  $\tilde{\mathbf{W}}[i] < -1000$  then
18:       continue ▷ Skip nodes with high negative weights
19:     end if
20:      $\Delta x \leftarrow \mathbf{x}_{j,0} - \mathbf{x}_{i,0}$ 
21:      $\Delta y \leftarrow \mathbf{x}_{j,1} - \mathbf{x}_{i,1}$ 
22:     edge_dist  $\leftarrow \sqrt{\Delta x^2 + \Delta y^2}$  ▷ Euclidean distance
23:     new_dist  $\leftarrow \text{dist}[j] + \text{edge\_dist}$ 
24:     if new_dist  $< \text{dist}[i]$  then
25:        $\text{dist}[i] \leftarrow \text{new\_dist}$ 
26:       parent[ $i$ ]  $\leftarrow j$ 
27:       PQ.push((new_dist,  $i$ ))
28:     end if
29:   end for
30: end while
31: path  $\leftarrow []$ 
32: if  $\text{dist}[T] = \infty$  then
33:   return  $\emptyset$  ▷ No path exists
34: end if
35: curr  $\leftarrow t$ 
36: while curr  $\neq -1$  do
37:   path.append(curr)
38:   curr  $\leftarrow$  parent[curr]
39: end while
40: path.reverse()
41: if path is empty or path[0]  $\neq s$  then
42:   return  $\emptyset$ 
43: end if
   return path
```

631 The planning of a new route is implemented as a modified Dijkstra algorithm over the place cell
632 graph, provided as connectivity matrix \mathbf{C} . Its particularity is the use of a weighting $\tilde{\mathbf{W}}$ of the nodes
633 according to the neuromodulation map. A description is reported in algorithm 5.4.

634 5.5 Environments

635 The game in which test the model has been developed with the python library Pygame, used under
 636 license GNU LGPL version 2.1 and available at <https://github.com/pygame/pygame>. The
 637 environment layout consisting in a customizable arrangement of vertical and horizontal hard walls
 638 with variable length and fixed width. Below in Figure 4 some samples are shown.

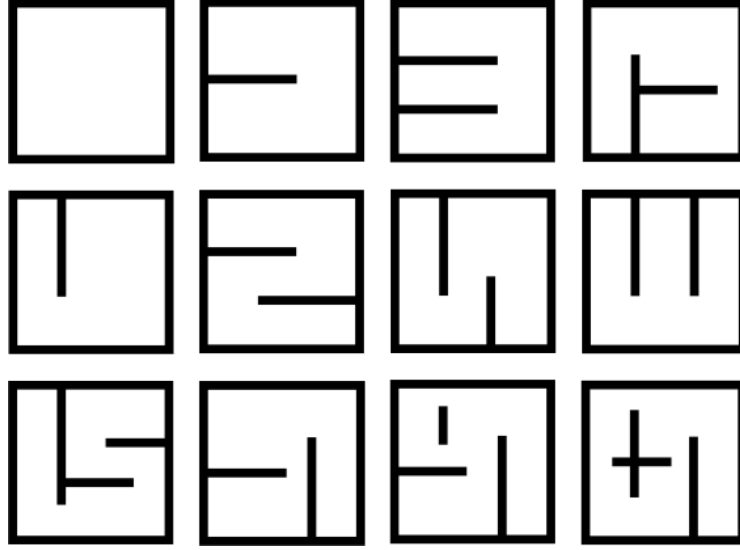


Figure 4: Sample of generated environments

639 The reward object is defined as a circle with size the 5% of the total environment area. When the
 640 agent position is within its boundary, it is provided a binary signal $R \sim \mathcal{B}(p_r)$ drawn from a Bernoulli
 641 with probability $p_r = 0.6$. The duration of the reward fetching is set to 2 time steps.

642 The agent object is defined as a square with size the 3.3% of the total environment area.

643 The testing protocol was inspired by the behavior of animals that venture into new territories in search
 644 of food. It was divided into two parts:

- 645 • **exploration phase:** the agent was placed in a random location within the environment for
 646 10'000 time steps. In this phase the reward is not present. Further, in order to force greater
 647 exploration of the environment, every 3'000 steps it was teleported to another random
 648 location. This external intervention was meant to mitigate the randomness in the exploratory
 649 behavioural strategy of the agent.
- 650 • **reward phase:** a reward is insert in a random location, and it is available to be discovered.
 651 When it is encountered, the agent is teleported to a random location within the environment,
 652 and after a fixed about 100 time step it is enabled its reward-seeking behaviour, in the form
 653 of goal-directed navigation. The total duration of this phase is set to 20'000 time steps.

654 An episode is defined as a continuous trajectory during the reward phase, namely a set of time steps
 655 starting from when the agent is place in a position until either it finds the reward or the simulation
 656 ends.

657 **Detour experiment** The protocol is modified such that after a fixed number of episodes the layout
 658 of the environment is changed, *e.g.* a wall is inserted. This experiment is meant to test the ability
 659 to reach the reward location by using the same cognitive map, and possibly update it with the new
 660 sensory information, such as the detection of the new boundaries.

661 **Changing reward experiment** During the reward phase, the reward location is changed after a
 662 fixed number of fetches.

1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [\[Yes\]](#)

Justification: the methods, experiments, and results match what is stated in the abstract and introduction

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [\[Yes\]](#)

Justification: the limits of the model architecture and of the experiments are taken into account within the Methods and Discussion sections.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [\[Yes\]](#)

Justification: the methodological premises of each experiment, with its result, are mentioned and discussed, in the Results and Discussion section.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [\[Yes\]](#)

Justification: the Appendix, together with the Methods section, reports enough information for a solid understanding of the functioning model and the experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
 - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
 - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
 - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
 - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: the codebase of this work is publicly available on github, and includes all script behind the model and experiments.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: the Methods, Results, and Appendix section include the necessary details accounting for the results.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: the statistical significance of the tests is reported in the figures.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).

- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: the specification of the computational resources is mentioned.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. Code of ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: the paper conforms to the NeurIPS Code of Ethics.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. Broader impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [NA]

Justification: to our knowledge, the paper does not involve nor impact any relevant societal aspect, for the main reason that it solely concerns computational modelling of neuronal and behavioural dynamics.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.

- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: this work does not involve any high risk data or models, and thus safeguards have not been deemed necessary.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: the creators, credits, and licensed have been properly mentioned and addressed.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.

- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. New assets

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: no new asset has been introduced, other than the code used for building and testing the model.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. Crowdsourcing and research with human subjects

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: the paper does not involve crowdsourcing nor human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. Institutional review board (IRB) approvals or equivalent for research with human subjects

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: the paper does not involve crowdsourcing nor human subject.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.

- 973 • We recognize that the procedures for this may vary significantly between institutions
974 and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the
975 guidelines for their institution.
976 • For initial submissions, do not include any information that would break anonymity (if
977 applicable), such as the institution conducting the review.

978 **16. Declaration of LLM usage**

979 Question: Does the paper describe the usage of LLMs if it is an important, original, or
980 non-standard component of the core methods in this research? Note that if the LLM is used
981 only for writing, editing, or formatting purposes and does not impact the core methodology,
982 scientific rigorousness, or originality of the research, declaration is not required.

983 Answer: [NA]

984 Justification: the usage of LLMs was limited to writing, editing, and formatting purposed.
985 There was no usage for developing the model or the research background.

986 Guidelines:

- 987 • The answer NA means that the core method development in this research does not
988 involve LLMs as any important, original, or non-standard components.
989 • Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>)
990 for what should or should not be described.