Replay as structural inference in the hippocampal-entorhinal system

Talfan Evans[123,#] and Neil Burgess[12,#]

1 Institute of Cognitive Neuroscience, UCL

2 Institute of Neurology, UCL

3 CoMPLEX, UCL

#correspondence

Abstract

Model-based decision making relies on the construction of an accurate representation of the underlying state-space, and localization of one's current state within it. One way to localize is to recognize the state with which incoming sensory observations have been previously associated. Another is to update a previous state estimate given a known transition. In practice, both strategies are subject to uncertainty and must be balanced with respect to their relative confidences; robust learning requires aligning the predictions of both models over historic observations. Here, we propose a dual-systems account of the hippocampal-entorhinal system, where sensory prediction errors between these models during *online* exploration of state space initiate *offline* probabilistic inference. *Offline* inference computes a *metric* embedding on grid cells of an *associative* place graph encoded in the recurrent connections between place cells, achieved by message passing between cells representing non-local states. We provide testable explanations for coordinated place and grid cell 'replay' as efficient message passing, and for distortions, partial rescaling and direction-dependent offsets in grid patterns as the confidence weighted balancing of model priors, and distortions to grid patterns as reflecting inhomogeneous sensory inputs across states.

Author Summary

- Minimising prediction errors between transition and sensory input (observation) models predicts partial rescaling and direction-dependent offsets in grid cell firing patterns.
- Inhomogeneous sensory inputs predict distortions of grid firing patterns during *online* localisation, and local changes of grid scale during *offline* inference.
- Principled information propagation during *offline* inference predicts coordinated place and grid cell 'replay', where sequences propagate between structurally related features.

Introduction

Grid cells in the medial entorhinal cortex (mEC), whose firing fields form a periodic hexagonal lattice across the environment, are thought to support path integration[1–3], whereas hippocampal place cells tend to have unimodal firing fields reflecting environmental cues such as boundaries[4,5]. Grid cell firing patterns are stable over time, suggesting corrective environmental inputs, possibly from place cells[6], but rely more on self-motion than place cell firing patterns[7] suggesting that environmental inputs are not fully corrective[2,8,9]. Given estimates of each input's uncertainty, Bayesian inference tells us how they should be optimally combined.

Although *online* learning (i.e. using only currently available sensory information) can converge under low PI and sensory noise, robust learning in the presence of noise requires minimizing the error between self-motion and environmental estimates of location across all state transitions[10,11]. Thus, historic observations must be stored and revisited *offline* (i.e. independently of current sensory inputs) to allow propagation of local environmental information to non-local but structurally connected regions of the cognitive map, e.g. as when adapting to a novel shortcut or barrier. This process can also be viewed as an embedding of sensory experience within a low-dimensional manifold (in this case, 2D space), as observed of place cells during sleep[12].

Building on previous work[11,13–15], we propose a dual-systems (*online-offline*) account of spatial inference in the hippocampal/entorhinal system, which we define as the process of identifying the

configuration of both one's own location (current state) and the location of environmental landmarks in space (c.f. 'SLAM'[10]). In familiar environments, *online* localization (identification of one's own position) is achieved by recursively combining self-motion and sensory inputs, which are mediated by learned *transition* and *observation* models, respectively. However, prediction errors between these models trigger *offline* inference over non-local states, facilitating fast learning of new or changed associative environmental structure, encoded *online* in place-place cell synaptic associations. We identify this *offline* inference with coordinated hippocampal/mEC 'replay'[16–20].

Our framework also provides algorithmic- and implementation-level explanations for observed features of grid cell firing in response to manipulations[7,9,21–24] or inhomogeneity[25–29] of environmental sensory input. Overall, these phenomena can be understood in a probabilistic framework, where minimization of prediction errors between the transition and observation models are traded against prior model beliefs.

Results

Probabilistic *online* localization with place and grid cells

Grid cells exist in 'modules' of cells, whose firing patterns have the same spatial scale and orientation relative to the environment, but differ in their spatial offsets[30]. The spatial scale increases in discrete steps along the dorso-ventral axis, suggesting that, across modules, GCs support a hierarchical representation of space[21,24,31–33]. Here, we consider a single module of GCs, whose activity represent a probability distribution over a periodic, discretized region of space (visualised as a topographically arranged sheet of cells; Fig. 1A).

The self-location distribution is maintained over time by recursively integrating sensory and self-motion inputs, accounting for their uncertainties (Fig. 1A). Firstly, the posterior distribution over agent location (grid module activity) from the previous time-step $\boldsymbol{G}$ is updated given noisy perceived movement $\hat{\boldsymbol{u}}$ via the *transition model* T (see Methods):

$$G'(\boldsymbol{x}) = \int T(\boldsymbol{x}, \boldsymbol{x}' | \hat{\boldsymbol{u}}, \boldsymbol{\alpha}) \cdot G(\boldsymbol{x}') d\boldsymbol{x}' \quad (1)$$

where $\boldsymbol{x}$ is the 2D coordinate of the agent location in *metric* space (corresponding to a particular grid cell) and $\boldsymbol{x}'$ the location at the previous time-step. Biophysically, T would be represented by a population of direction dependent 'shifter' cells with asymmetric recurrent weights[1] with a circulant structure[34] (Fig. S1C, see Methods) learned *apriori* (but see Refs. [35–37]). The rate of translation of activity on the grid sheet in response to movement $\hat{\boldsymbol{u}}$ is controlled by the *transition model* gain $\boldsymbol{\alpha} = [\alpha_x, 0; 0, \alpha_y]$, which might correspond to the strength of the associations to, or the speed dependence of, shifter or conjunctive cells[31,38,39] (see Supplementary Methods).

The *transition model* estimate $G'(\boldsymbol{x})$ is then refined by observations of environmental features, which map to *metric* space locations via *observation model* H:

$$G(\boldsymbol{x}) = \frac{1}{K}\, H(\boldsymbol{P}|\boldsymbol{x}) \cdot G'(\boldsymbol{x}) \tag{2}$$

where $\boldsymbol{P}$ is a vector of place cell firing rates, firing of place cell i representing the likelihood of the presence of a specific sensory feature, or combination thereof. In our simulations, these have unique locations in physical space $\boldsymbol{\mu}_i$ and receptive field widths $\boldsymbol{\Sigma}_{PC}^{i}$. Where the number of grid cells is large, the weights from place cell i to the grid module define a distribution for that feature's estimated location in *metric* space (Fig. 1A). The weighted projection of place cell activity by these weights defines the *observation model* H. K provides inhibitory normalization (see Supplementary Methods).

Online learning modifies the *observation model* to reflect the current *transition model* location estimate (induces synaptic changes in the place-grid cell connection weights via a BCM rule; Fig. 2A; Methods). *Online* learning produces stable grid patterns (due to the circulant structure of T) for a range of levels of PI and sensory noise, but convergence fails in higher noise regimes (Fig. S2B). After a short period of initial learning, stable grid patterns emerge in the integrated estimate, despite the pure PI estimate being too noisy and the sensory associations too immature to drive stable patterns, if operating independently (Fig. 1B).

Offline inference: The hippocampus as a probabilistic graph

Local, *online* learning is not robust in novel environments, because corrections to the estimated agent location (current grid cell activity), e.g. upon encountering familiar environmental features (place cell activity) associated to a different location on the grid module, also imply corrections to the encoding of feature observations along the preceding trajectory[10]. That is to say; local updates to the *cognitive map* also imply non-local, structurally associated changes. Formally, probabilistic spatial inference in this case requires finding the most likely configuration of *metric* space feature locations $\{\boldsymbol{b}_i\}_{i=1:N_P}$ (each $\boldsymbol{b}_i$ is the 2D coordinate of feature i in *metric* space, i.e. a place - grid cell association) and agent location $\boldsymbol{x}$ (the distribution over which is indicated by the grid cell firing rates) consistent with environmental sensory observations made along a given trajectory.

Theoretically, the configuration $\{\boldsymbol{b}_i\}_{i=1:N_P}$ can be recovered purely from the distances between pairs of environmental features[10] (Fig. 1F, *"square"*, *"ring"*). Importantly, despite both feature locations being susceptible to large absolute errors (due to noisy PI), the errors will be correlated such that pairwise distance measurements will decrease in variance with observations[10]. This method predicts characteristic failure modes when the pairwise distance information is ambiguous or incomplete (e.g., Fig. 1F; *"broken ring"*). New distance observations might also cause dramatic changes to the inferred configuration (e.g. the discovery of a shortcut). If the current *absolute* estimates of feature node (i.e. place cell) location are stored in the place-grid cell synaptic weights, we propose that the *relative distances* between pairs of features are stored in the recurrent weights between place cells in hippocampal region CA3.

Consider a spring network, where the edge between environmental feature nodes i and j represents a noisy pairwise observation with length reflecting pairwise distance and stiffness reflecting certainty[13,40] (Fig. 1F, S6A). Minimizing the elastic energy in the spring mesh system corresponds to finding the maximum of the joint likelihood $L(\cdot)$, which is a function of the feature locations in *metric* space $\{\boldsymbol{b}_i\}_{i=1:N_P}$ and the internal gain parameter $\boldsymbol{\alpha}$, given pairwise distance measurements with

Gaussian noise (δ_{ij}). This is equivalent to defining a probabilistic graphical model (see Methods) over the posterior:

$$L\left(\boldsymbol{b}_1, \dots, \boldsymbol{b}_{N_p}, \boldsymbol{\alpha}\right) \propto p(\boldsymbol{\alpha}|\boldsymbol{\alpha}_0) \prod_{(i,j)\in E} \psi\left(\boldsymbol{b}_i - \boldsymbol{b}_j, \boldsymbol{\alpha}\middle|\delta_{ij}\right) \prod_{i=1:N_p} B_i(\boldsymbol{b}_i) \tag{3}$$

where the current PC-GC weights B (B_i is the probability distribution of the location of feature i and $B_i(\boldsymbol{b}_i)$ is its value at *metric*/grid module location $\boldsymbol{b}_i$) act as priors on the feature node locations, the pairwise potential terms $\psi(\cdot)$ penalize the difference between *associative* pairwise distance measurements δ_{ij} made directly in environmental *stimulus* space, and the distance between their candidate locations in *metric* space $\boldsymbol{b}_i - \boldsymbol{b}_j$. E is the set of connected PCs (see Methods). Distance in *metric* space is also a function of the *transition model* gain ($\boldsymbol{\alpha}$), which has a Gaussian prior $p(\boldsymbol{\alpha}|\boldsymbol{\alpha}_0)$ (a larger $|\boldsymbol{\alpha}|$ will decrease the *metric* space distance for all pairs; see Methods). Maximizing the likelihood (finding the state of minimum energy in the spring network) model over all feature node pairs minimizes the total prediction error between *associative* and *metric* generative models of the world[15] (Fig. S7).

The *associative* distances can be straightforwardly learned during *online* exploration. Since Hebbian learning reflects coactivity, a trajectory exploring the environment uniformly results in synaptic strengths between place cells proportional to the spatial correlation between their receptive fields[41]. The Euclidean separation between their fields is then accessed via a simple transformation (see Methods; Fig. S1G). In this context, learning the PC-GC weights (modifying the *observation* model) during *online* localization corresponds to forming spatial priors over feature locations which anchor the structure, which would otherwise be translation or rotation invariant (since measurements are relative), learned during *offline* inference to constant locations on the grid-map. Taken together, our framework proposes a mapping onto anatomy of the joint *agent-feature* location distribution required for full probabilistic inference over environmental structure (Fig. S5; See overall algorithm in Table S1).

Partial grid pattern responses to environmental rescaling

Uniform rescaling of an environment will introduce a mismatch between the estimates of location from the *transition* and *observation* models (i.e. a 'prediction error'). To minimize these prediction errors, the *offline* system can either modify the *transition* model gain to match the current environmental input (Fig. 2A, bottom), or modify the mapping from environmental inputs to *metric* space in the *observation* model (Fig. 2A, middle; see Methods). The degree to which either is modified should reflect their relative confidences, specified by a 'transition confidence score' ($\mathrm{T_C} = \sigma^2_{Obs}/\sigma^2_{\alpha}$, the ratio of confidence in the *transition* vs *observation* models; see Methods). Similarly, a 'prior confidence score' $\mathrm{P_C}$ specifies how much the system will tolerate persistent prediction errors; if $\mathrm{P_C}$ is large, optimization may favour preserving prior configurations, as opposed to alignment of the current *transition* and *observation* models (see Methods).

We modelled experiments in which the physical environment[21,42] and perceived velocity through a visual virtual environment[7] were re-scaled, such that self-motion and sensory inputs conflicted. In both experiments, the rescaling of the grid patterns was *partial*, i.e. less than the magnitude of the physical or virtual manipulation, and less than those of place fields.

Both manipulations can be simulated by introducing a visual gain parameter α^{Visual} to the simulation of the environment (in both experiments it scales the amount of self-motion required to traverse the width of the perceived environment). Learned *associative* distances ($\boldsymbol{\delta}$) are also scaled by this parameter, reflecting its effect on the temporal overlap of place fields (see Methods Eq. 12). We simulated grid pattern rescaling responses over a range of transition confidence scores $\mathrm{T_C}$. When confidence in the *transition* model is high ($\mathrm{T_C} \rightarrow \infty$), grid patterns in the real world are unchanged when plotted against physical movement (but are changed when plotted in visual VR coordinates; Fig. 2B, first column). The opposite is true when confidence in the *observation* model is high: grid patterns are unchanged relative to the apparent environment (Fig. 2B, last column).

However, for intermediate $\mathrm{T_C}$ values (i.e. balanced confidence in the *transition* and *observation* models), the model predicts partial rescaling of the grid pattern relative to the size of the manipulation

(Fig. 2B, middle column), matching the observed grid patterns in both experiments[7,21] (and a similar third experiment in on a virtual linear track[43]; Fig. 2C).

Differential grid and place field responses to environmental reshaping

How does the *offline* system respond to more complex environmental deformations? When one wall of a familiar rectangular environment is rotated inwards by 45°[22], place fields near the wall shifted almost fully while fields further away remained largely stationary, consistent with place fields preferentially reflecting local environmental inputs[5]. In contrast, grid fields shifted only partially near to the manipulated wall. Using the observed place field shifts[23], we simulated the response of grid cells to the same manipulation (Fig. 3; see Supplementary Methods). Shifted place fields induce a misalignment between the *associative* distances and the distance between their encodings in *metric* space. The place field shifts are local and non-uniform, and so misalignment cannot be corrected by a global change to the *transition gain* $\boldsymbol{\alpha}$. Indeed, $\boldsymbol{\alpha}$ is not significantly modified during the optimization process, regardless of the T_C value. Instead, alignment between the *transition* and *observation* models is maximized by modifying the *observation model*, i.e. updating the locations of the place fields on the grid module.

If there were no confidence in the prior *observation* model ($P_C \rightarrow 0$), it would be modified offline to match the transition model, leaving the grid pattern unperturbed by the environmental change (Fig. 3B,C, top row). At the other extreme, favouring prior beliefs over recent observations (i.e. pairwise distances encoded during the manipulation trial) would result in an unchanged *observation* model, and grid field shifts that exactly mirror corresponding place field shifts (Fig. 3B,C, bottom row). In this regime, there would be permanent misalignment of the *transition* and *observation* models during *online* localization, producing noisy grid patterns, as when simulating a related experiment where grid distortions were observed in trapezoid environments[44] (Fig. S2D). Setting P_C to an intermediate value reproduces the experimentally observed partial shifting of grid fields (relative to the place fields[23]) when visualizing the structure encoded in the *observation* model (i.e. assuming low confidence in the transition model; Fig. 3B,C and D).

Direction dependent shifting of grid patterns during *online* localization

In addition to partial changes to grid scale in response to environmental rescaling, enduring misalignments between *observation* and *transition* models can result from strong model priors, which prevent complete adaptation of the *transition* model gain. These cause the *transition* estimate to consistently precede that of the *observation* model, in the current direction of travel (on a 1D track, Fig. 2E) during VR visual gain decrease trials or physical expansion of the environment.

In all three cases, the integrated estimate of location (Eq. 2) in the *online* model converges to a fixed distance ahead of the *observation* model estimate (in the direction of travel; Fig. 2E inset and 3D), causing the grid pattern in the real world to dynamically shift opposite the direction of travel, as observed experimentally[7,9]. Our model suggests that the offsets should be partial (smaller than implied by a hard-reset at the boundary) and not specifically require a recent boundary encounter (cf. Keinath et al.[9]). Dynamic shifting in the model will reduce with experience of the novel or manipulated environment, as model misalignment reduces, as observed experimentally[45].

Online and *offline* perceptual warping in spatial representations

With increasing experience of an environment, grid firing patterns exhibit both local scale changes[26] and global shear-like distortions[29], the latter associated with 7.5-8° offsets of one of the grid axes[29,44] to the walls of square environments. Both effects were present in our simulations and can be attributed distinctly to the *offline* map-learning and *online* localization components of our theoretical framework.

Firstly, we show that local changes to the grid scale[26], which are positively correlated with behavioural occupancy (animals spend more time in the middle of the environment), arise from the *offline* process of map (PC-GC connections) learning. These mapping-induced distortions can be further subdivided into two mechanisms, both of which induce local scale changes by biasing the pairwise distances recovered from the Hebbian learned recurrent connections in CA3 (Fig. 1D,4; Fig. S7).

Firstly, relative behavioural under-sampling of the place fields near the boundaries of the environment (using occupancy statistics from[26]; Fig. 4A, bottom) lead to weaker PC-PC connections, and consequent overestimation of their pairwise distances, producing local scale changes (Fig. S7C).

Secondly, since Hebbian learned connection weights between place cells reflect the correlation in their firing, and therefore their statistical discriminability[46], two place cells with broad receptive fields would develop a stronger connection than a pair with equal separation but narrower receptive fields (stronger connections correspond to shorter distances on the grid module, producing grid patterns with larger scales in the environment; Fig. 1D). Another recent study[47] suggests that place fields are narrower near the edges of an environment, consistent with greater precision when driven by more proximal environmental features[5] (Fig. 4A, top row). In our model, this produces weaker recurrent connections and a shrinking of the grid pattern at the edges of the environment following *offline* inference (Fig. 4D,E, top row).

Together, our results suggest that the cognitive 'distance' between two sensory features should be greater both when the absolute confidence in their spatial locations is greater (reflecting an increased statistical discriminability), or when those features are under-sampled relative to other features.

Although the action of both mechanisms are independent their effect is the same; both i) relative under-sampling of the transition between two adjacent states and ii) a reduced statistical discriminability between those states, both contribute to a weaker pairing of their representative place cells, resulting in greater separation between their encodings in *metric* space and a locally larger grid scale when 'read-out' in the firing pattern (a locally larger perception of distance).

In contrast, global shear-like distortions[29] and associated 7.5-8° offsets of one of the grid axes[29,44] can be interpreted as *localization* induced distortions during *online* exploration. In Stensola et al.[29], rats were introduced into the same corner of the box at the start of each trial; in Butler et al.[25], shearing developed following the introduction of reward[25]. In both experiments, shearing developed with increasing experience[25,29]. We hypothesized that these distortions reflect an increasing effect of non-

uniform environmental inputs to the grid module, either reflecting their natural distribution[25,29] or inhomogeneous behavioural sampling of environmental locations[26].

In our simulations, given a learned map, biasing the strength of sensory inputs at specific locations (e.g. one/two corners) during *online* exploration reproduced several experimentally characterized global distortions by causing a bias in the decoding of location (i.e. salient locations contribute a larger 'vote'; Fig. S3; see Supplementary Methods).

Probabilistic inference through HPC-mEC message passing

To this point we discussed, from a functional perspective, how the brain might optimize its internal representations to reflect the uncertainty of sensory information. But how might the brain perform this optimization? In the above analyses of *offline* inference, we numerically computed the maximally likely feature locations on the grid module. However, the system must also track the uncertainty in these estimates, which would require updating the place-grid cell weights (including those with firing fields far from the agent location). An update of the full weight distributions is generally intractable when the state space is large.

Belief propagation[48] is a technique for approximating this inference on graph structured data, and comprises two stages. First, a given feature node (i.e. a place cell) computes its location distribution (i.e. connections to the grid cells) $B_i(\boldsymbol{b}_i)$ by multiplying its prior $B_i^{(0)}(\boldsymbol{b}_i)$ with messages received from its connected neighbours (Fig. 5C; see Methods for details). A message $m_{i\to j}(\boldsymbol{b}_j)$ expresses neighbour node i's belief of node j's location, conditioned on its own distribution, and is dependent on the same pairwise potential terms $\psi(\cdot)$ in Eq. 3. The effect of a message is to favour distributions of nodes i and j which locate them at a radial distance equal to the associative distance δ_{ij}; causing messages to be expressed as rings centred on the belief of the broadcasting node (Fig. 5C). Resolving a feature's unique location then depends on aggregating messages from multiple neighbours (Fig. 5C). Computations are distributed, and importantly only require information that is local to each neuron.

Each node in the graph iterates between updating its belief and broadcasting messages, converging when new messages cease to change the beliefs of their recipient nodes. As expected, the reduction in

pairwise prediction error between associative distances and their corresponding distances in grid space (see Methods) over successive message iterations is accompanied by a sharpening of the distribution of each feature's location on the grid module (see Supplementary Methods; Fig. S1H).

Offline inference triggered by prediction errors

How might the *online* and *offline* systems interact? If the *online* system is sufficient to localize within pre-learned, simple or slowly changing environments, non-local reactivations of place cells would be unnecessary. However, more complex *offline* inference is required under more demanding circumstances, or in novel or changing environments. We hypothesize that *offline* or 'remote' inference is triggered by *prediction errors* between location estimates from the *transition* and *observation* models, respectively (Fig. 1G), defined in our model as the Kullback-Leibler divergence $\epsilon = \mathbb{KL}(G', H)$.

Prediction errors are large when the *observation model* prediction (weighted place cell input) is different and more sharply peaked than the *transition model* estimate (Fig. 1G; see Methods; prediction errors will not be generated in absence of incoming sensory information, as in darkness, when the observation model estimate is uncertain).

To illustrate our *dual-systems* (*online*+*offline*) hypothesis (Fig. S7), we simulated an agent navigating around a novel circular track (the *loop closure* task; Fig. 5). Completion of the first lap produces positive prediction errors between the sharply peaked input from feature inputs learned at the beginning of the trial, and the agent location estimate which is uncertain given the accumulation of PI noise (Fig. 5B).

Decrease in structural error (the difference between the place field separations and their encoded separations on the grid cell sheet) following *online*+*offline* inference was markedly larger than following *online* learning alone (Fig. 5E). The inferential power of this 'one-shot' learning process derives from consideration of the full covariance structure of the feature locations (captured by the CA3 connection weights between place cells), compared to the purely local learning occurring *online*.

297 The system was subsequently able to navigate with dramatically reduced error (Fig. 5Aiv),
298 eliminating prediction errors on subsequent lap completions (Fig. 5B; Supplementary Video 1).

299 Coordinated grid-place cell replay as structured information propagation

300 The *scheduling* of updates in belief propagation is important because messages that do not change the
301 beliefs of neighbours are redundant (Fig. 6A). We scheduled only the place cell whose belief had
302 changed most to broadcast a new message on each cycle (Fig. 6A; see Methods). This *max-update*
303 scheduling was more efficient than simple synchronous schemes, converging with fewer messages
304 (Fig. 6C; see Ref.[49]).

305 The sequences of place cells broadcasting messages during *offline* inference in the loop-closure
306 simulation have significant structure (6B). They tend to initially propagate backwards along the track
307 from the animal's current position, resembling the characteristic reverse hippocampal replay
308 following reward[17] (Fig. 6B), but also occasionally hop to new locations where remote sequences are
309 initiated[50] (Fig. 6B,F). These subsequent sequences showed an approximately equal distribution of
310 forward/reverse sweeps (Fig. 6D; see Methods; Supplementary Video 1).

311 Thus, hippocampal 'replay' may reflect correction of local regions of the cognitive graph given new
312 or 'surprising' information, as opposed to simple recapitulation of experience[51]. Sequences selectively
313 affect place cells whose beliefs are structurally affected, and terminate when this is no longer the case,
314 'hopping' to remote regions. This leads to smooth sequences in un-converged graphs (novel
315 environments) and more hoppy sequences with experience, where converged regions may be skipped
316 (Fig. 6F). These 'hops' marked the separation of 'replay' events into distinct sub-sequences (see
317 Methods). Multiple trajectories may also be played out in parallel (e.g. two trajectories alternating
318 under max-scheduling; Fig. 6B, middle, grey shading).

319 A neural model of coordinated place cell – grid cell replay

320 How might belief propagation for *offline* inference be implemented in spikes fired by place and grid
321 cells during replay? We propose a schematic model with a focus on function rather than biological
322 detail (e.g. our 'place cells' combine the recurrent connections of CA3 with the connections to mEC

323 of CA1). In the model, minimizing prediction errors between *associative* and *metric* generative
324 models corresponds to synchronizing the propagation of activity through CA3 and mEC, respectively
325 (Fig. 7; see Supplementary Methods; Supplementary Video 2). A 'message' is initiated by a place cell
326 spike, which propagates in CA3 via the Hebbian recurrent connections that encode place field
327 separations. In parallel, the same spike initiates activity at the corresponding location on the grid cell
328 module, which then propagates on the grid sheet as a traveling wave, using the same circuitry as path
329 integration in the *online* model and propagating at the same speed as spikes in CA3 (see Methods).
330 Hebbian-like learning strengthens connections from place cells to grid cell which simultaneously
331 receive input in CA3 and EC respectively (Fig. 7A), approximating the algorithmic message-passing
332 implementation (Fig. 7B, C). Firing of the broadcasting place cell is triggered by changes in its
333 synaptic weights to the grid cell population, reflecting correction of the observation model in response
334 to prediction error with the transition model (see Supplementary Methods).

335 Discussion

336 Building on previous work[11,13,15], we argue that the mEC-HPC system performs spatial inference in
337 two distinct regimes. Given a known 'cognitive map' (mapping sensory information to metric space),
338 probabilistic integration allows optimal estimation of current location by *online* combination of
339 uncertainty-weighted self-motion and environmental observations provided by *transition* and
340 *observation* models respectively (Fig. 1). Where these estimates deviate strongly, prediction errors
341 (Fig. 1G) trigger *offline* inference events (Fig. 5B), which propagate local environmental input to
342 remote but structurally associated states, producing coordinated (often sequential) reactivations in
343 place and grid cells (Fig. 6, 7). The effect of *offline* inference is to produce a 2D embedding of the
344 sensory information provided through the place cells, which may facilitate planning or generalization.
345 Although not modelled here, back-projections from grid to place cells, reflecting the metric
346 embedding of their place fields, might therefore also reduce uncertainty in place cells' firing,
347 producing increased spatial stability in their fields, as observed to occur during sleep[12].

348 Partial rescaling of grid patterns[7,21] and differential shifting of grid and place fields[23] in response to
349 manipulations of environment sensory input can be understood as joint optimization of *transition* and

350 *observation* models, balancing model priors with new observations. Where prediction errors persist,
351 direction dependent grid pattern shifts may emerge as a result of probabilistic integration of these
352 conflicting cues[7,9] (whereas boundary-dependent resetting[9] produces larger shifts than experimentally
353 observed and no rescaling; Fig. 2D, E).

354 We show that observed grid pattern distortions can be mechanistically linked to inhomogeneity in the
355 sampling or neural representation of the environment[25,26,29] (Figs. 4A, S3), which might be reflected in
356 behaviour[52]. Thus variation in the confidence, sampling or discriminability of sensory states will
357 produce local changes in grid scale, inducing non-Euclidean structure in the *metric* representation of
358 space (Fig. 1D, 4B). Our model also shows that distortions appear gradually with experience[29], as the
359 learned mapping from sensory features to metric space (the *observation* model) becomes more
360 confident relative to the estimate of location from path integration (the *transition* model). Given initial
361 learning, o*nline* localization errors (Fig. S3) should occur immediately following subsequent
362 manipulations to the environmental sensory input, whereas *offline* changes may occur over longer
363 timescales and correlate with replay of the manipulated states (Figs. 4,6; consistent with grid, but not
364 place fields reorganizing significantly during sleep[53]). However, although large prediction errors will
365 cause more easily detectable *offline* inference events, *offline* learning may occur continuously and not
366 necessarily reactivate distinct previously experienced spatial trajectories[51]. We note that strong
367 *associative* connectivity may also contribute to pattern completion, making the place cell
368 representation robust to cue removal[54].

369 Theoretical studies have demonstrated how the connectivity of the mEC *metric* space might emerge
370 from a low-dimensional embedding of sensory stimuli[55], predictive states[35] or from unsupervised
371 learning during navigational tasks[56]. A crucial difference in our model is that perceptually similar but
372 physically separated compartments will be represented distinctly[57], reflecting the vectorial translation
373 between them in the transition model (i.e., not simply reflecting the topological state transition
374 structure[35]). Another recent model showed that grid cell like responses can emerge from learning the
375 transition model that best predicts observed sensory stimuli[36]. We instead assume a fixed transition
376 structure but with a variable linear gain, consistent with continuous attractor models[31] where

translation of activity on the grid cell sheet is driven by cells with velocity-dependent firing rates[38,39]. Indeed, a recent study showed that velocity dependence in mEC firing is tied to environmental manipulations[39].

We propose that *offline* structural inference events correspond to coordinated HPC/mEC replay[16–19,58–60], which can be viewed as synchronizing predictions from *associative* (CA3) and *metric* (mEC) generative models (Fig. S7). In this way, structural changes to an environment can be propagated to non-local regions of the *metric* embedding, in contrast to models in which these states need to be physically revisited[13], consistent with the observation that replays do not necessarily repeat experienced trajectories[51]. Prediction errors between the two models may trigger replay events and corresponding sharp-wave ripples[61,62]. To our knowledge, this is the first functional model of coordinated place cell-grid cell replay[18] (although cf. Ref. [63]), and provides an alternative to reward-based theories[64,65] (we note that rewards may themselves represent salient sensory features, independent of their reward value).

Our model makes a number of experimentally testable predictions. Firstly, systematic manipulation of the discriminability of sensory cues distributed within an environment should produce predictable distortions to the grid pattern, observed with increasing experience of an environment. Secondly, replay should be more frequent after structural changes such as shortcuts, blockages or gain manipulations as in the experimental setup of Fig. 5. Thirdly, replay events triggered by specific unexpected sensory observations should become less frequent (Fig. 5B) and smooth (Fig. 6E, F) with continued experience, if the observations remain stable. Fourthly, multiple local replay events may occur in inter-leaved fashion (Fig. 6C, *middle*, grey shading). Fifthly, we predict the existence of travelling waves in grid cells (as a function of their spatial phase; see also[66,67], Fig. 7). Lastly, initial messages propagating from the animal's current location may not cause subsequent messages in remote regions of the graph which are already sufficiently converged (messages will not cause changes in the beliefs of their recipients), although activity in mEC will continue to propagate. Thus grid cell replay could thus be detectable in the absence of simultaneous place cell replay[19] (but place cell replay requires the grid cell *transition* model and so depends on mEC[68]).

Our proposed structure learning framework can account for diverse phenomena observed in the HPC-mEC system, and makes several novel, experimentally testable predictions.

Methods

Online recursive Bayesian estimation

The transition matrix T defines the probability of transitioning from agent location $\boldsymbol{x}'$ to location $\boldsymbol{x}$, and is a function of the perceived current velocity $\hat{\boldsymbol{u}}$ and *transition* model gain $\boldsymbol{\alpha} = [\alpha_x, 0; 0, \alpha_y]$. Since our *metric* space is periodic, T accounts for cyclic transitions $\boldsymbol{c}_{mn}$, with Gaussian noise proportional to the perceived velocity $\hat{\boldsymbol{u}} \sim \boldsymbol{u} + N(0, \mathrm{diag}(\boldsymbol{u})\boldsymbol{\Sigma}_{\mathrm{PI}})$:

$$T(\boldsymbol{x}, \boldsymbol{x}' | \hat{\boldsymbol{u}}, \boldsymbol{\alpha}) = \sum_{mn=-\infty}^{\infty} f(\boldsymbol{x} - \boldsymbol{x}' | \hat{\boldsymbol{u}} + \boldsymbol{c}_{mn}, \mathrm{diag}(\hat{\boldsymbol{u}})\boldsymbol{\alpha}^{-1}\boldsymbol{\Sigma}_{\mathrm{PI}}) \tag{7}$$

where $f(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a multivariate Gaussian PDF, $\boldsymbol{c}_{mn} = 2\boldsymbol{\alpha}^{-1}(m\boldsymbol{v}_1 + n\boldsymbol{v}_2)$ and $\boldsymbol{v}_1 = [\cos(\phi), \sin(\phi)]$ and $\boldsymbol{v}_2 = [\cos\left(\phi + \frac{\pi}{3}\right), \sin\left(\phi + \frac{\pi}{3}\right)]$ define the unit vectors of a hexagonal lattice[69] with grid pattern orientation ϕ and $\mathrm{diag}(\cdot)$ produces a diagonal matrix from a vector input. Since most of the mass is associated with shorter transitions, in practice we approximate the full distribution with a finite number of periodic summations (i.e. ignore the tails; 5 cycles in our simulations; Fig. S1B).

The *observation* model defines the likelihood of the current environmental sensory inputs (i.e. the population vector of place cell firing $\boldsymbol{P}$, where $p_i(\overline{\boldsymbol{x}}) = f(\overline{\boldsymbol{x}}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{PC}^{\mathrm{i}})$ is the firing rate distribution of place cell i over physical space $\overline{\boldsymbol{x}}$) given the predicted *metric* location $\boldsymbol{x}$**,** via a thresholded weighted sum: $H(\boldsymbol{P}|\boldsymbol{x}) = [\sum_i B_i(\boldsymbol{x})P_i]^+$. Here, $B_i(\boldsymbol{x})$ is the location distribution of landmark i in *metric* space, which would be encoded biophysically in the learned $[N_P \times N_G]$ matrix $\boldsymbol{B}$ of synaptic weights from place to grid cells (i.e. the gth row and i^{th} column of $\boldsymbol{B}$ is the distribution $B_i(\cdot)$ evaluated at the location of the gth grid cell). The normalization constant $K = \int G(\boldsymbol{x})d\boldsymbol{x}$ in Eq. 2 simply sums over the current grid cell activity and might biophysically be implemented by inhibitory interneurons.

Online learning of structural priors

In the *online* model, the place-grid cell weight matrix $\boldsymbol{B}$ is learned using the BCM rule:

$$\begin{aligned} \boldsymbol{B} \leftarrow \quad & \boldsymbol{B} + \tau_{PG}\boldsymbol{P}^{\top} \otimes [\boldsymbol{G}' \odot (\boldsymbol{G}' - \boldsymbol{\theta})] && (8a) \\ \boldsymbol{\theta} \leftarrow \quad & \boldsymbol{\theta} + \tau_{\theta}[(\boldsymbol{G} \odot \boldsymbol{G}) - \boldsymbol{\theta}] && (8b) \end{aligned}$$

where $\tau_{PG} = 1e - 4$ is the learning rate and $\boldsymbol{G}'$ and $\boldsymbol{H}$ are column vectors whose elements are the *transition* and *observation* models estimated at the locations of a finite set of grid cell locations. $\odot$ is the element-wise (Hadamard) product between two vectors and $\otimes$ is their outer product. The sliding threshold $\boldsymbol{\theta} \in R^{1\times N_G}$ provides adaptive synaptic normalization, where $\tau_{\theta} \approx 10\tau_{PG}$. Learning takes place between the *apriori* distribution $\boldsymbol{G}'$ and the current sensory observation $\boldsymbol{H}$ (i.e. before the observation correction to $\boldsymbol{G}'$).

The *offline* probabilistic graphical model

The pairwise potentials $\psi(\boldsymbol{b}_i - \boldsymbol{b}_j, \boldsymbol{\alpha}|\delta_{ij}) = \sum_{m,n=-\infty}^{\infty} \exp\left(-w_{ij}\left(\delta_{ij} - d_{mn}(\boldsymbol{b}_i - \boldsymbol{b}_j, \boldsymbol{\alpha})\right)^2\right)$ penalize differences in the pairwise distances encoded by *association* δ_{ij}, and those that would be computed by comparing their absolute encodings in *metric* space $d_{mn}(\cdot)$; i.e. they encourage a metric embedding that reflects the associative distance. The *metric* distance function $d_{mn}^2(\boldsymbol{x}, \boldsymbol{\alpha}) = (\boldsymbol{x} + \boldsymbol{c}_{mn})^{\mathrm{T}}\, \boldsymbol{\alpha}(\boldsymbol{x} + \boldsymbol{c}_{mn})$ defines the pairwise distance between the encoding of locations i and j in mEC in *metric* space, and is dependent on the gain factor $\boldsymbol{\alpha}$ of the *transition model*. Pairwise measurements are assumed to have confidence w_{ij} (inverse variance) that increases with decreasing inferred distance (i.e. $w_{ij} = 1/(\sigma_{PC} + \sigma_{PI}\delta_{ij})$). The *transition model* gain is assumed to have a Gaussian prior $p(\boldsymbol{\alpha}) = \exp\left(-w_{\alpha}(\boldsymbol{\alpha} - \boldsymbol{\alpha_0})^{\mathrm{T}}(\boldsymbol{\alpha} - \boldsymbol{\alpha_0})\right)$, the w_{α} term representing the confidence in the prior gain value $\boldsymbol{\alpha}_0$. The periodic offset term $\boldsymbol{c}_{mn}$ is the same as defined for the transition model.

Associative encoding in the hippocampus

The associative distances are recovered from the $[N_P \times N_p]$ synaptic weights in CA3 $\boldsymbol{A}$**.** Under a random-walk behavioural trajectory, the simple modified Hebbian learning rule:

$$\boldsymbol{A} \leftarrow \boldsymbol{A} + \tau_{PP}[\boldsymbol{P}\boldsymbol{P}^{\top} - \boldsymbol{A} \odot \boldsymbol{A}] \qquad (9)$$

where $\odot$ is the element-wise (Hadamard) product and τ_{PP} the learning rate. The synaptic weights can be shown to converge to $A_{ij} = \sqrt{\langle P_i(t)P_j(t)\rangle_t}$, the square-root of the correlation between the firing rates of two PCs[70]. Where place fields have uniform receptive field widths (σ_{PC}) and peak firing rates, the Euclidean distance between place fields *i* and *j* can be inferred via the simple transformation[46]:

$$\delta_{ij}^2 = -\log(A_{ij}) = \frac{(\mu_i - \mu_j)^2}{2\sigma_{PC}^2} \tag{10}$$

The recovered distance is therefore scaled by the receptive fields' variance (the Bhattacharyya distance[46]), and so relates to 'discriminability' (Fig. 1D). CA3 synapses effectively average over multiple pairwise measurements. By assuming that noise in the pairwise distance measurements scale linearly with distance, both the mean and variance of the Gaussian describing this distribution is efficiently encoded in a single PC-PC synapse.

Simplified analysis of the probabilistic graphical model likelihood

To characterize model predictions in the environmental rescaling[71] and gain change[7] experiments, we studied a reduced version of the full graphical model (Eq. 3; see Supplementary Methods for full derivation). In 1D, given a linear *observation model* $x = H(x') = K_1 + Kx'$ and a large number of evenly spaced place fields, Eq. 3 simplifies to (see Supplementary Methods):

$$-\log L = \ell \approx \left(1/\alpha^{Visual} - K/\alpha\right)^2 + \mathrm{P_C}(K - K^0)^2 + \mathrm{P_C} \cdot \mathrm{T_C}(1/\alpha - 1/\alpha^0)^2 \tag{11}$$

which can be solved analytically. A similar reduction was applied to the 2D case when considering differential shifts in grid and place fields[23].

Belief propagation for *offline* inference

Belief propagation[48] is an iterative, two-stage local message-passing scheme in which, at each iteration n, a feature node (i.e. a place cell) first updates its own belief (connections to the grid cells) $B_i^{(n)}$ by integrating messages from connected nodes $j \in \Gamma_i$ with its own prior belief $B_i^{(0)}$ (Fig. 5C):

$$B_i^{(n)}(\boldsymbol{b}_i) \propto B_i^{(0)}(\boldsymbol{b}_i) \prod_{j \in \Gamma_i} m_{j \to i}^{(n)}(\boldsymbol{b}_i) \tag{12a}$$

The message from node j to node i ($m_{j\to i}$) communicates its belief over the distribution of the locations of place cell i in grid cell space, conditioned on its own location distribution:

$$m_{j\to i}^{(n+1)}(\boldsymbol{b}_i) \propto \int \psi(\boldsymbol{b}_j - \boldsymbol{b}_i, \boldsymbol{\alpha}|\delta_{ji}) \cdot \left[\frac{B_j^{(n)}(\boldsymbol{b}_j)}{m_{i\to j}^{(n)}(\boldsymbol{b}_j)}\right] d\boldsymbol{b}_j \tag{12b}$$

where the pairwise potentials $\psi(\cdot)$ are the same as those described in the full likelihood function (Eq. 3). The graph converges when new messages cease to change the beliefs of their recipient nodes.

Scheduled message passing on the place cell graph

'Synchronous' belief propagation computes belief updates for each step before broadcasting all new messages in the next step. In simulations, we demonstrated that scheduling message broadcasts based on internal 'message tension' (divergence between previous and updated belief given new messages) produced faster and more accurate convergence (Fig. 6B; see also Elidan et al.[49]). Message tension between the node's previous $\boldsymbol{B}_i^{(n-1)}$ and updated $\boldsymbol{B}_i^{(n)}$ beliefs is defined as:

$$\mathcal{T}_i^n = \mathbb{JS}(B_i^{(n)} \parallel B_i^{(n-1)}) = \frac{1}{2}\left[\mathbb{KL}(B_i^{(n)} \parallel m) + \mathbb{KL}(B_i^{(n-1)} \parallel m)\right] \tag{13}$$

where $\mathbb{JS}$ is the Jensen-Shannon (symmetric K-L) divergence, where:

$$\begin{aligned} m &= \frac{1}{2}\left(B_i^{(n)} + B_i^{(n-1)}\right) \\ \mathbb{KL}(p \parallel q) &= -\int p(x) \cdot \log\left(\frac{q(x)}{p(x)}\right) dx \end{aligned} \tag{14}$$

When the message tension is below a pre-defined threshold $\mathcal{T}_{min}$, a node has converged and ceases to broadcast new messages. This mechanism is similar to the prediction error between *transition* and *observation* models used to trigger *offline* inference, with the exception that it uses the symmetric divergence measure $\mathbb{JS}$ rather than $\mathbb{KL}$.

Traveling waves in neural media

In simulations, the traveling waves in mEC are simulated explicitly by calculating the true messages conditioned on the sending nodes' current beliefs at each time-step (Eq. 5B). In the 'neural model'

(Fig. 7), messages were approximated as waves propagating radially from an initial stimulation on the mEC sheet using a modified mechanical wave model, as used to describe oscillations in water:

$$\frac{d\boldsymbol{v}^2}{dt} = c^2 \nabla'^2 \cdot [\boldsymbol{v}]^+ \tag{15}$$

where c is the speed of wave propagation, $[\cdot]^+$ is a threshold linear activation function, and the modified spatial Laplacian operator ∇' is a symmetric 2D Gaussian filter with variance equal to the PI noise (see Supplementary Methods for extended discussion).

Acknowledgements

We thank the groups of Caswell Barry and Francesca Cacucci for use of their data and Martin Pearson (UWE and Bristol Robotics Lab), Dan Bush, Andrej Bicanski and Tim Behrens for valuable theoretical discussions.

We acknowledge funding by the ERC Advanced grant NEUROMEM, the Wellcome Trust and the European Union's Horizon 2020 research and innovation programme under grant agreement No. 785907 Human Brain Project SGA2. The authors declare no competing financial interests.

References

1. Burak, Y. & Fiete, I. R. Accurate path integration in continuous attractor network models of grid cells. *PLoS computational biology* **5**, e1000291 (2009).
2. Fuhs, M. C. & Touretzky. A Spin Glass Model of Path Integration in Rat Medial Entorhinal Cortex. *Journal of Neuroscience* **26**, 4266–4276 (2006).
3. McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I. & Moser, M.-B. Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience* **7**, 663–678 (2006).
4. Lever, C., Burton, S., Jeewajee, A., O'Keefe, J. & Burgess, N. Boundary Vector Cells in the Subiculum of the Hippocampal Formation. *Journal of Neuroscience* **29**, 9771–9777 (2009).
5. Hartley, T., Burgess, N., Lever, C., Cacucci, F. & O'keefe, J. Modeling place fields in terms of the cortical inputs to the hippocampus. *Hippocampus* **10**, 369–379 (2000).

6. O'Keefe, J. & Burgess, N. Dual phase and rate coding in hippocampal place cells: Theoretical significance and relationship to entorhinal grid cells. *Hippocampus* **15**, 853–866 (2005).
7. Chen, G., Lu, Y., King, J. A., Cacucci, F. & Burgess, N. Differential influences of environment and self-motion on place and grid cell firing. *Nature communications* **10**, 630 (2019).
8. Hardcastle, K., Ganguli, S. & Giocomo, L. M. Environmental Boundaries as an Error Correction Mechanism for Grid Cells. *Neuron* (2015) doi:10.1016/j.neuron.2015.03.039.
9. Keinath, A. T., Epstein, R. A. & Balasubramanian, V. Environmental deformations dynamically shift the grid cell spatial metric. *eLife* **7**, e38169 (2018).
10. Durrant-Whyte, H. & Bailey, T. Simultaneous localization and mapping: part I. *IEEE Robotics & Automation Magazine* **13**, 99–110 (2006).
11. Milford, M. J., Wyeth, G. F. & Prasser, D. RatSLAM: a hippocampal model for simultaneous localization and mapping. in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on* vol. 1 403–408 (IEEE, 2004).
12. Guo, W., Zhang, J. J., Newman, J. & Wilson, M. A. Latent learning drives sleep-dependent plasticity in distinct CA1 subpopulations. *bioRxiv* (2020).
13. Ocko, S. A., Hardcastle, K., Giocomo, L. M. & Ganguli, S. Emergent elasticity in the neural code for space. *Proceedings of the National Academy of Sciences* **115**, E11798–E11806 (2018).
14. Penny, W. D., Zeidman, P. & Burgess, N. Forward and backward inference in spatial cognition. *PLoS computational biology* **9**, e1003383 (2013).
15. Evans, T. & Burgess, N. Coordinated hippocampal-entorhinal replay as structural inference. in *Advances in Neural Information Processing Systems* (2019).
16. Diba, K. & Buzsáki, G. Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience* **10**, 1241–1242 (2007).
17. Foster, D. J. & Wilson, M. A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* **440**, 680–683 (2006).
18. Ólafsdóttir, H. F., Carpenter, F. & Barry, C. Coordinated grid and place cell replay during rest. *Nature Neuroscience* **19**, 792–794 (2016).

19. O'Neill, J., Boccara, C., Stella, F., Schoenenberger, P. & Csicsvari, J. Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. *Science* **355**, 184–188 (2017).
20. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science* **261**, 1055–1058 (1993).
21. Barry, C., Hayman, R., Burgess, N. & Jeffery, K. J. Experience-dependent rescaling of entorhinal grids. *Nature Neuroscience* **10**, 682–684 (2007).
22. Barry, C., Ginzberg, L. L., O'Keefe, J. & Burgess, N. Grid cell firing patterns signal environmental novelty by expansion. *Proceedings of the National Academy of Sciences* **109**, 17687–17692 (2012).
23. Krupic, J., Bauza, M., Burton, S. & O'Keefe, J. Local transformations of the hippocampal cognitive map. *Science* **359**, 1143–1146 (2018).
24. Stensola, H. *et al.* The entorhinal grid map is discretized. *Nature* **492**, 72–78 (2012).
25. Butler, W. N., Hardcastle, K. & Giocomo, L. M. Remembered reward locations restructure entorhinal spatial maps. *Science* **363**, 1447–1452 (2019).
26. Hägglund, M., Mørreaunet, M., Moser, M.-B. & Moser, E. I. Grid-Cell Distortion along Geometric Borders. *Current Biology* (2019) doi:10.1016/j.cub.2019.01.074.
27. Hollup, S. A., Molden, S., Donnett, J. G., Moser, M.-B. & Moser, E. I. Accumulation of hippocampal place fields at the goal location in an annular watermaze task. *Journal of Neuroscience* **21**, 1635–1644 (2001).
28. Munn, R. G. K., Mallory, C. S., Hardcastle, K., Chetkovich, D. M. & Giocomo, L. M. Entorhinal velocity signals reflect environmental geometry. *bioRxiv* (2019) doi:10.1101/671222.
29. Stensola, T., Stensola, H., Moser, M.-B. & Moser, E. I. Shearing-induced asymmetry in entorhinal grid cells. *Nature* **518**, 207–212 (2015).
30. Hafting, T., Fyhn, M., Molden, S., Moser, M.-B. & Moser, E. I. Microstructure of a spatial map in the entorhinal cortex. *Nature* **436**, 801–806 (2005).
31. Fiete, I. R., Burak, Y. & Brookings, T. What Grid Cells Convey about Rat Location. *Journal of Neuroscience* **28**, 6858–6871 (2008).

32. Mathis, A., Herz, A. V. & Stemmler, M. Optimal population codes for space: grid cells outperform place cells. *Neural computation* **24**, 2280–2317 (2012).

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

34. Guanella, A., Kiper, D. & Verschure, P. A model of grid cells based on a twisted torus topology. *International journal of neural systems* **17**, 231–240 (2007).

35. Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. The hippocampus as a predictive map. *Nature Neuroscience* (2017) doi:10.1038/nn.4650.

36. Whittington, J., Muller, T., Mark, S., Barry, C. & Behrens, T. Generalisation of structural knowledge in the hippocampal-entorhinal system. in *Advances in Neural Information Processing Systems 31* (eds. Bengio, S. et al.) 8484–8495 (Curran Associates, Inc., 2018).

37. Widloski, J. & Fiete, I. R. A Model of Grid Cell Development through Spatial Exploration and Spike Time-Dependent Plasticity. *Neuron* **83**, 481–495 (2014).

38. Kropff, E., Carmichael, J. E., Moser, M.-B. & Moser, E. I. Speed cells in the medial entorhinal cortex. *Nature* **523**, 419–424 (2015).

39. Munn, R. G. K., Mallory, C. S., Hardcastle, K., Chetkovich, D. M. & Giocomo, L. M. Entorhinal velocity signals reflect environmental geometry. *Nat Neurosci* (2020) doi:10.1038/s41593-019-0562-5.

40. Milford, M. J. & Wyeth, G. F. Mapping a suburb with a single camera using a biologically inspired SLAM system. *IEEE Transactions on Robotics* **24**, 1038–1053 (2008).

41. Muller, R. U., Stead, M. & Pach, J. The hippocampus as a cognitive graph. *The Journal of general physiology* **107**, 663–694 (1996).

42. Burgess, N. & O'Keefe, J. Neuronal computations underlying the firing of place cells and their role in navigation. *Hippocampus* **6**, 749–762 (1996).

43. Campbell, M. G. *et al.* Principles governing the integration of landmark and self-motion cues in entorhinal cortical codes for navigation. *Nature neuroscience* **21**, 1096 (2018).

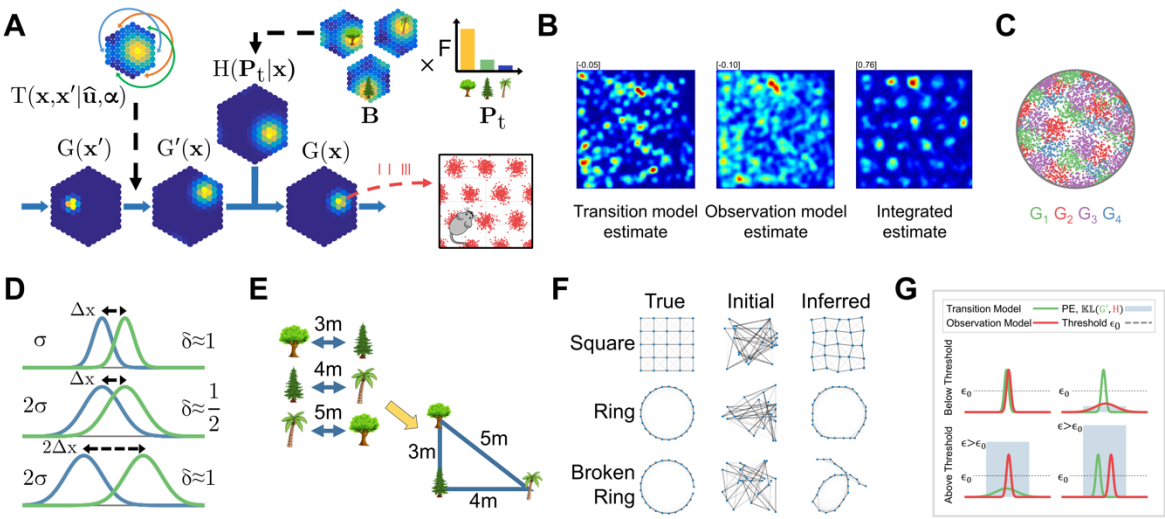
44. Krupic, J., Bauza, M., Burton, S., Barry, C. & O'Keefe, J. Grid cell symmetry is shaped by environmental geometry. *Nature* **518**, 232–235 (2015).

45. Hägglund, M. & Mørreaunet, M. Grid pattern development, distortions and topological defects may depend on distributed anchoring. *BioRxiv* 830158 (2019).
46. Bhattacharyya, A. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.* **35**, 99–109 (1943).
47. Tanni, S. & Barry, C. Relationship between place cell recruitment and the spatial scale of open-field environments revealed using a novel scalable recording apparatus. in *Society for Neuroscience* Program No. 689.18. (Society for Neuroscience, 2018).
48. Pearl, J. Reverend Bayes on inference engines: A distributed hierarchical approach. in *National Conference on Artificial Intelligence* 133—-136 (AIII Press, 1982).
49. Elidan, G., McGraw, I. & Koller, D. Residual belief propagation: Informed scheduling for asynchronous message passing. *arXiv preprint arXiv:1206.6837* (2012).
50. Davidson, T. J., Kloosterman, F. & Wilson, M. A. Hippocampal Replay of Extended Experience. *Neuron* **63**, 497–507 (2009).
51. Stella, F., Baracskay, P., O'Neill, J. & Csicsvari, J. Hippocampal Reactivation of Random Trajectories Resembling Brownian Diffusion. *Neuron* (2019) doi:10.1016/j.neuron.2019.01.052.
52. Bellmund, J. L. *et al.* Deforming the metric of cognitive maps distorts memory. *Nature human behaviour* 1–12 (2019).
53. Boccara, C. N., Nardin, M., Stella, F., O'Neill, J. & Csicsvari, J. The entorhinal cognitive map is attracted to goals. *Science* **363**, 1443–1447 (2019).
54. Nakazawa, K. *et al.* Requirement for hippocampal CA3 NMDA receptors in associative memory recall. *Science* **297**, 211–218 (2002).
55. Dordek, Y., Soudry, D., Meir, R. & Derdikman, D. Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. *eLife* **5**, (2016).
56. Banino, A. *et al.* Vector-based navigation using grid-like representations in artificial agents. *Nature* **557**, 429 (2018).
57. Carpenter, F., Manson, D., Jeffery, K., Burgess, N. & Barry, C. Grid Cells Form a Global Representation of Connected Environments. *Current Biology* 1–7 (2015) doi:10.1016/j.cub.2015.02.037.

58. Lee, A. K. & Wilson, M. A. Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron* **36**, 1183–1194 (2002).
59. Skaggs, W. E. & McNaughton, B. L. Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science* **271**, 1870–1873 (1996).
60. Wilson, M. A. & McNaughton, B. L. Reactivation of hippocampal ensemble memories during sleep. *Science* **265**, 676–679 (1994).
61. Buzsáki, G., Vanderwolf, C. H. & others. Cellular bases of hippocampal EEG in the behaving rat. *Brain Research Reviews* **6**, 139–171 (1983).
62. O'Keefe, J. Place units in the hippocampus of the freely moving rat. *Experimental neurology* **51**, 78–109 (1976).
63. Bicanski, A. & Burgess, N. A neural-level model of spatial memory and imagery. *eLife* **7**, e33752 (2018).
64. Mattar, M. G. & Daw, N. D. Prioritized memory access explains planning and hippocampal replay. (2018) doi:10.1101/225664.
65. Momennejad, I., Otto, A. R., Daw, N. D. & Norman, K. A. Offline replay supports planning in human reinforcement learning. *Elife* **7**, e32548 (2018).
66. Hasselmo, M. E. Neuronal rebound spiking, resonance frequency and theta cycle skipping may contribute to grid cell firing in medial entorhinal cortex. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **369**, 20120523 (2014).
67. Muller, L., Chavane, F., Reynolds, J. & Sejnowski, T. J. Cortical travelling waves: mechanisms and computational principles. *Nature Reviews Neuroscience* (2018).
68. Yamamoto, J. & Tonegawa, S. Direct medial entorhinal cortex input to hippocampal CA1 is crucial for extended quiet awake replay. *Neuron* **96**, 217–227 (2017).
69. Mathis, A., Stemmler, M. B. & Herz, A. V. Probable nature of higher-dimensional symmetries underlying mammalian grid-cell activity patterns. *Elife* **4**, e05979 (2015).
70. Dayan, P. & Abbott, L. F. *Theoretical neuroscience*. (Cambridge, MA: MIT Press, 2001).
71. Barry, C. & Burgess, N. Learning in a geometric model of place cell firing. *Hippocampus* **17**, 786–800 (2007).

657

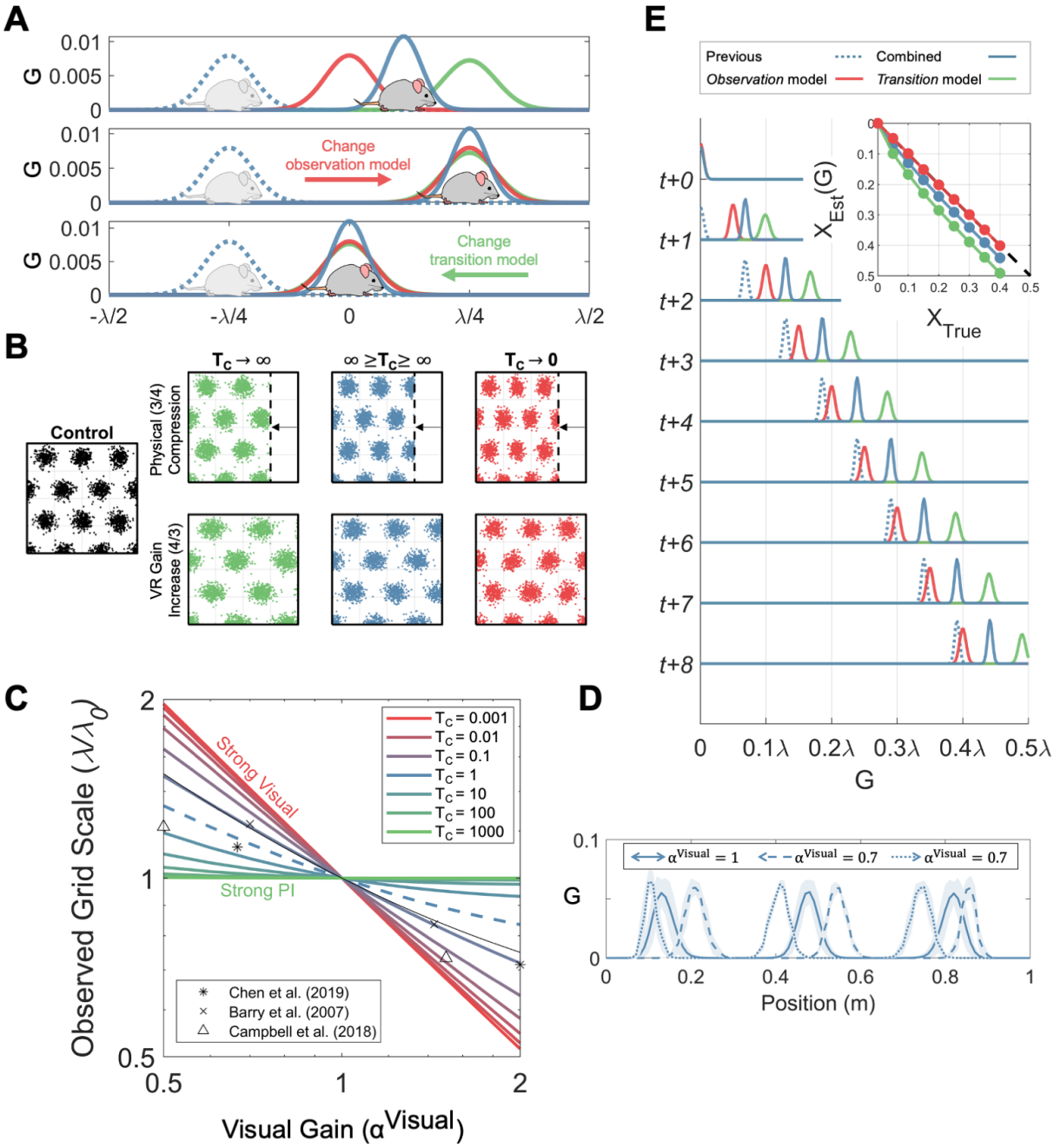
Figure 1. Online+offline localization and mapping. **A)** Illustration of recursive Bayesian integration. A probability distribution over current location ($G(\boldsymbol{x})$), represented by grid cell firing, is updated according to self-motion via the *transition model* ($\mathrm{T}(\mathbf{x}, \mathbf{x}'|\hat{\mathbf{u}}, \boldsymbol{\alpha})$) then refined by environmental inputs via the *observation model* ($H(\boldsymbol{P}_t|\boldsymbol{x}_t)$). **C)** Estimates based on the integration of noisy self-motion and environmental inputs may be stable, as shown in simulated grid cell firing rate maps (*right*), despite instability when using only self-motion (*left*) or environmental inputs after brief initial exploration (*middle*; numbers show gridness score). **C)** Simulated grid cells exhibit spatially offset grid-like firing patterns, due to toroidal connectivity, despite the absence of attractor dynamics. Right shows histogram of spatial phases. **D)** Inferred pairwise distances D are a function of the 'overlap' between place fields. **E)** Pairwise distances can be used to infer the structure of the world (the mapping of place fields onto the grid map). **F)** Given noisy initial priors ("Initial"), structural encodings are modified to reflect pairwise *associative* measurements. Inferred structure is sensitive to the topology of the environment (cf. "Ring" and "Broken Ring"). **G)** Illustration of the prediction error mechanism used to arbitrate between the *online* and *offline* systems (blue bars show prediction error ε, ε_0 is the minimum prediction error needed to trigger *offline* inference).



676

Figure 2. Minimizing prediction errors in the *offline* system: grid rescaling and direction-dependent offsets under manipulations of environmental size or VR gain. **A)** During *online* spatial localization, the *observation* model estimates location in metric space (activity on the grid sheet) via inputs from place cells driven by environmental features (red curve), the *transition* model updates the previous estimate (dashed blue curve) according to self-motion (green curve), producing a combined estimate (blue curve). Manipulations of the environment cause the predictions from both models to diverge. One way to minimize these prediction errors is to modify the *observation* model by changing the connection weights from place to grid cells (i.e. the mapping between environmental observations to estimated location in grid space). An alternative is to modify the *transition* model to reflect the *observation* model estimate (e.g. varying the 'gain' mapping self-motion to grid space). The degree to which both are modified is controlled by the relative strength of their respective prior confidences ($\mathrm{T_C}$). **B)** Joint optimization of the *observation* and *transition models* predict partial rescaling of grid patterns in response to increase in the VR gain (i.e. the rate of visual movement in response to physical movement on the ball; below) or compression of a real environment (above). When the system is confident in its self-motion ($\mathrm{T_C} \rightarrow \infty$, green), the *observation* model is modified to match the *transition* model (no change in grid scale plotted in the real or visual VR environment). When the system is confident in its environmental inputs ($\mathrm{T_C} \rightarrow 0$, red), the *transition* model adapts and grid scale follows the environmental change. Balanced model confidence produces intermediate rescaling (blue). **C)** Change in observed grid scale (grid pattern plotted against self-motion) depends on the *transition* confidence values $\mathrm{T_C}$. X-axis shows VR gain change or environmental compression, where $\alpha^{\mathrm{Visual}} = 1$ is a control trial. Y-axis shows observed change in grid scale (see Methods). Data points show corresponding values from Barry et al. (2007), Chen et al. (2019) and Campbell et al., (2019), which suggest an approximately equal weighting of *transition* and *observation* model priors for both gain decreases and increase trials ($\mathrm{T_C} \approx 1$). **D)** Firing rate map of a grid cell on the linear track in control (solid line) and VR gain decrease trials ($\alpha^{Visual} = 0.7$),

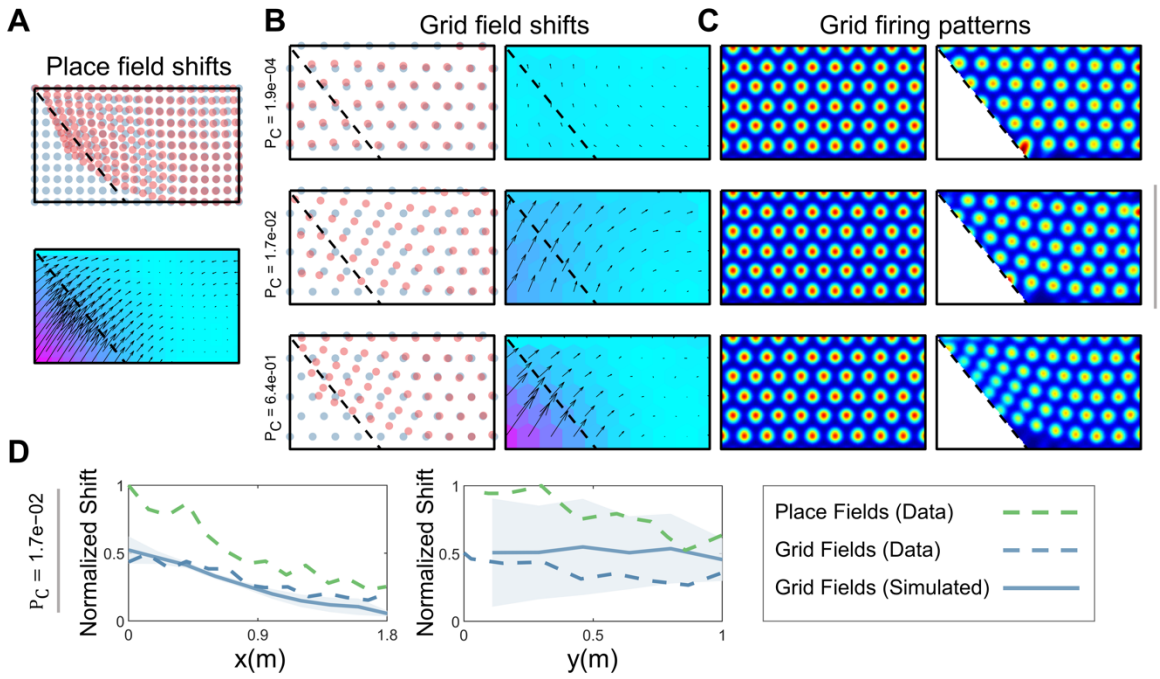
plotted in the visual VR environment. In the gain decrease condition, the grid fields are stably shifted to fire earlier when running left (dashed line) or right (dotted line), i.e. towards the location indicated by the transition model from that indicated by vision. **E)** The direction-dependent shift in estimated location in the grid module in a VR gain decrease trial ($\alpha^{Visual} = 2/3$) stabilises at a fixed distance. Sequence of eight updates of estimated location on the grid module G(x) when running to the right in a VR visual gain decrease trial, colours as in A. The *transition* model (green) predicts a location ahead of that from the *observation* model on the grid sheet (red; driven by visual input) because of the visual gain decrease. Combining these estimates produces an intermediate distribution (blue). At each new update, the prediction from the *transition* model builds on the shift of the previous combined estimate (not the previous transition model estimate) so that the distance between the observation model estimate and the combined estimate stabilizes at a fixed value, producing a fixed direction-dependent offset of the grid patterns in environmental coordinates (see the differences in the location estimates between models, Inset).


A
G
0.01
0.005
0
Change observation model
Change transition model
-λ/2
-λ/4
0
λ/4
λ/2
B
Control
Tc → ∞
∞ ≥Tc≥ ∞
Tc → 0
Physical (3/4) Compression
VR Gain Increase (4/3)
C
Observed Grid Scale (λV/λ0)
Visual Gain (αVisual)
Strong Visual
Strong PI
Tc = 0.001
Tc = 0.01
Tc = 0.1
Tc = 1
Tc = 10
Tc = 100
Tc = 1000
Chen et al. (2019)
Barry et al. (2007)
Campbell et al. (2018)
D
G
Position (m)
αVisual = 1
αVisual = 0.7
αVisual = 0.7
E
Previous
Combined
Observation model
Transition model
t+0
t+1
t+2
t+3
t+4
t+5
t+6
t+7
t+8
XEst(G)
XTrue
0.1λ
0.2λ
0.3λ
0.4λ
0.5λ

718

719

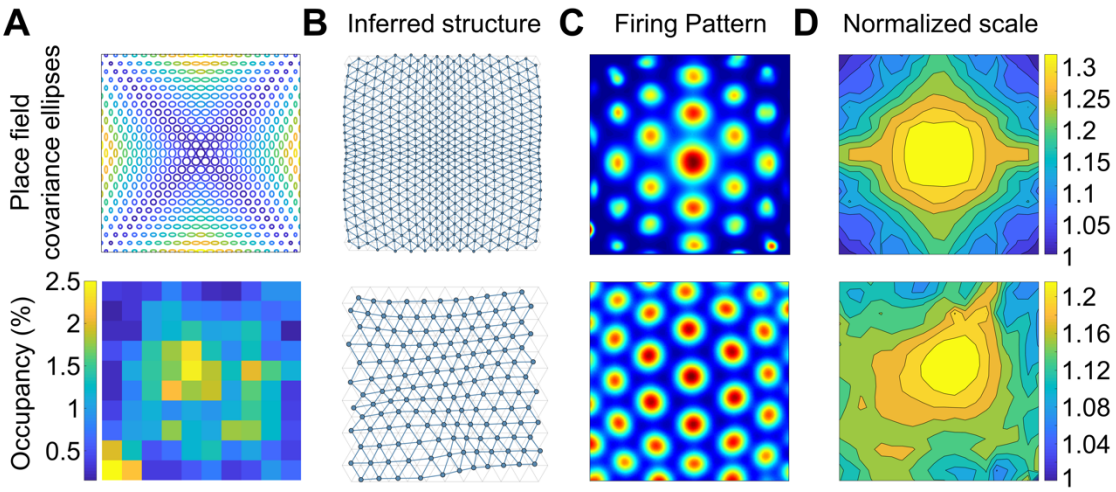
Figure 3. Effects of local deformation of a rectangular environment on grid patterns following *offline* inference (Krupic et al., 2018). Dots indicate place/grid field locations, arrows indicate shifts before/after *offline* structural optimization and coloured hexagons indicate magnitude of shift vector. Grid patterns show estimate generated by the *observation* model (weighted place cell activity). **A)** Place field shifts were measured from Krupic et al. (2018) and interpolated and smoothed. **B-C)** Place field shifts cause immediate prediction errors (PE) between the new pairwise place field distances and the distances between their corresponding grid locations. If the *observation / transition* model priors are weak ($\mathrm{P_C} \rightarrow 0$), PEs are eliminated during *offline* inference by updating the *observation* model (top row; unlike in Fig. 4B, modifying the global gain did not remove the effect of local distortions). Complete adaptation in the *observation* model leads to an unperturbed grid pattern, i.e. grid fields will not shift. Alternatively, strong model priors ($\mathrm{P_C} \rightarrow \infty$) prevent adaptation to new environmental inputs, leading to distorted grid patterns when driven purely by the observation model, whose field shifts match those of the place cells (bottom row). When the model priors are balanced against the new pairwise observations, the *observation* model is partially adjusted, producing partial grid field shifts (which are smaller than those of the place fields; middle row). Partial (middle) or no (bottom) adjustment to new observations preserve mismatches between the *transition* and *observation* models, which would result in direction-dependent offsets (see Figure 4E) and irregular firing patterns (see Figure S2D). NB the confidence in the prior model may depend on location, e.g. if there is strong anchoring to the wall prior connections from place cells with fields near the wall may be stronger than those with fields further away. **D)** Comparison of experimentally observed place and grid field shifts with simulation corresponding to the middle row of **B-C**.


A
Place field shifts
B
Grid field shifts
C
Grid firing patterns
P_C = 1.9e−04
P_C = 1.7e−02
P_C = 6.4e−01
D
P_C = 1.7e−02
Normalized Shift
x(m)
y(m)
Place Fields (Data)
Grid Fields (Data)
Grid Fields (Simulated)

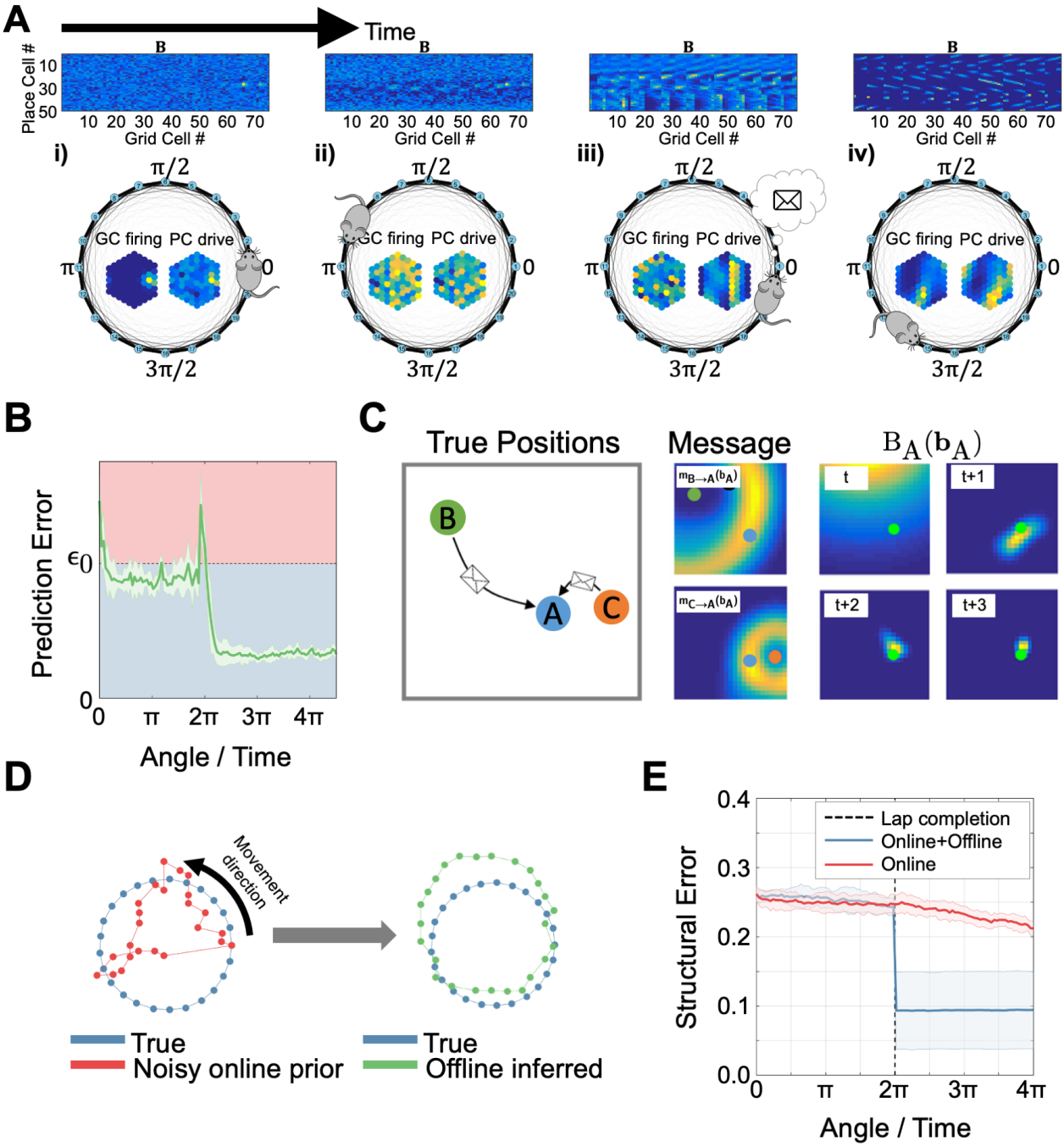
743

744

Figure 4. Distortions to grid patterns caused by inhomogeneous environmental input during offline inference. **A)** Simulated variation in place field shape due to proximity to boundaries (above), or inhomogeneous sampling of locations (Hagglund et al,. 2019; below) results in distortions to the inferred pairwise distances in CA3 (**B**). These distortions lead to distortions to grid scale due to adjustment of the observation model during offline inference (**C-D**).



752 **Figure 5.** Illustration of the dual-systems model: prediction errors and replay in the loop-
753 closure task (see also Supplementary Video 1). **A)** Place cell – grid cell (PC-GC) connection
754 weights (above) as agent runs around a circular track for the first time (below, GC activity
755 and the input from PCs to GCs, both shown on topographically organised sheet of cells,
756 inset). **i)** Confidence in the initial location is high, such that coactive GC and PC fields form
757 strong associations (the GC firing distribution is peaked and the inputs from place cells with
758 fields at the beginning of the track are strong). **ii)** The agent navigates around the track,
759 accumulating self-motion error, leading to diffuse GC firing. **iii)** Prediction errors (PE) on lap
760 completion (when the initially learned precise PC input arrives) triggers an *offline* inference
761 event (see main text and Supp. Video 1 for details). **iv)** On subsequent laps of the track, PC-
762 GC weights are sharply tuned following *offline* inference, allowing effective localization. **B)**
763 PE is reduced on completion of subsequent laps due to alignment of the *transition* and
764 *observation models* (i.e. environmental inputs and self-motion updating of GC activity
765 coincide). **C)** Illustration of belief propagation. Place cell A receives messages from PCs B
766 and C. Messages take the form of rings, describing a preferred distance about the current
767 locations of B and C with variance reflecting the confidence in the message (the variance of
768 pairwise distance estimates with Gaussian noise). The intersection of the messages
769 uniquely determines the location of A over time. NB A will also be broadcasting messages
770 back to B and C. **D)** True structure (blue), structure encoded by noisy path integration (left,
771 red; i.e. the location of the peaks of the weights from each place cell to the grid cell sheet)
772 and structure inferred after loop-closure (right, green). **E)** *Offline* inference allows one-shot
773 learning when compared to the online system alone.


A
Time
Place Cell #
Grid Cell #
i)
ii)
iii)
iv)
GC firing PC drive
B
Prediction Error
Angle / Time
C
True Positions
Message
D
Movement direction
True
Noisy online prior
True
Offline inferred
E
Structural Error
Lap completion
Online+Offline
Online
Angle / Time

774

775

Figure 6. Principled message scheduling during *offline* inference generates sequences of place cell activity. **A)** Place cells (i.e. graph nodes representing conjunctions of environmental features) are connected via their pairwise potentials (ψ_{AB}, ψ_{BC}), which penalize the mismatch between *associative* and *metric* pairwise distances (δ_{ij} and d_{ij} respectively). ($t = 0$) Environmental sensory input ($m_{S\to A}(b_A)$ causes an update to the belief of place cell A (i.e. updating its synaptic weights to the grid sheet) by making it fire in a new location $b_A = 1$. ($t = 1$) Place cell A sends a message to B expressing its belief over the location of B, given its own (new) location and the *associative* distance δ_{AB}, causing B to update its belief. ($t = 2$) Messages from B to A and C only cause C to update its belief, so only C broadcasts at the next time-step. **B)** Examples of PC reactivation sequences in loop-closure task for different values of path integration noise σ_{PI}^2 (and therefore pairwise measurement confidence, since $w_{ij} = 1/(\sigma_{PC}^2 + \sigma_{PI}^2\delta_{ij})$; where $\sigma_{PC}^2 = 1e - 4$). Multiple local sequences can occur in interleaved fashion (Middle, grey shading) and become longer and smoother when pairwise measurements are less confident (Right; also **E**, **F**). **C)** The 'Max-Entropy' schedule (i.e., only the place cell with max entropy change fires in the next-step) converges faster than when all PCs broadcast messages at each time-step. **D)** Forward and reverse sequences occurred approximately equally often.

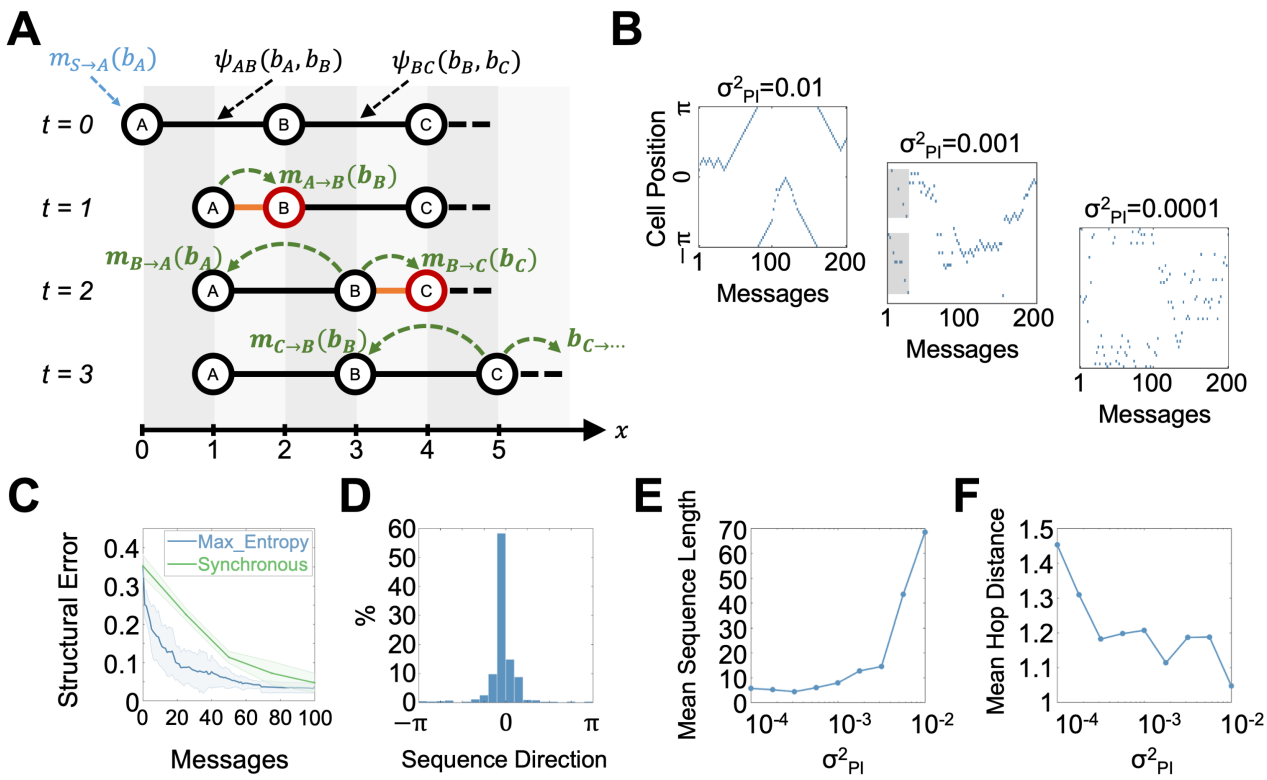

A
$m_{S\to A}(b_A)$
$\psi_{AB}(b_A, b_B)$
$\psi_{BC}(b_B, b_C)$
t = 0
t = 1
t = 2
t = 3
$m_{A\to B}(b_B)$
$m_{B\to A}(b_A)$
$m_{B\to C}(b_C)$
$m_{C\to B}(b_B)$
$b_{C\to\cdots}$
x
B
σ^2_{PI}=0.01
σ^2_{PI}=0.001
σ^2_{PI}=0.0001
Cell Position
Messages
C
Structural Error
Max_Entropy
Synchronous
Messages
D
%
Sequence Direction
E
Mean Sequence Length
σ^2_{PI}
F
Mean Hop Distance
σ^2_{PI}

Figure 7. Schematic neural mechanism for probabilistic message passing (see also Supplementary Video 2). **A)** Illustration in 1D. The broadcasting place cell P_B sends a spike to receiving place cell P_R via recurrent connections in CA3, and also initiates a travelling wave at the corresponding location on the grid cell module via their connections there. (Left) No new learning occurs when the spike and travelling wave arrive at P_R and its corresponding grid location (GC_7) at the same time, as P_R-GC_7 connection will already be strong. (Middle) If the CA3 spike arrives at P_R ahead of the travelling wave reaching GC_7, the synaptic associations of P_R are adjusted towards the currently active GC_5 (updating the belief, by increasing P_R-GC_5 and decreasing P_R-GC_7, see *dW*, below). (Right) Same as middle, except that GC wave reaches GC_7 before the CA3 spike reaches P_R. **B)** Propagating messages as travelling waves in mEC. A neural simulation of travelling waves with a modified Laplacian diffusion kernel (Wave) closely approximates the probabilistic propagation of activity (True), reflecting the accumulation of self-motion noise in the broadening of the wave front. **C)** Comparison of the algorithmic and neural belief propagation performance.

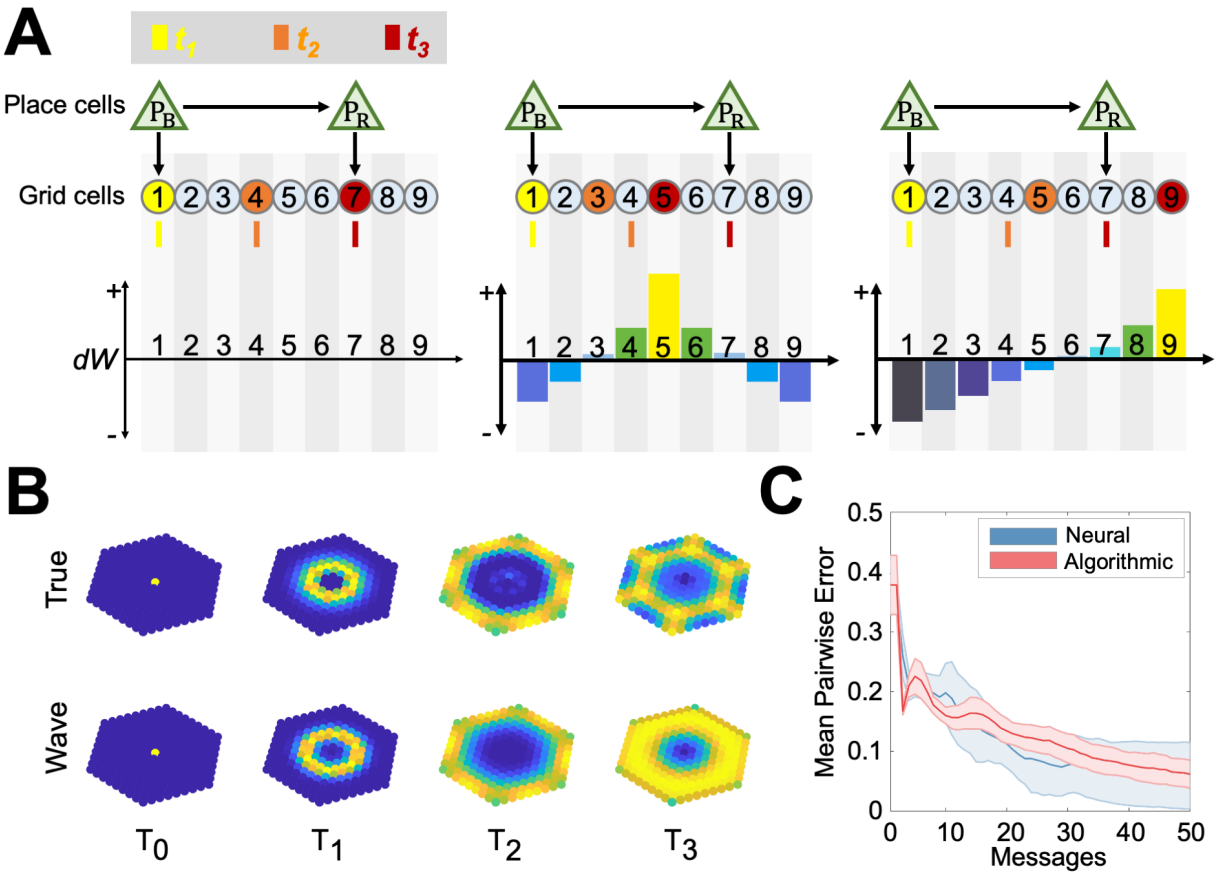

A
t_1
t_2
t_3
Place cells
P_B
P_R
Grid cells
1 2 3 4 5 6 7 8 9
dW
+
-
B
True
Wave
T_0
T_1
T_2
T_3
C
Neural
Algorithmic
Mean Pairwise Error
0.5
0.4
0.3
0.2
0.1
0
0 10 20 30 40 50
Messages

Figure S1 Implementational details of the *online* system. **A** The mapping from real to grid cell space can be considered as the subtraction of a mapping vector $[i \cdot \boldsymbol{v}_1 + j \cdot \boldsymbol{v}_2]$, where each grid field can be described as a point on a 2D lattice with basis vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Each grid field has associated with it a Voronoi domain, defined as the region within which any point is closest to the corresponding grid field. When mapping from grid-to-real space, the vector of the closest grid field is subtracted. **B** The wrapped Normal distribution is a summation of the likelihoods of the current position estimate being at any one of an infinite number of periodic tilings (here, five wrappings are shown). **C** The grid cells are connected periodically to produce a 'twisted-torus' topology. **D** Illustration of the shifter cell mechanism. Each 'readout' grid cell is connected to four 'shifter' grid cells and a single self-connecting cell. **E** Illustration of the difference between wrapped and non-wrapped Gaussian distributions. **F** Correspondence between the neural shifter cell mechanism and the algorithmic transition function. **G** Pairwise distances between place fields can be inferred from the strengths of Hebbian connections. **H** Mean entropy in the beliefs of each place cell over their encoded location in grid space (encoded in place-grid cell connection weight distributions decreases with iterations during *offline* inference.

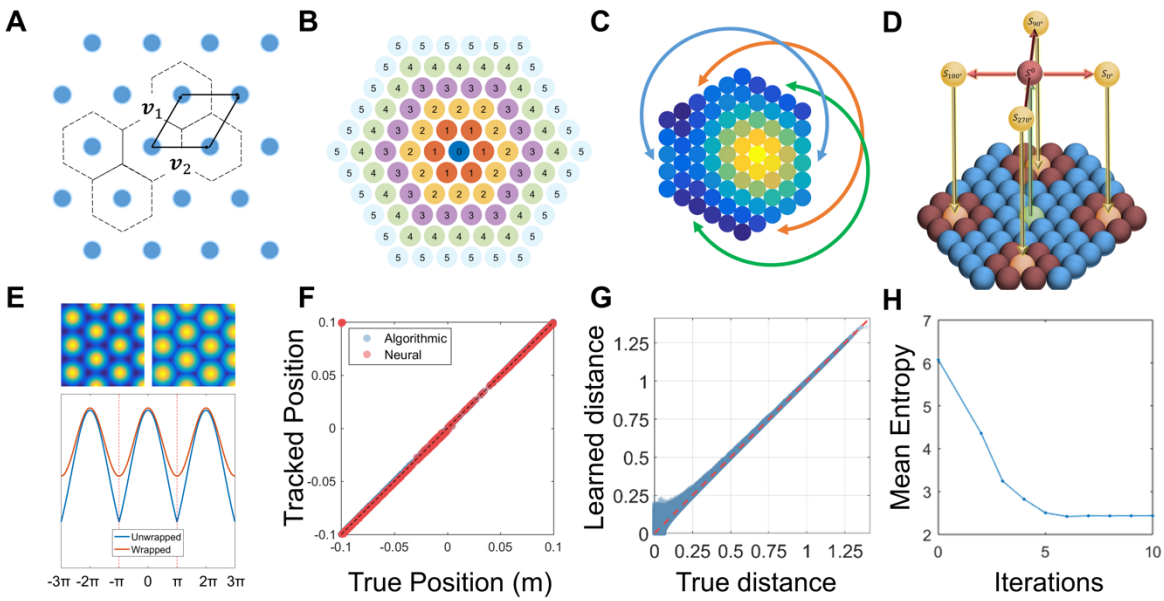


Figure S2 A Convergence of grid patterns and place-grid cell weights over three periods (columns). The place-grid cell weights (*bottom row*, colour denotes normalized connection strength) converge fastest near to the boundaries (as seen in development; Muessig et al., 2015) and corners of an environment, since trajectories through are more directionally constrained. Top row shows corresponding grid patterns (colour denotes normalized firing rate). **B** Convergence of stable grid patterns over sensory noise (σ_{PC}) and path integration noise (σ_{PI}). Colour denotes grid score (Sargolini et al., 2006). *P* indicates number of place cells used in simulations. **C-D** Environmental deformations caused by prior structural beliefs. **C** Place fields generated by the boundary vector cell model (Hartley et al., 2000). An animal's perception of a trapezoidal environment (*right*) may be influenced by place-BVC associations learned in a previous rectangular environment (*left*). In the trapezoid, simulated place fields shift with the wall (**D**), causing similar distortion of the grid pattern.

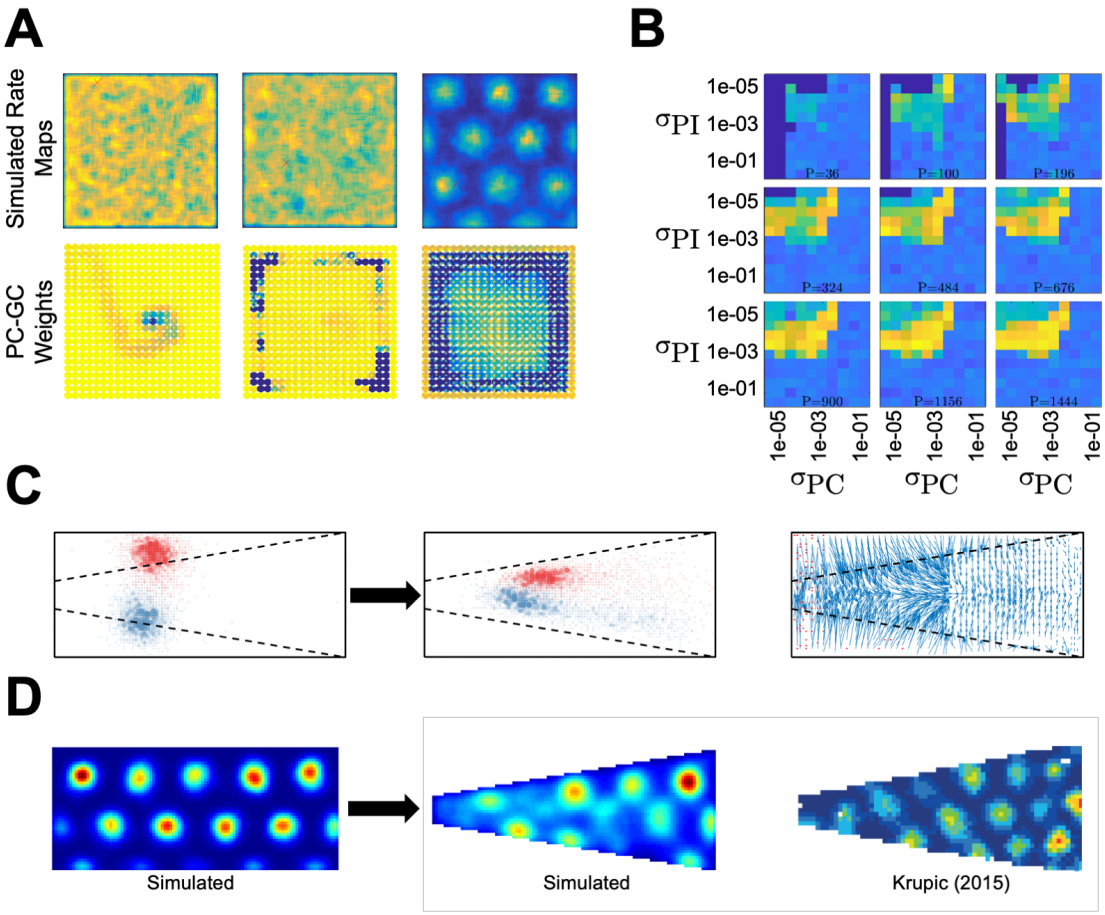


Figure S3. Grid distortions due to inhomogeneous environmental inputs in the *online* model. **A)** When environmental inputs are concentrated in one corner of the environment (*top-middle panel*), the resultant grid cell firing rate maps undergo a shearing distortion which produces an orientation offset (A_{min}; *bottom-left panels*). This offset increases with experience, as grid cell firing becomes increasingly dependent on the maturing sensory inputs (*bottom-middle panel*), matching experiments in which rats always entered the box at the same corner (*top-left panels*; Stensola et al., 2015) The size of the experimental and simulated offsets are similar (*right panels*). **B)** Simulated distortions based on an exponential decay in place cell input from one corner as a function of the decay parameter (R). **C)** Concentrated place field input along one wall and both corners (*bottom left panel*) causes another distortion pattern (*bottom right panel*) also observed experimentally (*top* adapted from Stensola et al., 2015).

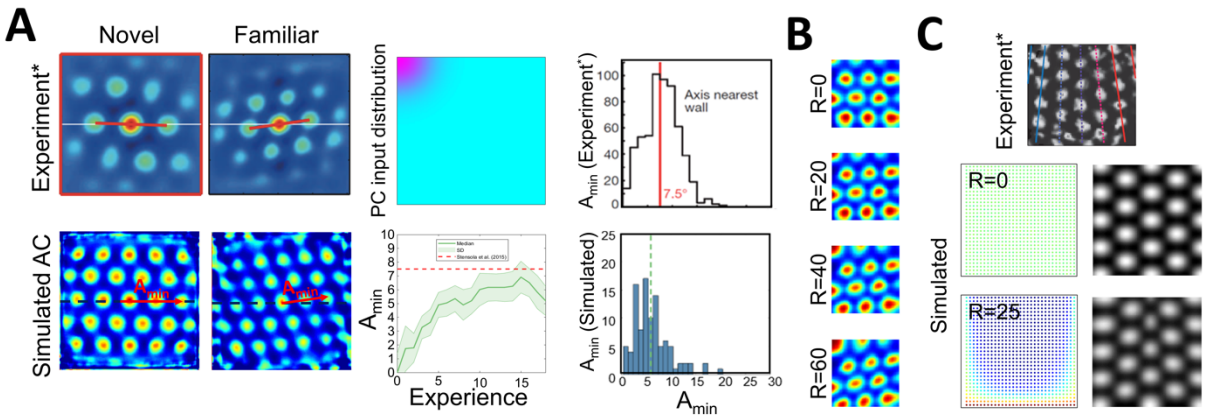


Figure S4 Mathematical analysis of shearing of the grid pattern due to inhomogeneous environmental inputs. **A** The orientation resulting from the shearing operation can be calculated by analysing the angles of the sheared hexagonal lattice describing the centroids of the grid fields. **B, C** The analysis predicts that the orientation offset should be dependent on both the strength of the place fields' density / firing rate imbalance R (see Supplementary Methods 1.7 and Fig. 2) (**B**) and their tuning widths (**C**).

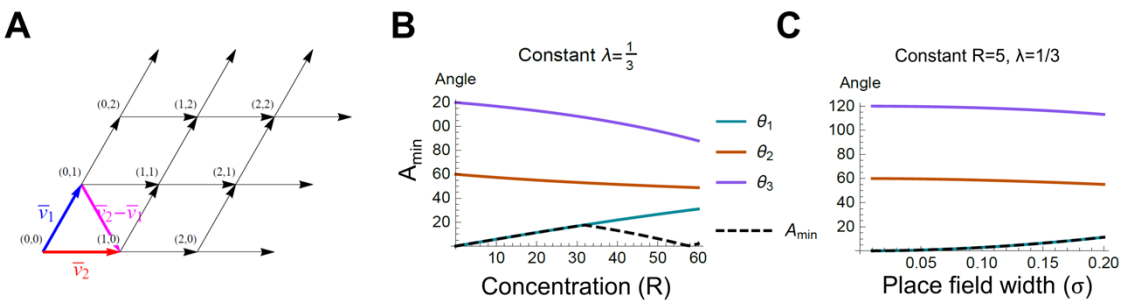


Figure S5 Anatomy of a SLAM system. The joint location-map probability distribution (**A**) is represented in the firing rates and synaptic weights within the HPB-mEC system (**B**).

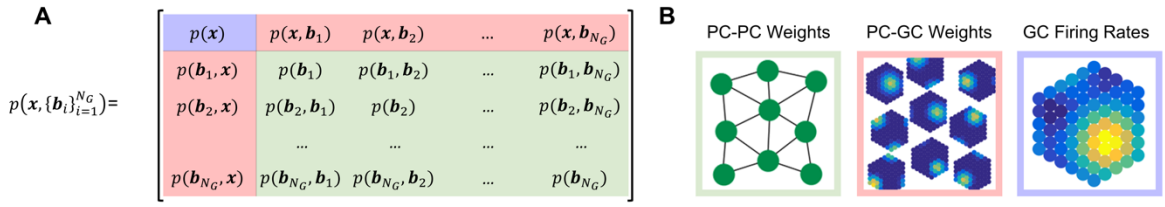


Figure S6 A Spring network analogy of the associative structure of an environment. Edge 'stiffness' is inversely proportional to the variance in the Gaussian observation. **B** 'Wrapping' physical space to encoded location on the grid sheet. Each colour indicates the tiling of the based grid sheet's domain in real space. **C** Pairwise distances near the edges of the environment are overestimated due to under-sampling when the agent preferentially explores the middle of an environment. Colours denote the distance of the pair of PCs i and j from the walls of the 1x1m^2 environment $d_{wall} = \sum_{p=i,j} \frac{1}{2}(\min(x_p, 1 - x_p) + \min\ (y_p, 1 - y_p))$. **D** Resulting local scale is proportional to the occupancy. C When the grid scale is smaller than the size of structure being encoded, we can think of 'wrapping' the structure onto the grid sheet. Here, colours denoted different tilings of the base metric tile (the Voronoi region of a given grid cell).

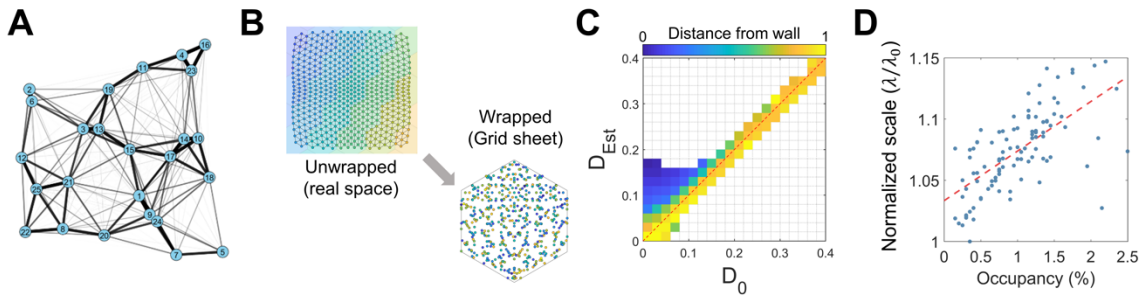


Figure S7 An alternative view of the *online* and *offline* models. Place cells P are driven by real-world stimulus S. During *online* exploration, the *associative* generative model is learned, but does not generate predictions. During *offline* inference, the *metric* generative model is corrected towards the predictions being generated by the *online* model, which becomes a surrogate for sensory stimuli as would be generated by the real world model.

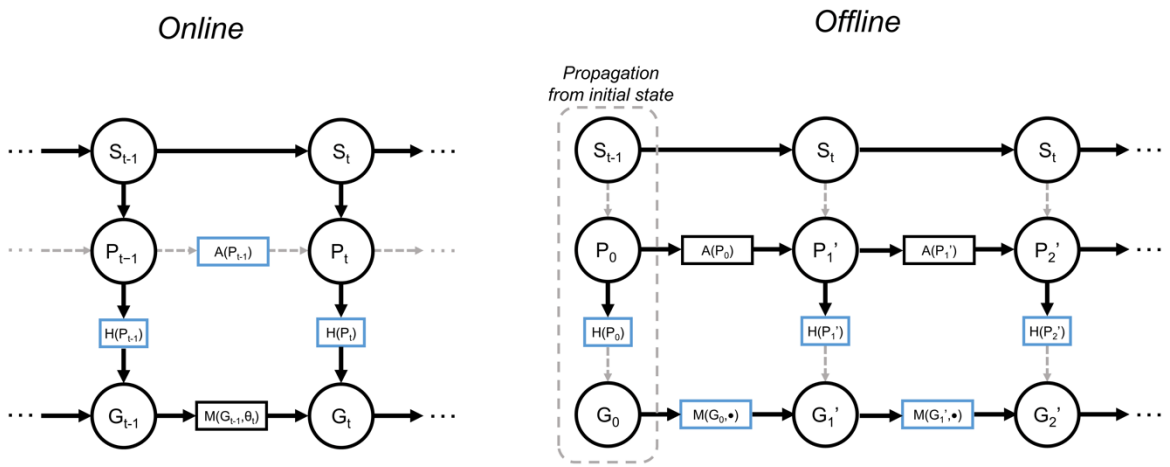


Table S1 Algorithm detailing the overall dual-systems hypothesis of *online* and *offline* localization and learning in the HPC/mEC.

Algorithm: *Online* localization and mapping with prediction error initiated *offline* inference

```
initialize: A~𝕌(0,Δ), B~𝕌(0,Δ)                    % Initialize weights to small values ~Δ
while k<K do:                                      % For the duration of the simulation
    do_movement_update();                          % Update state via path integration
    make_observation();                            % Compute PC firing
    if ε_k > ε_0 do:                               % If prediction error, do offline inference
        while any τ_i<τ_min do:                    % Loop until cells below tension threshold
            u=argmin(τ);                           % Find node with max. message tension
        compute_and_broadcast_message(u);          % Broadcast message to neighbours of u
        for t ∈ Neighbours(u) do:                  % Loop over neighbours of u
            update_belief(t);                      % Update belief of node t
            update_message_tensions(t);            % Compute change in belief of node t
        update_PC_GC_weights();                    % Do associative to metric map learning
    update_PC_PC_weights();                        % Update associative weights
    do_measurement_update();                       % Correct state estimate
```