Title: Hippocampal replay facilitates the formation of entorhinal grid cells

Bo Zhang[1,2], and Jia Liu[2*]

[1] Beijing Academy of Artificial Intelligence, Beijing, 100084, China

[2] Tsinghua Laboratory of Brain and Intelligence & Department of Psychology, Tsinghua University, Beijing, 100084, China

*Correspondence shall be addressed to: JL, liujiathu@tsinghua.edu.cn

Abstract

The phenomenon of neuronal replay, the sequential reactivation of hippocampal place cells for past experiences, have been proposed to organize learned knowledges into a cognitive map. Here we used the spin-glass model to simulate the formation of the hexagonal pattern of grid cells, the metric of cognitive map, when exploring novel environments under the influence of neuronal replay. We found a significant enhancement of grid periodicity particularly in smaller grid scales after the application of reverse replay. This suggests that reversed replay plays a role in constructing a high-resolution cognitive map by expanding the horizon in reverse directions, thereby capturing the global structure of the environment. In conclusion, our study highlights the directional attribute of reverse replay in the formation of the cognitive map through facilitating grid cells' hexagonal patterns in the entorhinal–hippocampus system.

Main text (introduction)

Episodic memory, which is closely associated with hippocampus (Tulving and Markowitsch, 1998; Burgess et al., 2002), is metaphorically characterized as mental time travel (Tulving, 2002), with hippocampal neuronal replay to remember the past and predict the future (Schacter et al., 2007).

The foundation for the mental travel is the cognitive map (Tolman, 1948; Moser et al., 2017), which represents the relation of episodes under the metric of entorhinal grid cells (Hafting et al., 2005; McNaughton et al., 2006; Behrens et al., 2018). That is, each location in the cognitive map is presumably represented by the combinational code of periodic hexagonal activation patterns at multiple scales of grid cells (Bush et al., 2015). Recent studies have revealed an interplay between the neuronal replay and the cognitive map that the former helps to organize the abstract knowledge from environmental topology to construct the latter. For example, the reverse reactivation of sequential patterns is associated with the unique structure of object sequences (Wu and Foster, 2014; Liu et al., 2019); the impaired replay reported in humans (Nour et al., 2021) and rodents (Suh et al., 2013) with schizophrenia showed deficits in the structured mental representations. On the other hand, the hexagonal pattern is gradually formed when rodents placed in a novel environment start to explore the environment (Barry et al., 2012), whereas replays are observed immediately after the first lap of the exploration on a novel track (Foster & Wilson 2006). The coincidence in the emergence of both grid cells' hexagonal pattern and place cells' neuronal replay leads to a hypothesis that the replay plays a critical role in forming the hexagonal pattern of grid cells, through which the cognitive map is constructed. Here, we investigated the mechanism underlying the interplay by examining the relation between the neuronal replay and the grid cells (Figure 1c). To turn the cooccurrence into a cause-and-effect relation, the neuronal replay needs to be directly and selectively manipulated (e.g., turning on and off intentionally) for the impact on the formation of hexagonal pattern of grid cells, which is difficult to achieve in neurophysiological studies. Therefore, we employed computational modelling to illustrate the role of neuronal replay on grid cells. Specifically, the spin-glass model (Fuhs et al., 2006), a type of continuous attractor network that has been widely used in simulating gird cells (Burak and Fiete, 2009; Couey et al.,

2013; Shipston-Sharman et al., 2016; Khona & Fiete, 2022), was used while a virtual gorilla arranged to act as the animal navigating in novel environments (Fig.1d, see model parameters in Table.1 and the details in methods). We found a significantly enhanced spatial periodicity of grid cells' firing pattern after applying the reverse replay of past experiences to the model, particularly effective for smaller grid scales (SS), suggesting the role of neuronal replay on the construction of high-resolution cognitive map (Schuck and Niv, 2019).

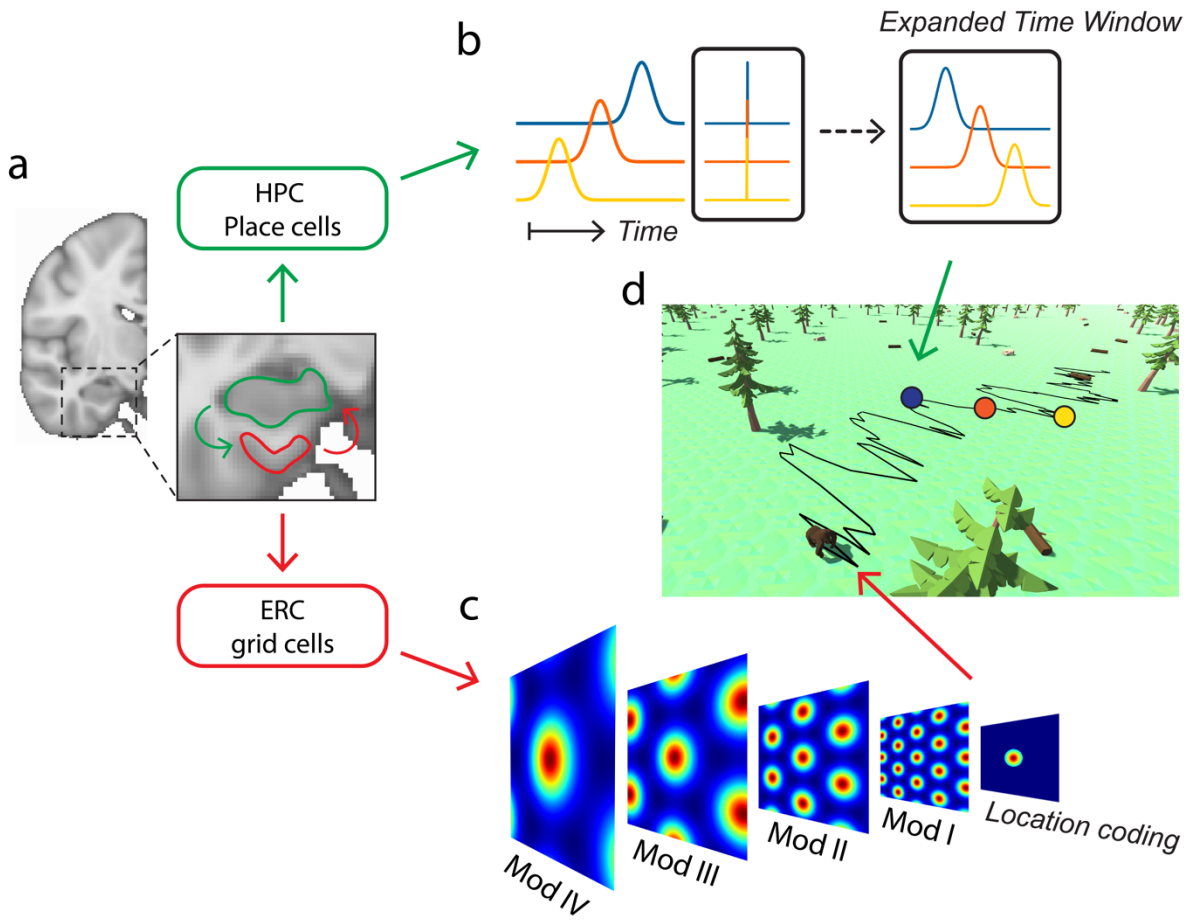

Figure 1. Anatomical and functional schema of the HPC-ERC system. a-c) hippocampal replay of place cells that represents past experience has been hypothesized to support the creation of cognitive map organized by the grid cells in the ERC, where a single location could be projected by unique populational coding of grid scales. d) coincidence of the neuronal replay and the grid cells when animal (acted by a gorilla) explores the novel environment. Colored

circles on the trajectory represent the place cells as shown in b), the location sequence organized by the place cells is replayed in time-compressed manner at pause location (green arrow).

Main text (results)

The spin-glass model (Fuhs and Touretzky, 2006) was used to simulate the formation of the hexagonal pattern of grid cells, where the lattice of spins (i.e., attractors) is interpreted as a two-dimensional grid, with each spin conjunctive by the head direction signals that cover full movement direction angles (0° - 360°) interacting with other spins through the excitation-inhibition dynamics of lattice network (see model parameters in Table.1 and the details in methods). In the study, a virtual gorilla was arranged to act as an animal navigating in an open environment (Fig.1d), with the movement velocity used as the input of model. The grid scales were determined by the spatial frequency of excitation-inhibition weight function, four weight functions were predefined to adapt with the capacity of high-resolution spatial coding (Fig.1c & Fig.S5, scale [i.e., Mod] I, II, III, and IV; see method section for details; Fiete et al., 2008; Stemmler et al., 2015; Bush et al., 2015). To examine the effect of neuronal replay in forming the hexagonal pattern of grid cells, we first examined the models with reverse replay (Fig.2a-c), which is for the consolidation of past experiences (Foster & Wilson, 2006; Ambrose et al., 2016; Gillespie et al., 2021). That is, during grid cell simulation, the movement sequence with the direction reversed relative to the actual movement direction were applied to the model at each pause location. To simulate the rapid formation of the hexagonal pattern of grid cells in coding visual space (Staudigl et al., 2018; Julian et al., 2018), the number of gorilla's movement directions was constrained, with the gorilla randomly moving in the open space towards either 70°,115°, and 160° during virtual navigation. After 15 independent simulations, the spatial periodicities of grid cells, indexed by Gridness (supplemental fig.3), were assessed.

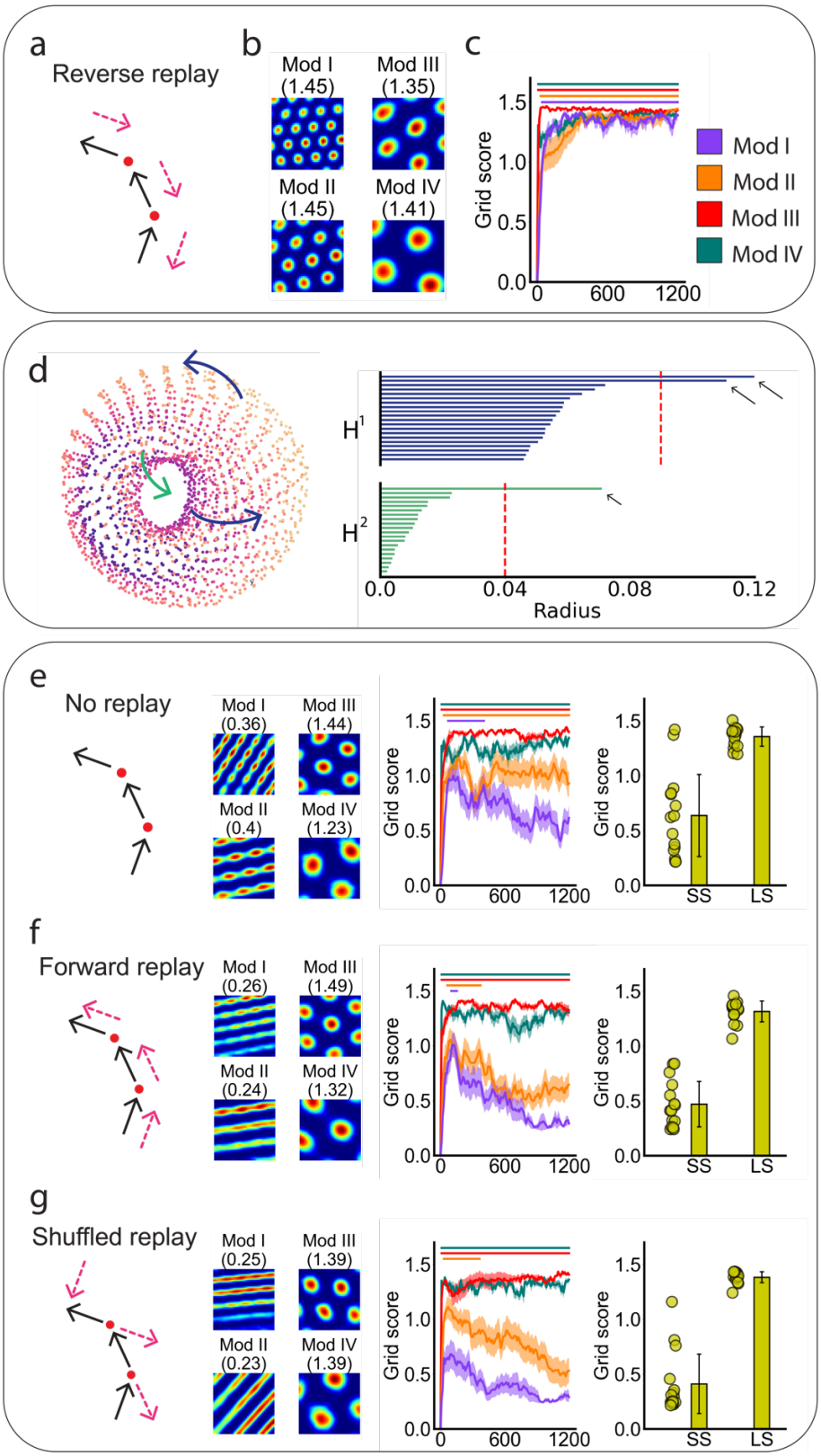
a
Reverse replay
b
Mod I (1.45)
Mod III (1.35)
Mod II (1.45)
Mod IV (1.41)
c
Grid score
1.5
1.0
0.5
0.0
0
600
1200
Mod I
Mod II
Mod III
Mod IV
d
H1
H2
0.0
0.04
0.08
0.12
Radius
e
No replay
Mod I (0.36)
Mod III (1.44)
Mod II (0.4)
Mod IV (1.23)
SS
LS
f
Forward replay
Mod I (0.26)
Mod III (1.49)
Mod II (0.24)
Mod IV (1.32)
g
Shuffled replay
Mod I (0.25)
Mod III (1.39)
Mod II (0.23)
Mod IV (1.39)

Figure 2. Effect of neuronal replay on the formation of multi-scale grid cells. a) schemas of animal trajectories, black arrows denote the direction of animal movement, red arrows denote the direction of neuronal replay. b) example of the multi-scale grid cells generated by spin-glass model; the number indicates Gridscore. c) population result, performed by 15 independent simulations, shows the fluctuation of Gridness during model iterations. The solid lines indicate the significance of Gridness during grid cell simulation ($p<0.05$, initial threshold, $p < 0.05$; corrected for multiple comparisons), error bars denote standard error. d) example of torus-like structure obtained by dimensionality reduction from the population activity of 484 grid cells of Mod IV. Persistent cohomology analysis shows the barcodes with 20 longest lifetimes were shown for the 1D hole (H^1, blue) and 2D hole (H^2, green), red dash line indicates the threshold of significance corrected for multiple comparison (radius = 0.083 / 0.024 for H^1 / H^2, respectively). e-g) the models simulated by no replay, forward replay, and the reverse replay with shuffled direction applied, respectively, the bar charts show the overall Gridness across four scales. Note that all models were applied with same number of iterations. SS and LS denotes small- and large-scale, respectively.

We found that the spatial periodicity of grid cells in all grid scales increases rapidly from the beginning of grid cell simulation, and then reached regular hexagonal patterns with only a minimum number of exploration, suggesting that, with reverse replay, the hexagonal patterns of grid cells can be formed rapidly (Fig. 2b & c, mean Gridness of Mod I / Mod II / Mod III / IV = 1.39 / 1.44 / 1.43 / 1.39, respectively; $p<0.05$, initial threshold, $p < 0.05$; corrected for multiple comparisons). To further validate the spatial periodicity of grid cells generated by the spin-glass model, we examined the topological structure of population activity from a total of 1936 grid cells (i.e., 484 grid cells for each scale, see method for details). Consistent with the finding reported by Gardner et al (2022), the toroidal structure, which embedded the population activity of grid cells (Fig.2d), was replicated as indicated by the significant lifetimes of barcode for two one-dimensional holes (the blue arrows, which indicate the inner and outer hole of torus) and for one two-dimensional hole (the green arrow, which indicates the cavity in the interior of torus) ($p<0.01$,

initial threshold, $p < 0.01$; corrected for multiple comparisons). Those results confirmed the regular firing pattern of multi-scales grid cells generated by the spin-glass model with reverse replay despite underlying limited environmental exploration.

To examine the necessity of the reverse replay, we next performed the simulation without replay, in which no direction-reversed movements were applied. Interestingly, deficits in grid periodicity were revealed by three control analysis particularly for grid Mod I & II (Fig. 2e, mean Gridness of Mod I / Mod II = 0.49 / 0.77), where grid periodicities showed a short increase at the beginning of simulation, and then followed by a sharp decrease during the remaining iterations, suggesting that reverse replay is necessary. In contrast, regular spatial periodicities were observed for the grid cells with large-scale (LS, Mod III & IV, mean Gridness of Mod I / Mod II = 1.42 / 1.29; $p<0.05$, initial threshold, $p < 0.05$; corrected for multiple comparisons). Next, we examined whether replay along is sufficient by applying forward replay, which occurs before actual movement and is thought to predict the future (Carr et al., 2011; Pfeiffer, 2020). A similar pattern to the no-replay condition was found (Fig. 2f, mean Gridness of Mod I / Mod II = 0.29 / 0.64; mean Gridness of Mod III & IV = 1.32 / 1.30; $p<0.05$, initial threshold, $p < 0.05$; corrected for multiple comparisons), suggesting that the replay in forming the hexagonal pattern is direction dependent. Finally, we examined the structure of reverse replay by shuffling the sequence of the movement while keeping the direction reversed in replay. Again, this scrambling of the spatial-temporal structure of reverse replay significantly reduced the Gridness in grid Mod I & Mod II (Fig. 2g, mean Gridness of Mod I / Mod II 0.29 / 0.53) and had little effect on grid Mod III & Mode IV (mean Gridness of Mod III / Mod IV = 1.40 / 1.36; $p<0.05$, initial threshold, $p < 0.05$; corrected for multiple comparisons), suggesting that the replay is also structure dependent. Taken together, the reverse replay, which

offers the sequential structure opposite to the actual experience, is particularly effective for the formation of the hexagonal pattern in small grid scales as opposed to that in large scales (figure 2e-g right; Mod I & II versus Mod III & IV: $t(14) = 7.02 / 13.89 / 13.28$ for no-replay, forward-replay, and shuffled-replay, respectively, $p<0.001$, two-sided), suggesting the necessity of reversed replay in constructing a high-resolution cognitive map.

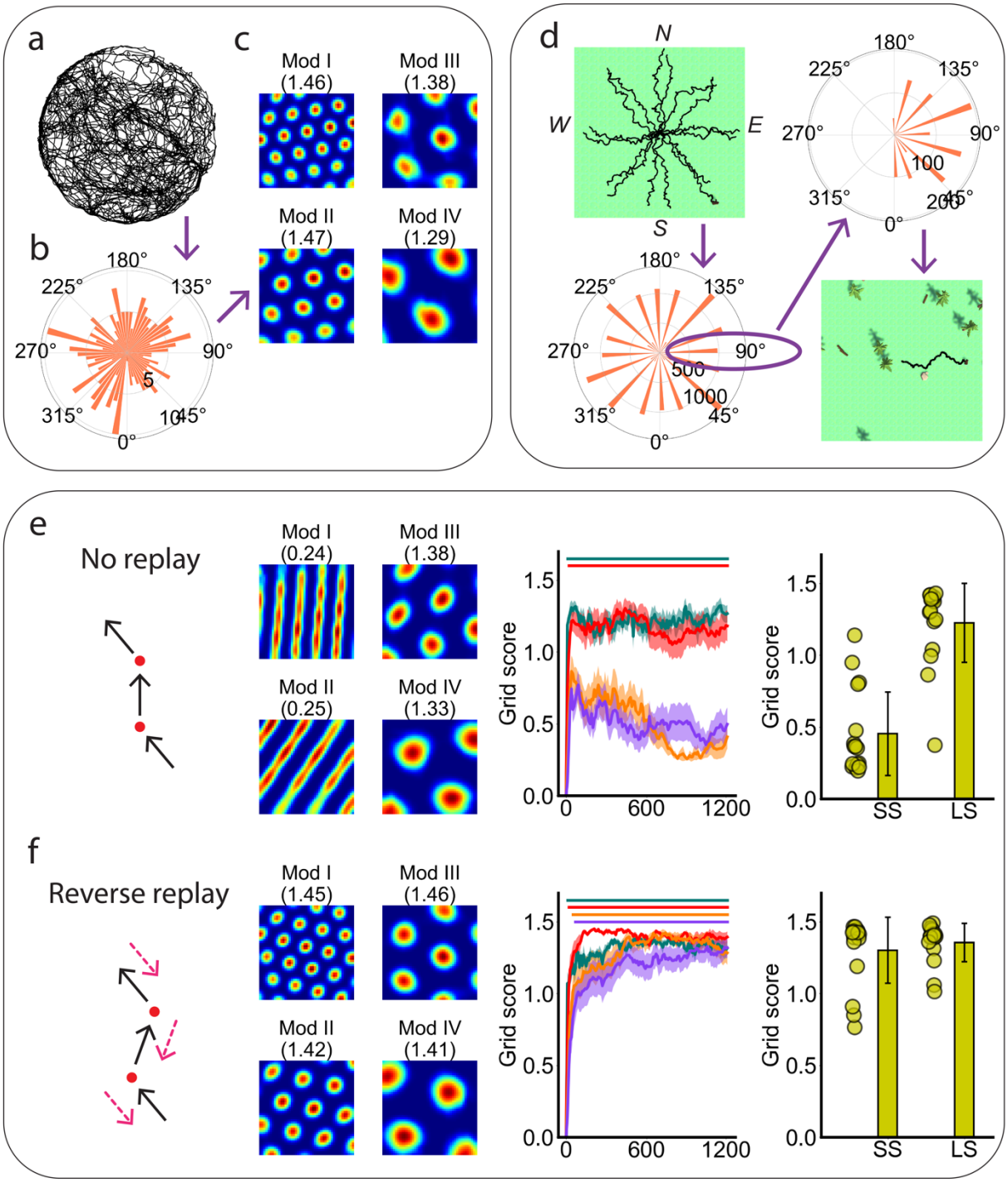

Figure 3. Effect of the amount of exploration on the formation of multi-scale grid cells. a) animal trajectories obtained from open-source data (Hafting et al., 2005), b) polar plot denotes the distribution of movement directions ($p>0.05$, Rayleigh test), c) example of the multi-scale grid cells generated by spin-glass model without neuronal replay applied; the number indicates Gridscore. d) overlapped trajectories of eight global directions with 45° apart projected on the

visible area of environment with a size of 15 × 15 virtual meters, polar plot denotes the distribution of movement directions ($p>0.05$, Rayleigh test). In addition to each of eight global directions, a noise of -90° to 90° was applied to animal's movement. For example, the 90° trajectory (purple circle) contains the directions from 0°-180° while animal moves towards East (2nd column). e-f) the models simulated by no replay and reverse replay applied, respectively, the line and bar charts indicate the population results performed by 16 independent simulations. Note that all models were applied with same number of iterations. SS and LS denotes small- and large-scale, respectively.

The facilitation of reverse replay on the formation of grid periodicity may be achieved by anchoring the self into a space with horizon expanded by the reverse direction to capture the global structure of the space, and a high-resolution map is thus constructed with each potential target location represented by the population vector field of place cells (Ormond and O'Keefe, 2022). This mechanism is theoretically equivalent to thoroughly exploration of the environment without replay. To test this idea, we used real animal trajectories published previously (Hafting et al., 2005), where the movement directions of animals distributed uniformly around 0° - 360° (Fig. 3a & b, $p=0.23$, Rayleigh test), to drive the model without replay. In line with our intuition, regular spatial periodicities were observed for all grid scales (Fig.3c, Gridness = 1.46/1.47/1.38/1.29 for Mod I / II / III / IV, respectively. $p<0.05$, initial threshold, $p < 0.05$; corrected for multiple comparisons), supporting our hypothesis on the mechanism underlying the facilitation of reverse replay on grid periodicity in maximizing the organization of cognitive map.

An alternative explanation, though, is that in our previous experiments we only used three directions with a very limited number of explorations, and therefore, with more directions and a large number of explorations, grid periodicity may be achieved without replay as well. To test this alternative, the model with hundreds of movement directions but without replay was examined, as

the gorilla was arranged to move towards one of eight global directions that covers 0°-360° with a noise randomly selected from a range of -90° to 90° for each footstep (Fig.3d and Fig.S3, $p<0.001$, Rayleigh test), which ensures a large number of explorations and no opposite direction possible. We found that without replay, grid periodicity in small grid scales was impaired (Fig.3e, mean Gridness of Mod I / Mod II = 0.49 / 0.41, respectively), which was significantly irregular as compared to that in large grid scales (Fig.2e; mean Gridness of Mod III / Mod IV = 1.18 / 1.27, respectively; $p<0.05$, initial threshold, $p < 0.05$; corrected for multiple comparisons; Mod I & II versus Mod III & IV: $t(15) = 7.47$, $p<0.001$, two-sided). In contrast, the quasi-periodicity of small grid scales was eliminated when reverse replay was applied (Fig.3f, mean Gridness = 1.32 / 1.28 / 1.39 / 1.32, for Mod I / II / III / IV, respectively. $p<0.05$, initial threshold, $p < 0.05$; corrected for multiple comparisons), which showed significantly higher spatial periodicity than the one without replay (Fig.3f, $t(15) = 8.85$, $p<0.001$, two-sided). Taken together, our results revealed that the role of reverse replay in constructing global structure of cognitive map rather than the local structures.

Main text (discussion)

The present study aimed to examine the impact of hippocampal replay on the formation of the hexagonal pattern of grid cells through computational modelling. We found that by applying the reversed activation sequence of past experiences to the spin-glass model, the grid regularity was significantly enhanced for smaller grid scales. Interestingly, the facilitating effect of the replay is direction- and structure-dependent, as the replay with forward or scrambled reverse-structure of past experiences had little effect on grid cells, similar to the effect of no-replay model. Moreover, although the grid regularity can be formed through a comprehensive exploration of the environment without replay (see also Kang and Balasubramanian, 2019; Couey et al., 2013), a

minimal number of replays were sufficient to establish the periodic regularity of grid cells, echoing the discovery of the rapid visual mapping of grid cell for efficient adaptation to novel environments (Staudigl et al., 2018; Julian et al., 2018). This efficiency obtained from the reverse replay likely stems from the possibility that in navigation systems, representing both the direction of movement and its opposite may result in the construction of a global structure of environment through the formation of periodic representation centered on the self-location. This is important for animals navigating in environments with an irregular structure, such as a one-dimensional linear virtual track (Yoon et al., 2016) or a circular track (Jacob et al., 2019), where intrinsic grid cell periodicity was maintained.

An intriguing phenomenon revealed in our study is the selective enhancement of the spatial periodicity of small grid scales, as opposed to large grid scales, by the replay. This finding is consistent with a report of reaction-diffusion model (Yang et al., 2004), where the growth of Turing instability is accompanied with increased spatial frequency. However, further examination of the temporally differential dynamics of the spin-glass model under the modulation of the replay is required. In neuroscience, the present finding aligns nicely with the functional gradient along the anterior (ventral)-posterior (dorsal) axis of the hippocampus (Strange et al., 2014; Schiller et al., 2015). Specifically, the anterior hippocampus encodes coarse-grained experiences possibly represented by large grid scales, while the posterior hippocampus encodes fine-grained experiences by small grid scales. It is noteworthy that the posterior hippocampus is a common site where neuronal replay is recorded (Bendor & Wilson, 2012; Carr et al., 2012; Gillespie et al., 2021). Taken together, these findings suggest that the establishment of a high-resolution cognitive map of encoding episodes in the posterior hippocampus may be contingent upon the precise metric

of small grid scales, which is facilitated by the presence of the replay in the same region. This conjecture is supported by the observation that hippocampal replay is embedded within sharp wave ripples, transient high-frequency local field potentials (Buzsáki, 2015), where the precise locations are likely decoded by high-frequency oscillations (Ambrose et al., 2016). Therefore, our finding of decoupled relation of the replay to large grid scales and small grid scales may shed light on the neural anatomy of the hippocampus.

Another noteworthy aspect of the interaction between the replay and grid cells is the interplay between two anatomically separated regions, the HPC and ERC, which requires close collaboration between two regions. Indeed, the HPC and ERC are intensively connected as a whole system, with broad projections from the ERC terminating on the long axis of the HPC, including the subiculum and the CA1-3 fields (Witter and Amaral, 1991; van Groen et al., 2003). Our finding supplements the unilateral impact from the HPC to the ERC, in which the reverse replay of hippocampal place cells facilitates the formation of entorhinal grid cells, consistent with previous studies showing that lesions in the HPC disrupt the hexagonal pattern of entorhinal grid cells (Bonnevie et al., 2013). On the other hand, lesions in the ERC partially impair the spatial firing of hippocampal place cells (Hales et al., 2014; Schlesiger et al., 2015), demonstrating the impact of the ERC on the HPC. Indeed, past experiences are organized by the metric of grid cells to construct the cognitive map, which can in turn guide future movement indexed by the forward replay of place cells in the HPC (Wu and Foster, 2014; Pfeiffer, 2020). Therefore, it is possible the other unilateral connection from ERC to HPC serves as the grid metrics for guiding place cells in path integration, possibly through forward replay. In conclusion, the interplay of grid cells and place cells with both reverse and forward replays may be at the heart of the HPC-ERC system for

navigation in mental spaces, including the visual space (Nau et al., 2018), object space (Constantinescu., 2016), and reward space (Sosa and Giocomo, 2021).

Data availability

Raw data reported in the present study are available from the corresponding author upon reasonable request.

Code availability

Python and .NET code for grid cell simulation and virtual navigation are available at https://github.com/ZHANGneuro/Hippocampal-replay-facilitates-the-formation-of-entorhinal-grid-cells. Scripts for the analysis reported in the present study are available from the corresponding author upon reasonable request.

Conflict of interest

None declared

Author Contributions

B.Z. and J.L. conceived the study; B.Z. built the computational model and performed the analysis under the supervision of J.L.; B.Z. and J.L. wrote the paper.

Acknowledgments

This work was supported by the National Key R&D Program of China (2020AAA0105200), the National Natural Science Foundation of China (31861143039), Tsinghua University Guoqiang

Institute (2020GQG1016), Tsinghua University Qiyuan Laboratory, Beijing Academy of Artificial Intelligence (BAAI), and China Postdoctoral Science Foundation (2022M710470).

Supplemental Material

“Hippocampal replay facilitates the formation of entorhinal grid cells”

Bo Zhang and Jia Liu

Methods

Task and design

Spatial navigations were simulated in a 3D squared open environment created by Unity (Unity Technologies, San Francisco). For each simulation of navigations, a borderless novel environment was created by randomly placed trees, tree trunks, and stones on a flat grassy ground. A virtual gorilla acts as the animal and was arranged to perform either limited or extended explorations in the environment, during which gorilla's navigational experiences were passive driven by pre-determined allocentric directions, each navigational experience in a given direction was constituted by either 5 footsteps, 10 footsteps, or 15 footsteps with the distance of 1 virtual meter (resulting in 5, 10, 15 virtual meters for each experience, respectively), a pause location was placed in the gap between every two movement directions to simulate the animal decision point (Fig.2e-g). In the limited exploration, the number of directions were constrained to three with 45° apart (70°,115°, and 160° with 0° pointing towards East), the direction of each footstep was randomly selected. In the extended exploration, gorilla was arranged to move with hundreds of directions by following either the rat trajectories derived from open-source data (Hafting et al., 2005; https://www.ntnu.edu/kavli/research/grid-cell-data), or the trajectories predefined by one out of eight global directions with 45° apart (0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°). In the latter case, gorilla's direction in each footstep of each simulation were determined by the global direction with an additional noise randomly selected from a range of -90° to 90°. For example, given the simulation with a global direction of 90°, Gorilla's actual directions would be ranged from 0° - 180° (Fig. 3d).

Continuous attractor network

Spin-glass model (Fuhs and Touretzky, 2006), which has been termed as the continuous attractor network and widely validated in the entorhinal gird cells simulation (Burak and Fiete, 2009; Couey et al., 2013; Shipston-Sharman et al., 2016; Khona & Fiete, 2022), was coded by Torch (Version 1.13) for each navigational simulation with Gorilla's virtual trajectories. To examine the impact of hippocampal replay on the spatial periodicity of grid cells, spin-glass model allows us to intentionally turn on or turn off the hippocampal replay and quantify its efficiency on grid periodicity. To do it, we examined the models, in addition to the no-replay model, additionally with the reverse replay, the forward replay, and the reverse replay with direction-shuffled. In the reverse and forward replay, Gorilla's past experience, constituted by the 5-, 10-, or 15-location sequence, were manipulated by either direction-reversed or direction-consistent relative to actual experience, and were then added into model next or prior to the past experience, respectively. To examine the necessity of reverse replay, we extensively examined an direction-shuffled model, where the reversed directions were shuffled across the footsteps of past experience before adding it into model next to the actual experience. To ensure the stability of model inference, a total of 15 and 16 (twice for each global direction) models were conducted for the limited and extended explorations, respectively, with the number of footsteps were constrained by 1200 for all the models. One model was performed using rat trajectory, where the position sequence was down-sampled to 1200 data points.

In spin-glass model, the membrane voltage ξ, a two-dimension matrix with size 45-by-45 pixel corresponding to the layout of environment, was created using Python (version 3.9) to represent a single neuron i. The evolution of matrix ξ over time is governed by a differential equation, which demonstrates a recurrently updating process of membrane voltage ξ_i by the summation of the

firing rate of all potential neuron j via the excitation and inhibition weight matrix W underlying the velocity v, a time constant τ, and a gaussian noise ε (equation 1).

$$\tau \frac{d\xi_i}{dt} = -\xi_i + \sum_j W_{ij} f_j + v_i + \varepsilon \tag{1}$$

The firing rate f for neuron i is derived from the membrane voltage ξ via a nonlinear transfer function (equation 2), where η represents the threshold of ξ. The firing rate f for neuron i is inhibited to 0 if its membrane voltage is not exceeding the threshold, otherwise neuron i is excited with a firing rate transferred via the square root of membrane voltage ξ.

$$f_i = \begin{cases} \sqrt{\xi_i} & \xi_i > \eta \\ 0 & \zeta_i \leqslant \eta \end{cases} \tag{2}$$

The excitation and inhibition weight W, represented by a two-dimensional periodic matrix with the size of 45-by-45 pixels, determines the connection strength of each neuron $i - j$ pair (equation 3). The weight W was predefined by averaging 360 planar waves with each representing the head direction signal for a given direction (see Fig.S5 & Table.S1 for details). Specifically, one planar wave was defined by $\cos(\omega(\nu_x X + \nu_y Y))$. The term $\nu_x X + \nu_y Y$ represents a rotated two-dimensional activity gradient of the head direction pointing to the given spatial direction θ, where ν_x and ν_y are defined by $\cos(\theta)$ and $\sin(\theta)$.

$$\mathrm{W}_{\mathrm{ij}} = \sum_{\theta=0}^{359} \alpha \cos(\omega(\upsilon_{\mathrm{x}} \mathrm{X} + \upsilon_{\mathrm{y}} \mathrm{Y})) \tag{3}$$

The spatial frequency ω in equation 3 was determined by a finite sequence of grid scales σ, where $\sigma = \{10,13,19,26\}$. Each of grid scales was derived by computing the Euclidean distance between neuron i and j corresponding to the 45-by-45-pixel space. For example, the grid scale of 10 represents 10 pixels (approximately 3.3 virtual meters), which is the distance between the peaks of two adjacent firing fields of a grid cell. Those grid scales were configured with the ratio at approximately 1.4 (Kang and Balasubramanian, 2019).

$$\omega = 2pi/\sigma \tag{4}$$

During grid cell simulation, the velocity v in equation 1, the only input of model, was computed for each footstep based on the predefined allocentric movement directions for Gorilla and the footstep of 1 virtual meter in virtual environment (corresponding to 3 pixels in neural space).

<u>Edge effect of grid cells</u>

The edge effect is an intrinsic issue of spin-glass model, which brings the distortion of hexagonal pattern of grid cells around the edge of space because the progressively weaker weights would be not able to reinforce the grid fields while moving toward the edge (Fuhs and Touretzky, 2006). Samsonovich and McNaughton (1997) deal with the issue of edge effects by hypothesizing a wrapped around connection of the neurons close to the edges, that is, the edges of neural space could be connected to each other to form a toroidal attractor. In this case, when an existing firing field of grid cell reaches the edge of space and moves away, while on the opposite side of space a "new" firing field was formed due to the excitation-inhibition dynamics of synaptic interactions. In further, grid scales computationally determines when the newly-formed firing field moves in

the space after the existing firing field moves out. Inspired by the property of spatial periodicity, the firing strength for the new-coming firing field could be predicted by the phase of a given grid scale. Therefore, in the present study, we temporally solved the edge effects of the grid cells generated by spin-glass model by interpolation (Fig.S6). Specifically, for a given grid pattern, the grid orientation was first derived by computing the autocorrelation map. Second, six coordinates along the three main axis of grid pattern were computed for each target pixel underlying trigonometric function. Finally, one coordinate was randomly selected and the firing strength in the coordinate was interpolated into phase of the newly-generated firing field after excluding the coordinates outside of space. With this procedure, grid cell population could be derived by phase-shifting the hexagonal pattern of single neuron.

<u>Validation of of grid periodicity</u>

The spatial periodicity of gird cell, as indicated by Gridness, was computed by following the procedure reported previously (Sargolini et al., 2006; Barry & Burgess, 2017). Specifically, for each neuron generated by the models, the spatial autocorrelation map was estimated by equation 5, where $\lambda(x, y)$ denote the activity strength at location (x, y), while τ_x and τ_y denote the spatial shifts of the firing pattern from the neuron. With the rotation of autocorrelation map by 30°, 60°, 90°, 120°, and 150°, Gridness was computed as the difference between the lowest correlation of 60° and 120° and the highest correlations of 30°, 90° and 150°.

$$r(\tau_x,\tau_y) = \frac{n\sum\lambda(x,y)\,\lambda(x-\tau_x,y-\tau_y) - \sum\lambda(x,y)\,\lambda(x-\tau_x,y-\tau_y)}{\sqrt{n\sum\lambda(x,y)^2(\sum\lambda(x,y))^2}\sqrt{n\sum\lambda(x-\tau_x,y-\tau_y)^2(\sum\lambda(x-\tau_x,y-\tau_y))^2}} \quad (5)$$

For the regularity of grid cell population, we examined the toroidal structure of population activity of grid cells using persistent homology (Ghrist, 2008), which was performed by the package Ripser++ (Version 1.1.2) for GPU acceleration (Zhang et al., 2020; Bauer, 2021). The principle behind persistent homology is to identify the toroidal characteristics from the global topological feature constituted by point-clouds (the firing maps of grid cell) in 3D Euclidean space. Specifically, we followed the outlines published previously (Gardner et al., 2022). At each grid scale, principal component analysis was performed on the population activity of grid cells, then the first six principal components were further dimension-reduced by uniform manifold approximation into 3D projection. In the 3D space, the data points were Z transformed and were replaced with simplicial complexes. Upon increasing the radius of simplicial complexes, the complexes are gradually connected and form certain global structure. To examine the emergence of toroidal structure, the life spans (the barcodes) of two 1D holes (the inner and outer hole of torus) and one 2D hole (the cavity in the interior of torus) were tracked throughout the duration of holes from it appear to disappear. Assuming that the life time of a given hole were long enough, the hole would be considered to be contained robustly by the global structure of connected complexes.

Statistics

The significance of grid periodicity was examined for each of grid scale and each model. To do it, we used the field-shuffling permutation (Fig.S8; Krupic et al., 2012), which was performed based on the grid cells generated by reverse-replay model, which show the regular hexagonal patterns for all grid scales. Specifically, for each neuron from grid cell population, the long-range spatial structure (the spatial relationship of firing fields) was disrupted by shuffling the firing

fields 100 times with the local structure (the pattern of firing field) preserved. After the permutation distribution was produced for the given neuron, uncorrected threshold was derived from the 95th percentile of the Gridness distribution. Next, the threshold corrected for multiple comparison was determined again by the 95th percentile of the uncorrected thresholds derived from grid cell population. Similar procedure was used for persistent cohomology analysis. For each grid scale, the initial uncorrected threshold and the threshold corrected for multiple comparison was indicated by the 99th and 95th percentile of the barcode distribution from single and population neurons, respectively. The significance thresholds of toroidal structure for 1D hole (H1) and 2D hole (H2) were averaged across grid scales. The Gridness computed from each of grid scales and models were compared using two sample t-test, all the t-tests used in the analysis were two-tailed.

Figure S1. Grid cell population generated by spin-glass model using animal trajectories acquired from open-source data (Hafting et al., 2005). For each scale, 24 grid cell samples were shown as example and were ranked by their horizontal and vertical phases.

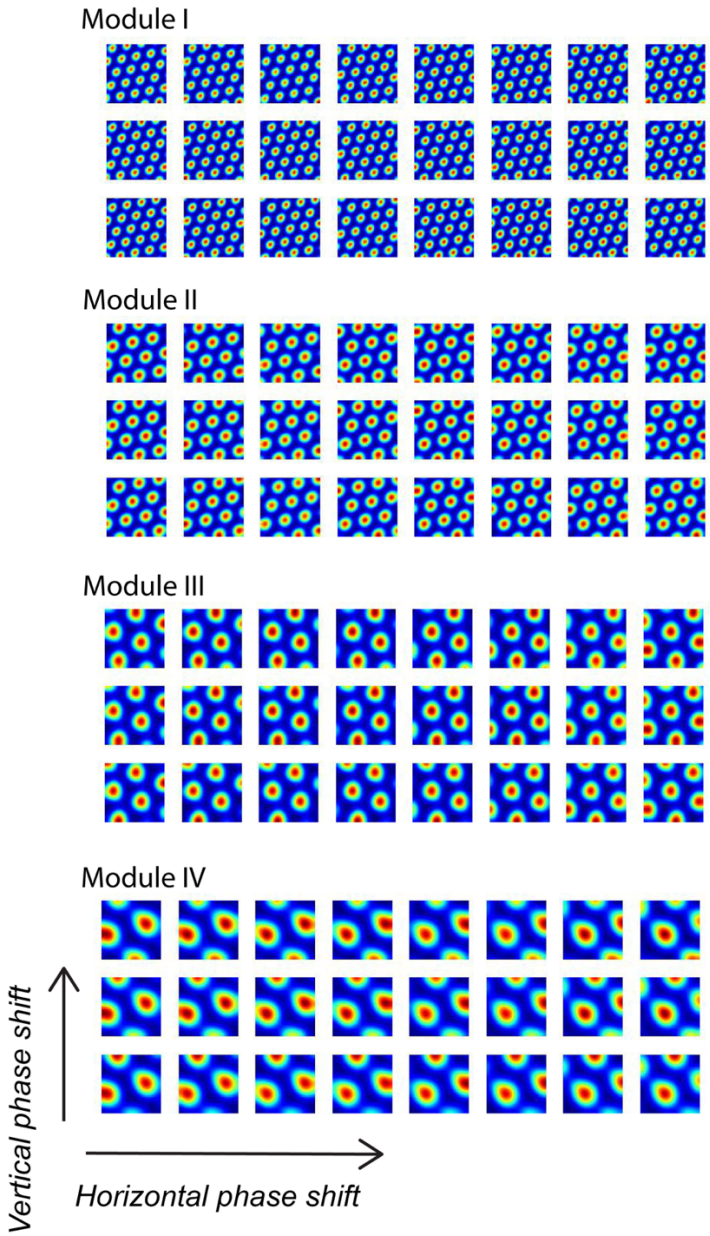

Figure S2. No effect of the movement distance prior to pause location were found on the spatial periodicity of grid cells in either Mod I ($t(14) = 1.25$, $p = 0.24$), Mod II ($t(14) = 0.90$, $p = 0.56$), Mod III ($t(14) = 0.34$, $p = 0.99$), or Mod IV ($t(14) = 0.34$, $p = 0.99$). Black line represents the mean curve across grid scales. Shaded area denotes standard error. All tests were two-sided.

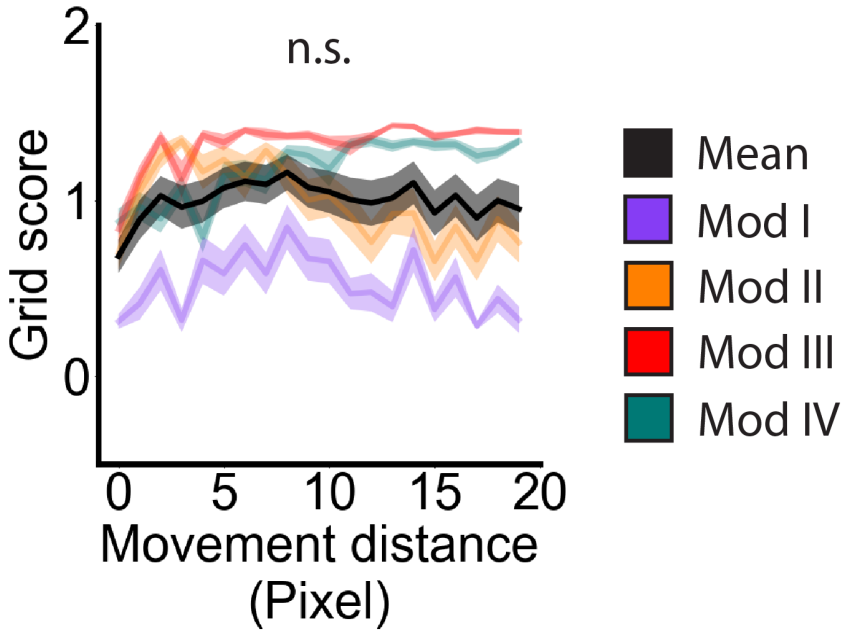

Figure S3. a, b) Extended comparison of Gridness scores in corresponding to Fig. 2 and 3. (15 independent simulations for each condition).

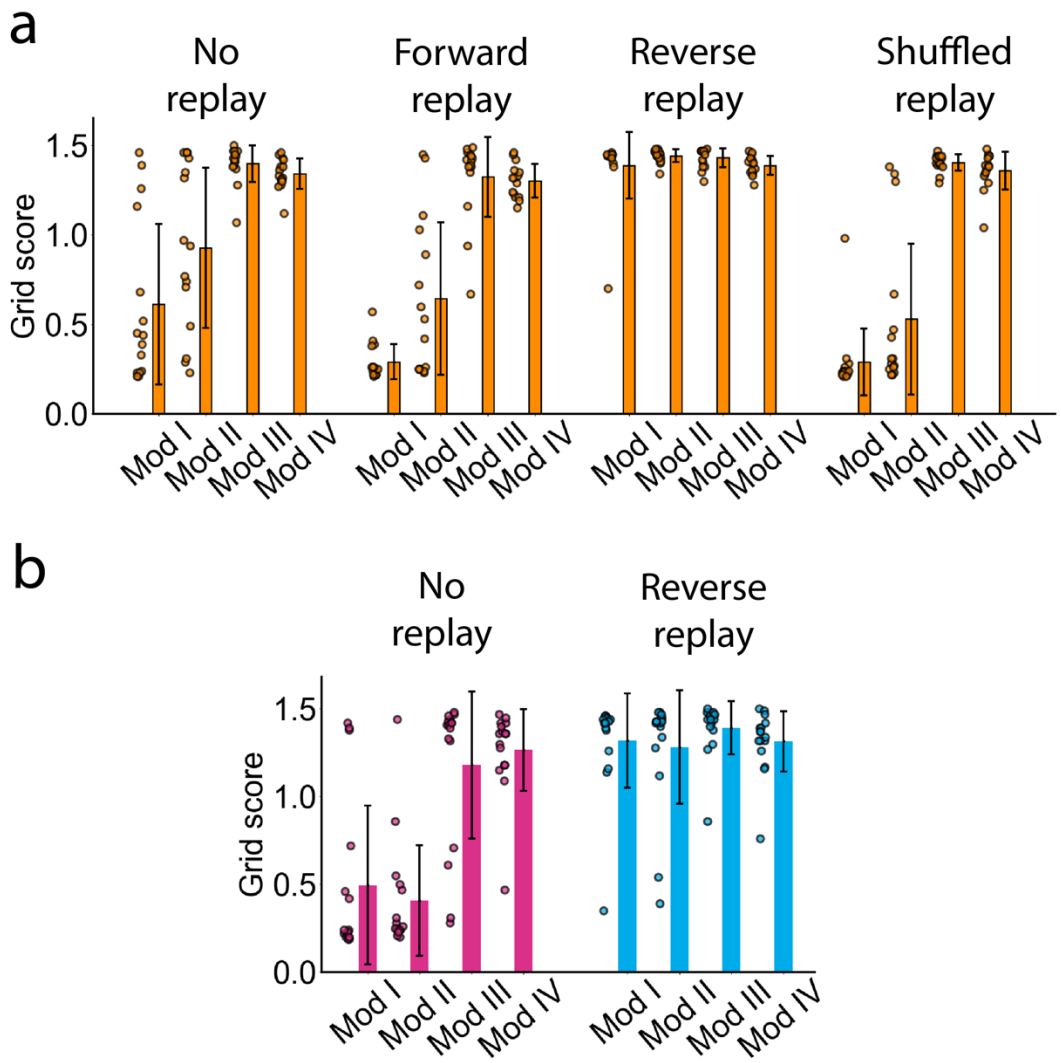

Figure S4. Persistent cohomology analysis on the population activity (484 grid cells) of each grid scale, which were derived from the model with reverse replay applied shown in Fig.2. The barcodes with 20 longest lifetimes were shown for the 1D hole (H1, blue) and 2D hole (H2, green), red dash lines indicate the threshold of significance corrected for multiple comparison.

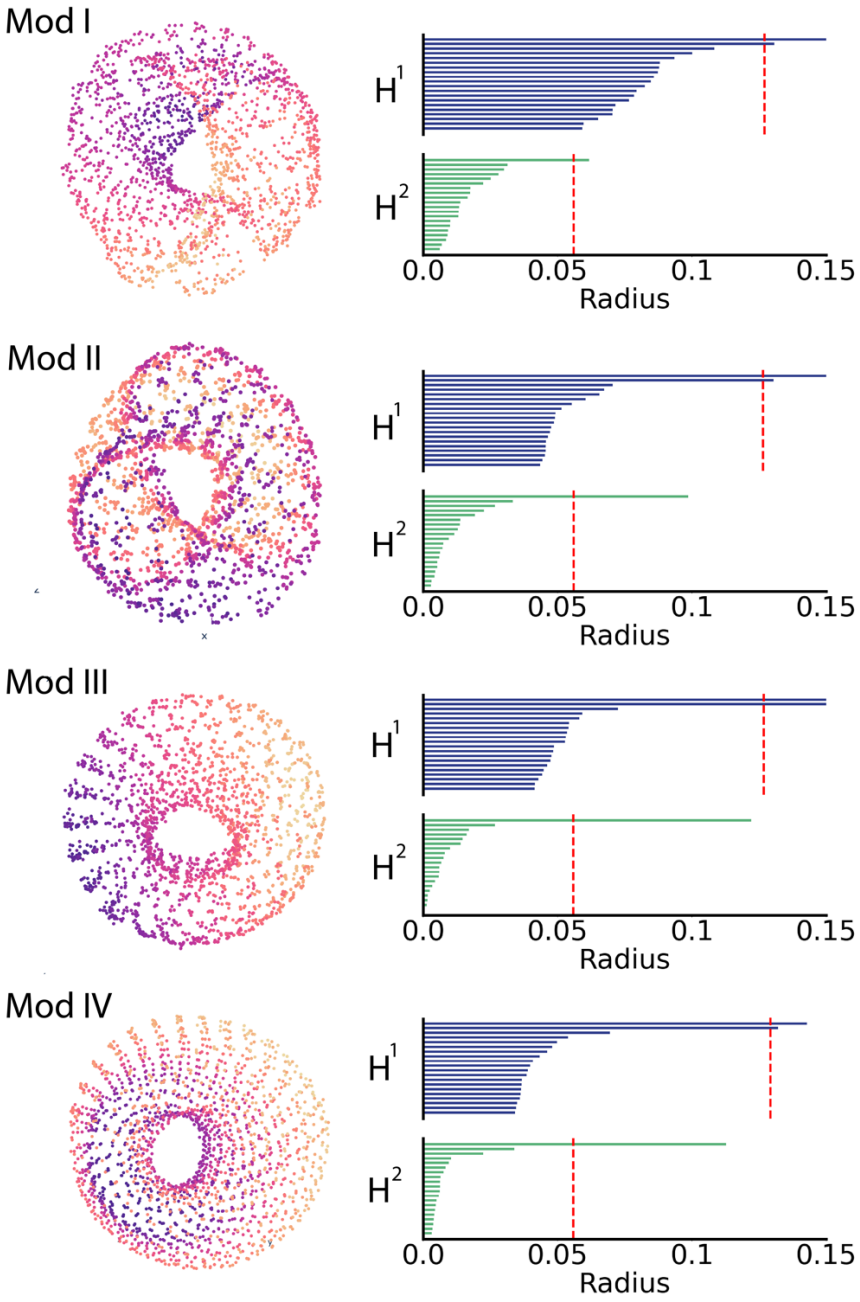

Figure S5. Illustration of the formation of grid cell using spin-glass model. a) Example of the weight matrix (grid Mod I), which represents the conjunctive head directions of 0° - 360° for a given allocentric location. Each head direction was embedded by a planar wave that represents the unique direction across locations on the two-dimensional space. b) the formation of hexagonal pattern with model iterations (T), regular grid periodicity could be seen by T = 60.

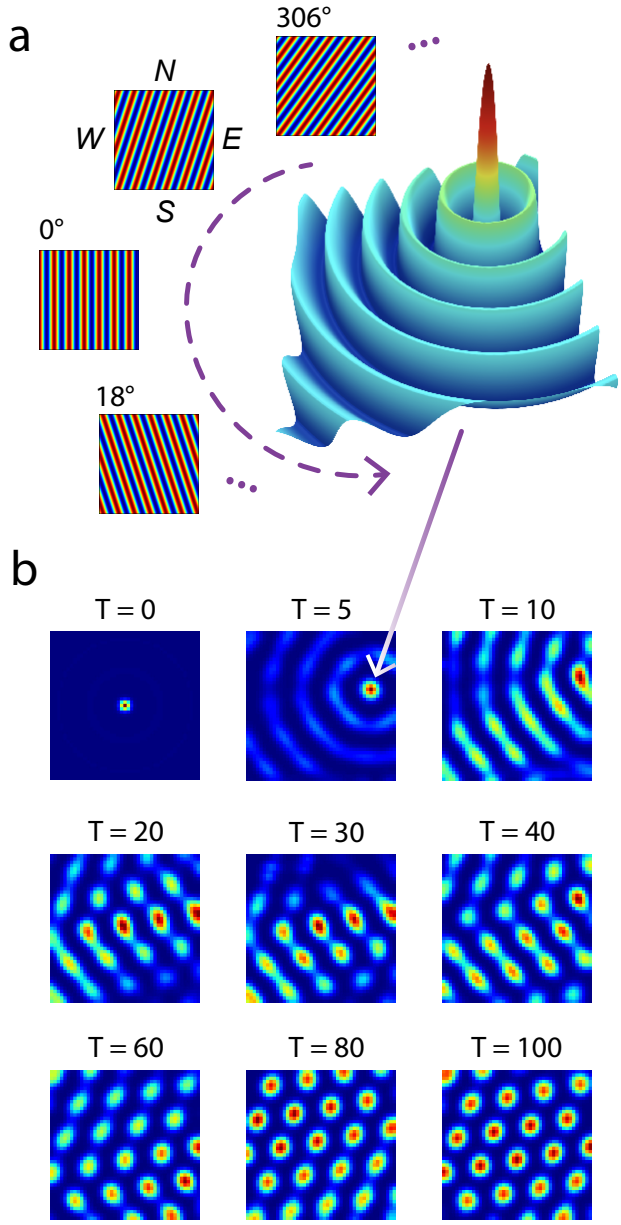

Figure S6. Simulation of grid cell population using interpolation. a) the target grid cell, b) the grid orientation was computed from the autocorrelation map of target grid cell, c) an example of interpolation procedure. Given the twisted nature of grid cell, each pixel in the blank area (colored by white) was filled after the phase shift of grid cell towards North by coping the value from one of six adjacent locations computed using trigonometric function. d) the simulation of grid cell with phase shifted relative to the target.

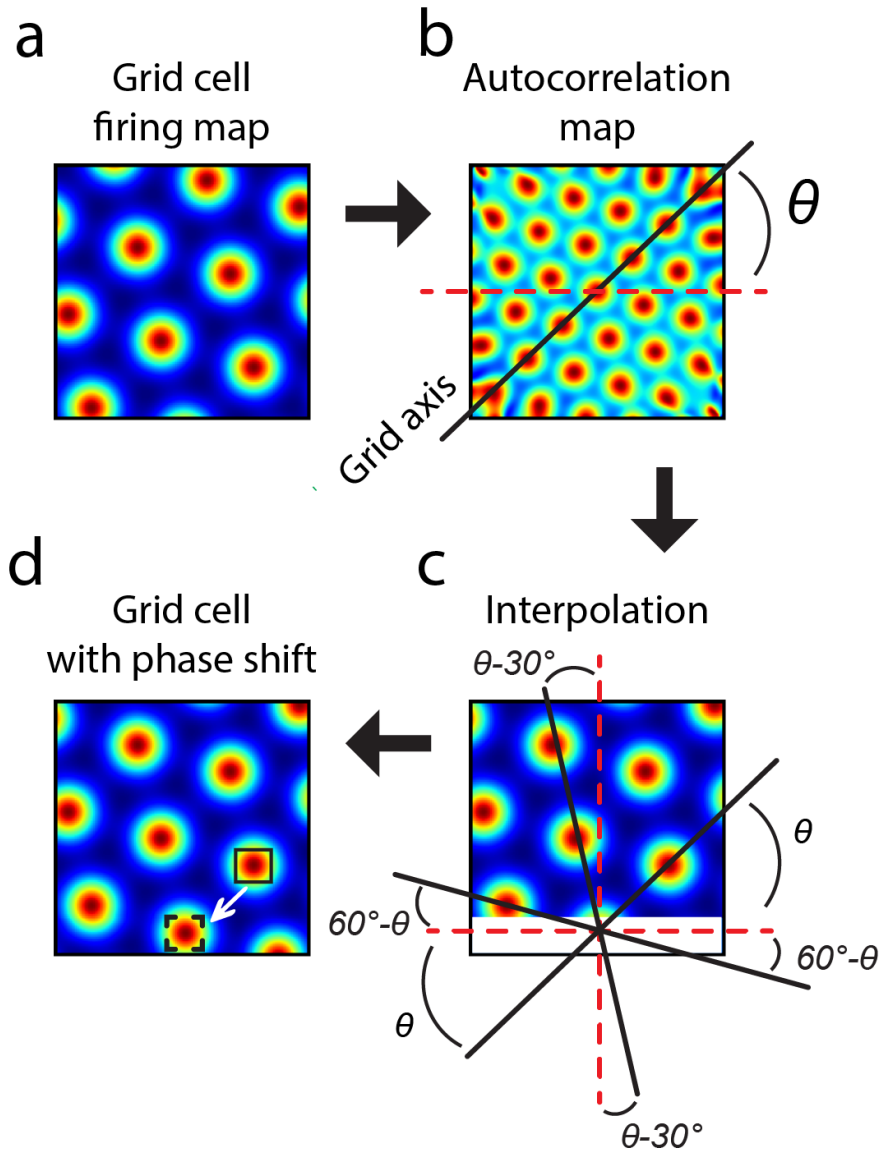

Figure S7. Sixteen independently simulated virtual navigations for each of no-replay and reverse-replay model. Each virtual navigation was dominated by a unique global direction with the noise ranged from -90° to 90° for each footstep. Left column, the distribution of movement directions, black solid line indicates the mean of directions, Right column, gorilla trajectories during virtual navigation.

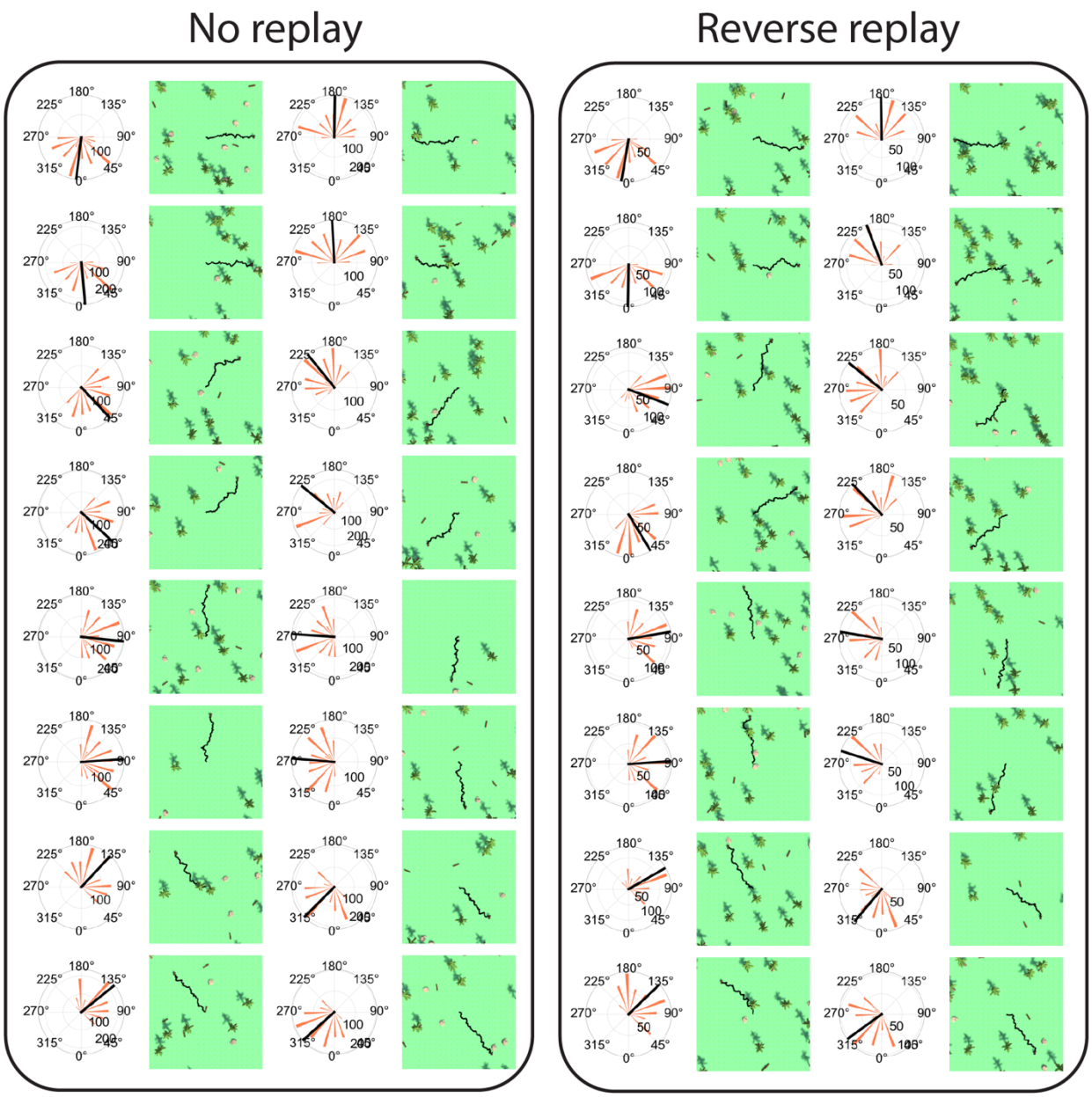

Figure S8. Procedure of field-shuffling permutation. 1) a grid cell example, 2) peaks of the firing field of grid cell, 3) local structure of grid cell identified by Voronoi polygons, 4,5) one firing field that was randomly re-positioned during field-shuffling, (6) shuffled firing pattern of grid cell, (7) Gridness distribution acquired from 100 field-shuffling, black dash line represents the initial (uncorrected) threshold at the 99th percentile of Gridness distribution, (8) the uncorrected Gridness distribution derived from grid cell population, black dash line represents the corrected threshold for multiple comparison at the 95th percentile of the distribution.

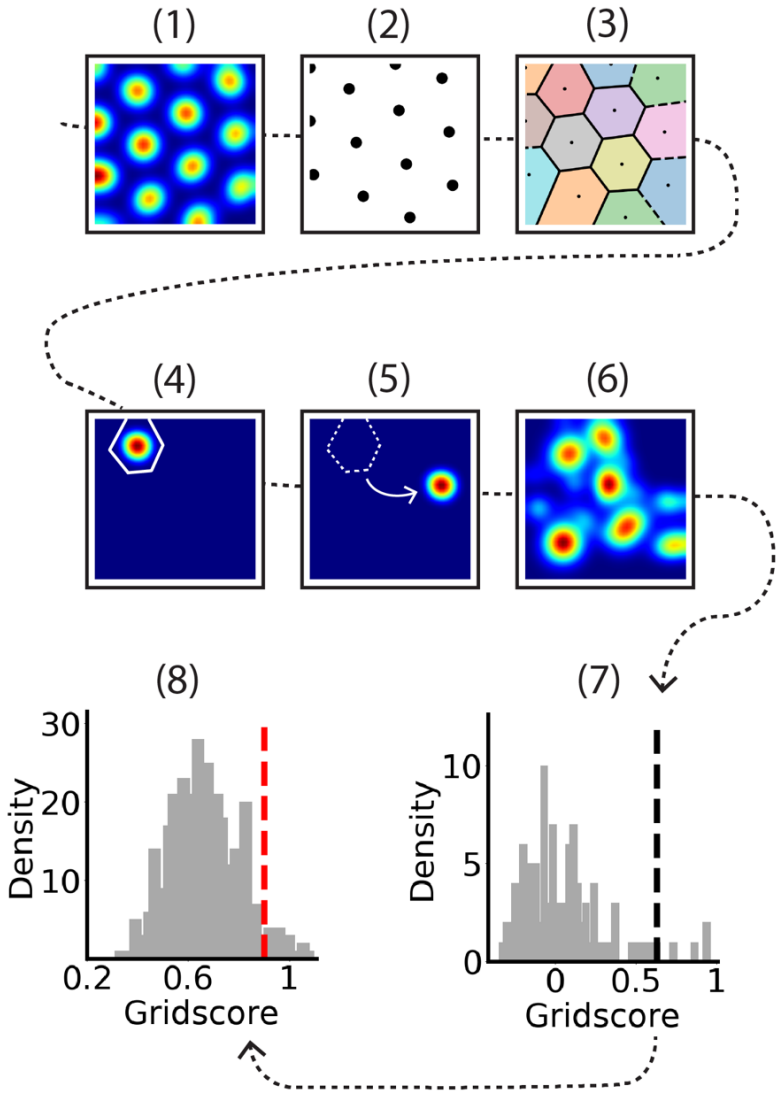

Table S1. Parameters of spin-glass model.

Parameter	Variable	Value
Membrane voltage threshold	η	0
Time constant	τ	10
Weight matrix	W	
Positive peak		0.003
Negative peak		-0.001
Mean		0
Amplitude of planar wave	α	0.333
Grid scales	$\vec{\sigma}$	{10, 13, 19, 26}
Space size	N	45 pixels

Reference

Ambrose, R. E., Pfeiffer, B. E., & Foster, D. J. (2016). Reverse replay of hippocampal place cells is uniquely modulated by changing reward. Neuron, 91(5), 1124-1136.

Barry, C., & Burgess, N. (2017). To be a Grid Cell: Shuffling procedures for determining "Gridness". BioRxiv, 230250.

Bauer, U. (2021). Ripser: efficient computation of Vietoris–Rips persistence barcodes. Journal of Applied and Computational Topology, 5(3), 391-423.

Bendor, D., & Wilson, M. A. (2012). Biasing the content of hippocampal replay during sleep. Nature neuroscience, 15(10), 1439-1444.

Burak, Y., & Fiete, I. R. (2009). Accurate path integration in continuous attractor network models of grid cells. PLoS computational biology, 5(2), e1000291.

Barry, C., Ginzberg, L. L., O'Keefe, J., & Burgess, N. (2012). Grid cell firing patterns signal environmental novelty by expansion. Proceedings of the National Academy of Sciences, 109(43), 17687-17692.

Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. Neuron, 100(2), 490-509.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. Neuron, 35(4), 625-641.

Bonnevie, T., Dunn, B., Fyhn, M., Hafting, T., Derdikman, D., Kubie, J. L., ... & Moser, M. B. (2013). Grid cells require excitatory drive from the hippocampus. Nature neuroscience, 16(3), 309-317.

Buzsáki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. Hippocampus, 25(10), 1073-1188.

Bush, D., Barry, C., Manson, D., & Burgess, N. (2015). Using grid cells for navigation. Neuron, 87(3), 507-520.

Couey, J. J., Witoelar, A., Zhang, S. J., Zheng, K., Ye, J., Dunn, B., ... & Witter, M. P. (2013). Recurrent inhibitory circuitry as a mechanism for grid formation. Nature neuroscience, 16(3), 318-324.

Carr, M. F., Jadhav, S. P., & Frank, L. M. (2011). Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. Nature neuroscience, 14(2), 147-153.

Carr, M. F., Karlsson, M. P., & Frank, L. M. (2012). Transient slow gamma synchrony underlies hippocampal memory replay. Neuron, 75(4), 700-713.

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. (2016). Organizing conceptual knowledge in humans with a gridlike code. Science, 352(6292), 1464-1468.

Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature, 440(7084), 680-683.

Fiete, I. R., Burak, Y., & Brookings, T. (2008). What grid cells convey about rat location. Journal of Neuroscience, 28(27), 6858-6871.

Fuhs, M. C., & Touretzky, D. S. (2006). A spin glass model of path integration in rat medial entorhinal cortex. Journal of Neuroscience, 26(16), 4266-4276.

Gardner, R. J., Hermansen, E., Pachitariu, M., Burak, Y., Baas, N. A., Dunn, B. A., ... & Moser, E. I. (2022). Toroidal topology of population activity in grid cells. Nature, 602(7895), 123-128.

Gillespie, A. K., Maya, D. A. A., Denovellis, E. L., Liu, D. F., Kastner, D. B., Coulter, M. E., ... & Frank, L. M. (2021). Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. Neuron, 109(19), 3149-3163.

Ghrist, R. (2008). Barcodes: the persistent topology of data. Bulletin of the American Mathematical Society, 45(1), 61-75.

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. Nature, 436(7052), 801-806.

Hales, J. B., Schlesiger, M. I., Leutgeb, J. K., Squire, L. R., Leutgeb, S., & Clark, R. E. (2014). Medial entorhinal cortex lesions only partially disrupt hippocampal place cells and hippocampus-dependent place memory. Cell reports, 9(3), 893-901.

Jacob, P. Y., Capitano, F., Poucet, B., Save, E., & Sargolini, F. (2019). Path integration maintains spatial periodicity of grid cell firing in a 1D circular track. Nature Communications, 10(1), 1-13.

Julian, J. B., Keinath, A. T., Frazzetta, G., & Epstein, R. A. (2018). Human entorhinal cortex represents visual space using a boundary-anchored grid. Nature neuroscience, 21(2), 191-194.

Khona, M., & Fiete, I. R. (2022). Attractor and integrator networks in the brain. Nature Reviews Neuroscience, 1-23.

Kang, L., & Balasubramanian, V. (2019). A geometric attractor mechanism for self-organization of entorhinal grid modules. Elife, 8.

Krupic, J., Burgess, N., & O'Keefe, J. (2012). Neural representations of location composed of spatially periodic bands. Science, 337(6096), 853-857.

Liu, Y., Dolan, R. J., Kurth-Nelson, Z., & Behrens, T. E. (2019). Human replay spontaneously reorganizes experience. Cell, 178(3), 640-652.

Moser, E. I., Moser, M. B., & McNaughton, B. L. (2017). Spatial representation in the hippocampal formation: a history. Nature neuroscience, 20(11), 1448-1464.

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

Nau, M., Schröder, T. N., Bellmund, J. L., & Doeller, C. F. (2018). Hexadirectional coding of visual space in human entorhinal cortex. Nature neuroscience, 21(2), 188-190.

Nour, M. M., Liu, Y., Arumuham, A., Kurth-Nelson, Z., & Dolan, R. J. (2021). Impaired neural replay of inferred relationships in schizophrenia. Cell, 184(16), 4315-4328.

Ormond, J., & O'Keefe, J. (2022). Hippocampal place cells have goal-oriented vector fields during navigation. Nature, 607(7920), 741-746.

Pfeiffer, B. E. (2020). The content of hippocampal "replay". Hippocampus, 30(1), 6-18.

Suh, J., Foster, D. J., Davoudi, H., Wilson, M. A., & Tonegawa, S. (2013). Impaired hippocampal ripple-associated replay in a mouse model of schizophrenia. Neuron, 80(2), 484-493.

Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S., & Ranganath, C. (2015). Memory and space: towards an understanding of the cognitive map. Journal of Neuroscience, 35(41), 13904-13911.

Strange, B. A., Witter, M. P., Lein, E. S., & Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. Nature Reviews Neuroscience, 15(10), 655-669.

Samsonovich, A., & McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. Journal of Neuroscience, 17(15), 5900-5920.

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M. B., & Moser, E. I. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. Science, 312(5774), 758-762.

Sosa, M., & Giocomo, L. M. (2021). Navigating for reward. Nature Reviews Neuroscience, 1-16.

Shipston-Sharman, O., Solanka, L., & Nolan, M. F. (2016). Continuous attractor network models of grid cell firing based on excitatory–inhibitory interactions. The Journal of physiology, 594(22), 6547-6557.

Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. Science, 364(6447), eaaw5181.

Stemmler, M., Mathis, A., & Herz, A. V. (2015). Connecting multiple spatial scales to decode the population activity of grid cells. Science Advances, 1(11), e1500816.

Staudigl, T., Leszczynski, M., Jacobs, J., Sheth, S. A., Schroeder, C. E., Jensen, O., & Doeller, C. F. (2018). Hexadirectional modulation of high-frequency electrophysiological activity in the human anterior medial temporal lobe maps visual space. Current Biology, 28(20), 3325-3329.

Schlesiger, M. I., Cannova, C. C., Boublil, B. L., Hales, J. B., Mankin, E. A., Brandon, M. P., ... & Leutgeb, S. (2015). The medial entorhinal cortex is necessary for temporal organization of hippocampal neuronal activity. Nature neuroscience, 18(8), 1123-1132.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. Nature reviews neuroscience, 8(9), 657-661.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: role of the hippocampus. Hippocampus, 8(3), 198-204.

Tolman, E. C. (1948). Cognitive maps in rats and men. Psychological review, 55(4), 189.

van Groen, T., Miettinen, P., & Kadish, I. (2003). The entorhinal cortex of the mouse: organization of the projection to the hippocampal formation. Hippocampus, 13(1), 133-149.

Vaz, A. P., Wittig Jr, J. H., Inati, S. K., & Zaghloul, K. A. (2020). Replay of cortical spiking sequences during human memory retrieval. Science, 367(6482), 1131-1134.

Wu, X., & Foster, D. J. (2014). Hippocampal replay captures the unique topological structure of a novel environment. Journal of Neuroscience, 34(19), 6459-6469.

Witter, M. P., & Amaral, D. G. (1991). Entorhinal cortex of the monkey: V. Projections to the dentate gyrus, hippocampus, and subicular complex. Journal of Comparative Neurology, 307(3), 437-459.

Yoon, K., Lewallen, S., Kinkhabwala, A. A., Tank, D. W., & Fiete, I. R. (2016). Grid cell responses in 1D environments assessed as slices through a 2D lattice. Neuron, 89(5), 1086-1099.

Yang, L., Zhabotinsky, A. M., & Epstein, I. R. (2004). Stable squares and other oscillatory Turing patterns in a reaction-diffusion model. Physical review letters, 92(19), 198303.

Zhang, S., Xiao, M., & Wang, H. (2020). GPU-accelerated computation of Vietoris-Rips persistence barcodes. arXiv preprint arXiv:2003.07989.