1 Replay as structural inference in the hippocampal-entorhinal

2 system

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

4 1 Institute of Cognitive Neuroscience, UCL

5 2 Institute of Neurology, UCL

6 3 CoMPLEX, UCL

7 #correspondence

8 **Abstract**

9 Model-based decision making relies on the construction of an accurate representation of the

10 underlying state-space, and localization of one's current state within it. One way to localize is to

11 recognize the state with which incoming sensory observations have been previously associated.

12 Another is to update a previous state estimate given a known transition. In practice, both strategies are

13 subject to uncertainty and must be balanced with respect to their relative confidences; robust learning

14 requires aligning the predictions of both models over historic observations. Here, we propose a dual-

15 systems account of the hippocampal-entorhinal system, where sensory prediction errors between these

16 models during *online* exploration of state space initiate *offline* probabilistic inference. *Offline*

17 inference computes a *metric* embedding on grid cells of an *associative* place graph encoded in the

18 recurrent connections between place cells, achieved by message passing between cells representing

19 non-local states. We provide testable explanations for coordinated place and grid cell 'replay' as

20 efficient message passing, and for distortions, partial rescaling and direction-dependent offsets in grid

21 patterns as the confidence weighted balancing of model priors, and distortions to grid patterns as

22 reflecting inhomogeneous sensory inputs across states.

1

Author Summary

23

24 - Minimising prediction errors between transition and sensory input (observation) models

25 predicts partial rescaling and direction-dependent offsets in grid cell firing patterns.

26 - Inhomogeneous sensory inputs predict distortions of grid firing patterns during *online*

27 localisation, and local changes of grid scale during *offline* inference.

28 - Principled information propagation during *offline* inference predicts coordinated place and

29 grid cell 'replay', where sequences propagate between structurally related features.

Introduction

30

31 Grid cells in the medial entorhinal cortex (mEC), whose firing fields form a periodic hexagonal lattice

32 across the environment, are thought to support path integration[1–3], whereas hippocampal place cells

33 tend to have unimodal firing fields reflecting environmental cues such as boundaries[4,5]. Grid cell

34 firing patterns are stable over time, suggesting corrective environmental inputs, possibly from place

35 cells[6], but rely more on self-motion than place cell firing patterns[7] suggesting that environmental

36 inputs are not fully corrective[2,8,9]. Given estimates of each input's uncertainty, Bayesian inference tells

37 us how they should be optimally combined.

38 Although *online* learning (i.e. using only currently available sensory information) can converge under

39 low PI and sensory noise, robust learning in the presence of noise requires minimizing the error

40 between self-motion and environmental estimates of location across all state transitions[10,11]. Thus,

41 historic observations must be stored and revisited *offline* (i.e. independently of current sensory inputs)

42 to allow propagation of local environmental information to non-local but structurally connected

43 regions of the cognitive map, e.g. as when adapting to a novel shortcut or barrier. This process can

44 also be viewed as an embedding of sensory experience within a low-dimensional manifold (in this

45 case, 2D space), as observed of place cells during sleep[12].

46 Building on previous work[11,13–15], we propose a dual-systems (*online-offline*) account of spatial

47 inference in the hippocampal/entorhinal system, which we define as the process of identifying the

48 configuration of both one's own location (current state) and the location of environmental landmarks

49 in space (c.f. 'SLAM'[10]). In familiar environments, *online* localization (identification of one's own

50 position) is achieved by recursively combining self-motion and sensory inputs, which are mediated by

51 learned *transition* and *observation* models, respectively. However, prediction errors between these

52 models trigger *offline* inference over non-local states, facilitating fast learning of new or changed

53 associative environmental structure, encoded *online* in place-place cell synaptic associations. We

54 identify this *offline* inference with coordinated hippocampal/mEC 'replay'[16–20].

55 Our framework also provides algorithmic- and implementation-level explanations for observed

56 features of grid cell firing in response to manipulations[7,9,21–24] or inhomogeneity[25–29] of environmental

57 sensory input. Overall, these phenomena can be understood in a probabilistic framework, where

58 minimization of prediction errors between the transition and observation models are traded against

59 prior model beliefs.

Results

60

Probabilistic *online* localization with place and grid cells

61

62 Grid cells exist in 'modules' of cells, whose firing patterns have the same spatial scale and orientation

63 relative to the environment, but differ in their spatial offsets[30]. The spatial scale increases in discrete

64 steps along the dorso-ventral axis, suggesting that, across modules, GCs support a hierarchical

65 representation of space[21,24,31–33]. Here, we consider a single module of GCs, whose activity represent a

66 probability distribution over a periodic, discretized region of space (visualised as a topographically

67 arranged sheet of cells; Fig. 1A).

68 The self-location distribution is maintained over time by recursively integrating sensory and self-

69 motion inputs, accounting for their uncertainties (Fig. 1A). Firstly, the posterior distribution over

70 agent location (grid module activity) from the previous time-step $\boldsymbol{G}$ is updated given noisy perceived

71 movement $\hat{\boldsymbol{u}}$ via the *transition model T* (see Methods):

72
$$G'(\boldsymbol{x}) = \int T(\boldsymbol{x}, \boldsymbol{x}'|\hat{\boldsymbol{u}}, \boldsymbol{\alpha}) \cdot G(\boldsymbol{x}')d\boldsymbol{x}' \tag{1}$$

73   where x is the 2D coordinate of the agent location in *metric* space (corresponding to a particular grid

74   cell) and x' the location at the previous time-step. Biophysically, T would be represented by a

75   population of direction dependent 'shifter' cells with asymmetric recurrent weights[1] with a circulant

76   structure[34] (Fig. S1C, see Methods) learned *apriori* (but see Refs. [35–37]). The rate of translation of

77   activity on the grid sheet in response to movement $\hat{u}$ is controlled by the *transition model* gain $\alpha =$

78      $[\alpha_x, 0; 0, \alpha_y]$, which might correspond to the strength of the associations to, or the speed dependence

79   of, shifter or conjunctive cells[31,38,39] (see Supplementary Methods).

80   The *transition model* estimate $G'(x)$ is then refined by observations of environmental features, which

81   map to *metric* space locations via *observation model H*:

82
$$G(x) = \frac{1}{K} \ H(P|x) \cdot G'(x) \tag{2}$$

83   where P is a vector of place cell firing rates, firing of place cell i representing the likelihood of the

84   presence of a specific sensory feature, or combination thereof. In our simulations, these have unique

85   locations in physical space μ_i and receptive field widths Σ_{PC}^i. Where the number of grid cells is large,

86   the weights from place cell i to the grid module define a distribution for that feature's estimated

87   location in *metric* space (Fig. 1A). The weighted projection of place cell activity by these weights

88   defines the *observation model H*. K provides inhibitory normalization (see Supplementary Methods).

89   *Online* learning modifies the *observation model* to reflect the current *transition model* location

90   estimate (induces synaptic changes in the place-grid cell connection weights via a BCM rule; Fig. 2A;

91   Methods). *Online* learning produces stable grid patterns (due to the circulant structure of T) for a

92   range of levels of PI and sensory noise, but convergence fails in higher noise regimes (Fig. S2B).

93   After a short period of initial learning, stable grid patterns emerge in the integrated estimate, despite

94   the pure PI estimate being too noisy and the sensory associations too immature to drive stable

95   patterns, if operating independently (Fig. 1B).

Offline inference: The hippocampus as a probabilistic graph

96

97 Local, *online* learning is not robust in novel environments, because corrections to the estimated agent

98 location (current grid cell activity), e.g. upon encountering familiar environmental features (place cell

99 activity) associated to a different location on the grid module, also imply corrections to the encoding

100 of feature observations along the preceding trajectory[10]. That is to say; local updates to the *cognitive*

101 *map* also imply non-local, structurally associated changes. Formally, probabilistic spatial inference in

102 this case requires finding the most likely configuration of *metric* space feature locations

103 $\{\boldsymbol{b}_i\}_{i=1:N_P}$ (each $\boldsymbol{b}_i$ is the 2D coordinate of feature i in *metric* space, i.e. a place - grid cell association)

104 and agent location $\boldsymbol{x}$ (the distribution over which is indicated by the grid cell firing rates) consistent

105 with environmental sensory observations made along a given trajectory.

106 Theoretically, the configuration $\{\boldsymbol{b}_i\}_{i=1:N_P}$ can be recovered purely from the distances between pairs

107 of environmental features[10] (Fig. 1F, *"square"*, *"ring"*). Importantly, despite both feature locations

108 being susceptible to large absolute errors (due to noisy PI), the errors will be correlated such that

109 pairwise distance measurements will decrease in variance with observations[10]. This method predicts

110 characteristic failure modes when the pairwise distance information is ambiguous or incomplete (e.g.,

111 Fig. 1F; *"broken ring"*). New distance observations might also cause dramatic changes to the inferred

112 configuration (e.g. the discovery of a shortcut). If the current *absolute* estimates of feature node (i.e.

113 place cell) location are stored in the place-grid cell synaptic weights, we propose that the *relative*

114 *distances* between pairs of features are stored in the recurrent weights between place cells in

115 hippocampal region CA3.

116 Consider a spring network, where the edge between environmental feature nodes i and j represents a

117 noisy pairwise observation with length reflecting pairwise distance and stiffness reflecting

118 certainty[13,40] (Fig. 1F, S6A). Minimizing the elastic energy in the spring mesh system corresponds to

119 finding the maximum of the joint likelihood $L(\cdot)$, which is a function of the feature locations in *metric*

120 space $\{\boldsymbol{b}_i\}_{i=1:N_P}$ and the internal gain parameter $\boldsymbol{\alpha}$, given pairwise distance measurements with

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

122   the posterior:

123

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

124   where the current PC-GC weights B (B_i is the probability distribution of the location of feature i and

125     $B_i(\boldsymbol{b}_i)$ is its value at *metric*/grid module location $\boldsymbol{b}_i$) act as priors on the feature node locations, the

126   pairwise potential terms $\psi(\cdot)$ penalize the difference between *associative* pairwise distance

127   measurements δ_{ij} made directly in environmental *stimulus* space, and the distance between their

128   candidate locations in *metric* space $\boldsymbol{b}_i - \boldsymbol{b}_j$. E is the set of connected PCs (see Methods). Distance in

129   *metric* space is also a function of the *transition model* gain ($\boldsymbol{\alpha}$), which has a Gaussian prior $p(\boldsymbol{\alpha}|\boldsymbol{\alpha}_0)$

130   (a larger $|\boldsymbol{\alpha}|$ will decrease the *metric* space distance for all pairs; see Methods). Maximizing the

131   likelihood (finding the state of minimum energy in the spring network) model over all feature node

132   pairs minimizes the total prediction error between *associative* and *metric* generative models of the

133   world[15] (Fig. S7).

134   The *associative* distances can be straightforwardly learned during *online* exploration. Since Hebbian

135   learning reflects coactivity, a trajectory exploring the environment uniformly results in synaptic

136   strengths between place cells proportional to the spatial correlation between their receptive fields[41]. The

137   Euclidean separation between their fields is then accessed via a simple transformation (see Methods;

138   Fig. S1G). In this context, learning the PC-GC weights (modifying the *observation* model) during *online*

139   localization corresponds to forming spatial priors over feature locations which anchor the structure,

140   which would otherwise be translation or rotation invariant (since measurements are relative), learned

141   during *offline* inference to constant locations on the grid-map. Taken together, our framework proposes

142   a mapping onto anatomy of the joint *agent-feature* location distribution required for full probabilistic

143   inference over environmental structure (Fig. S5; See overall algorithm in Table S1).

Partial grid pattern responses to environmental rescaling

144

145 Uniform rescaling of an environment will introduce a mismatch between the estimates of location

146 from the *transition* and *observation* models (i.e. a 'prediction error'). To minimize these prediction

147 errors, the *offline* system can either modify the *transition* model gain to match the current

148 environmental input (Fig. 2A, bottom), or modify the mapping from environmental inputs to *metric*

149 space in the *observation* model (Fig. 2A, middle; see Methods). The degree to which either is

150 modified should reflect their relative confidences, specified by a 'transition confidence score' ($T_C =$

151 $\sigma^2_{Obs}/\sigma^2_\alpha$, the ratio of confidence in the *transition* vs *observation* models; see Methods). Similarly, a

152 'prior confidence score' P_C specifies how much the system will tolerate persistent prediction errors; if

153 P_C is large, optimization may favour preserving prior configurations, as opposed to alignment of the

154 current *transition* and *observation* models (see Methods).

155 We modelled experiments in which the physical environment[21,42] and perceived velocity through a

156 visual virtual environment[7] were re-scaled, such that self-motion and sensory inputs conflicted. In

157 both experiments, the rescaling of the grid patterns was *partial*, i.e. less than the magnitude of the

158 physical or virtual manipulation, and less than those of place fields.

159 Both manipulations can be simulated by introducing a visual gain parameter α^{Visual} to the simulation

160 of the environment (in both experiments it scales the amount of self-motion required to traverse the

161 width of the perceived environment). Learned *associative* distances ($\boldsymbol{\delta}$) are also scaled by this

162 parameter, reflecting its effect on the temporal overlap of place fields (see Methods Eq. 12). We

163 simulated grid pattern rescaling responses over a range of transition confidence scores T_C. When

164 confidence in the *transition* model is high ($T_C \rightarrow \infty$), grid patterns in the real world are unchanged

165 when plotted against physical movement (but are changed when plotted in visual VR coordinates; Fig.

166 2B, first column). The opposite is true when confidence in the *observation* model is high: grid

167 patterns are unchanged relative to the apparent environment (Fig. 2B, last column).

168 However, for intermediate T_C values (i.e. balanced confidence in the *transition* and *observation*

169 models), the model predicts partial rescaling of the grid pattern relative to the size of the manipulation

7

170   (Fig. 2B, middle column), matching the observed grid patterns in both experiments[7,21] (and a similar

171   third experiment in on a virtual linear track[43]; Fig. 2C).

Differential grid and place field responses to environmental reshaping

173   How does the *offline* system respond to more complex environmental deformations? When one wall

174   of a familiar rectangular environment is rotated inwards by $45°$[22], place fields near the wall shifted

175   almost fully while fields further away remained largely stationary, consistent with place fields

176   preferentially reflecting local environmental inputs[5]. In contrast, grid fields shifted only partially near

177   to the manipulated wall. Using the observed place field shifts[23], we simulated the response of grid

178   cells to the same manipulation (Fig. 3; see Supplementary Methods). Shifted place fields induce a

179   misalignment between the *associative* distances and the distance between their encodings in *metric*

180   space. The place field shifts are local and non-uniform, and so misalignment cannot be corrected by a

181   global change to the *transition gain* α. Indeed, α is not significantly modified during the optimization

182   process, regardless of the T_C value. Instead, alignment between the *transition* and *observation* models

183   is maximized by modifying the *observation model*, i.e. updating the locations of the place fields on

184   the grid module.

185   If there were no confidence in the prior *observation* model ($P_C \rightarrow 0$), it would be modified offline to

186   match the transition model, leaving the grid pattern unperturbed by the environmental change (Fig.

187   3B,C, top row). At the other extreme, favouring prior beliefs over recent observations (i.e. pairwise

188   distances encoded during the manipulation trial) would result in an unchanged *observation* model, and

189   grid field shifts that exactly mirror corresponding place field shifts (Fig. 3B,C, bottom row). In this

190   regime, there would be permanent misalignment of the *transition* and *observation* models during

191   *online* localization, producing noisy grid patterns, as when simulating a related experiment where grid

192   distortions were observed in trapezoid environments[44] (Fig. S2D). Setting P_C to an intermediate value

193   reproduces the experimentally observed partial shifting of grid fields (relative to the place fields[23])

194   when visualizing the structure encoded in the *observation* model (i.e. assuming low confidence in the

195   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).

221 Firstly, relative behavioural under-sampling of the place fields near the boundaries of the environment

222 (using occupancy statistics from[26]; Fig. 4A, bottom) lead to weaker PC-PC connections, and

223 consequent overestimation of their pairwise distances, producing local scale changes (Fig. S7C).

224 Secondly, since Hebbian learned connection weights between place cells reflect the correlation in

225 their firing, and therefore their statistical discriminability[46], two place cells with broad receptive fields

226 would develop a stronger connection than a pair with equal separation but narrower receptive fields

227 (stronger connections correspond to shorter distances on the grid module, producing grid patterns with

228 larger scales in the environment; Fig. 1D). Another recent study[47] suggests that place fields are

229 narrower near the edges of an environment, consistent with greater precision when driven by more

230 proximal environmental features[5] (Fig. 4A, top row). In our model, this produces weaker recurrent

231 connections and a shrinking of the grid pattern at the edges of the environment following *offline*

232 inference (Fig. 4D,E, top row).

233 Together, our results suggest that the cognitive 'distance' between two sensory features should be

234 greater both when the absolute confidence in their spatial locations is greater (reflecting an increased

235 statistical discriminability), or when those features are under-sampled relative to other features.

236 Although the action of both mechanisms are independent their effect is the same; both i) relative

237 under-sampling of the transition between two adjacent states and ii) a reduced statistical

238 discriminability between those states, both contribute to a weaker pairing of their representative place

239 cells, resulting in greater separation between their encodings in *metric* space and a locally larger grid

240 scale when 'read-out' in the firing pattern (a locally larger perception of distance).

241 In contrast, global shear-like distortions[29] and associated 7.5-8° offsets of one of the grid axes[29,44] can

242 be interpreted as *localization* induced distortions during *online* exploration. In Stensola et al.[29], rats

243 were introduced into the same corner of the box at the start of each trial; in Butler et al.[25], shearing

244 developed following the introduction of reward[25]. In both experiments, shearing developed with

245 increasing experience[25,29]. We hypothesized that these distortions reflect an increasing effect of non-

246   uniform environmental inputs to the grid module, either reflecting their natural distribution[25,29] or

247   inhomogeneous behavioural sampling of environmental locations[26].

248   In our simulations, given a learned map, biasing the strength of sensory inputs at specific locations

249   (e.g. one/two corners) during *online* exploration reproduced several experimentally characterized

250   global distortions by causing a bias in the decoding of location (i.e. salient locations contribute a

251   larger 'vote'; Fig. S3; see Supplementary Methods).

Probabilistic inference through HPC-mEC message passing

253   To this point we discussed, from a functional perspective, how the brain might optimize its internal

254   representations to reflect the uncertainty of sensory information. But how might the brain perform this

255   optimization? In the above analyses of *offline* inference, we numerically computed the maximally

256   likely feature locations on the grid module. However, the system must also track the uncertainty in

257   these estimates, which would require updating the place-grid cell weights (including those with firing

258   fields far from the agent location). An update of the full weight distributions is generally intractable

259   when the state space is large.

260   Belief propagation[48] is a technique for approximating this inference on graph structured data, and

261   comprises two stages. First, a given feature node (i.e. a place cell) computes its location distribution

262   (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

263   from its connected neighbours (Fig. 5C; see Methods for details). A message $m_{i \to j}(\boldsymbol{b}_j)$ expresses

264   neighbour node i's belief of node j's location, conditioned on its own distribution, and is dependent

265   on the same pairwise potential terms $\psi(\cdot)$ in Eq. 3. The effect of a message is to favour distributions

266   of nodes i and j which locate them at a radial distance equal to the associative distance δ_{ij}; causing

267   messages to be expressed as rings centred on the belief of the broadcasting node (Fig. 5C). Resolving

268   a feature's unique location then depends on aggregating messages from multiple neighbours (Fig. 5C).

269   Computations are distributed, and importantly only require information that is local to each neuron.

270   Each node in the graph iterates between updating its belief and broadcasting messages, converging

271   when new messages cease to change the beliefs of their recipient nodes. As expected, the reduction in

272   pairwise prediction error between associative distances and their corresponding distances in grid

273   space (see Methods) over successive message iterations is accompanied by a sharpening of the

274   distribution of each feature's location on the grid module (see Supplementary Methods; Fig. S1H).

275   ## *Offline* inference triggered by prediction errors

276   How might the *online* and *offline* systems interact? If the *online* system is sufficient to localize within

277   pre-learned, simple or slowly changing environments, non-local reactivations of place cells would be

278   unnecessary. However, more complex *offline* inference is required under more demanding

279   circumstances, or in novel or changing environments. We hypothesize that *offline* or 'remote'

280   inference is triggered by *prediction errors* between location estimates from the *transition* and

281   *observation* models, respectively (Fig. 1G), defined in our model as the Kullback-Leibler divergence

282     $\epsilon = \mathbb{KL}(G', H)$.

283   Prediction errors are large when the *observation model* prediction (weighted place cell input) is

284   different and more sharply peaked than the *transition model* estimate (Fig. 1G; see Methods;

285   prediction errors will not be generated in absence of incoming sensory information, as in darkness,

286   when the observation model estimate is uncertain).

287   To illustrate our *dual-systems* (*online+offline*) hypothesis (Fig. S7), we simulated an agent navigating

288   around a novel circular track (the *loop closure* task; Fig. 5). Completion of the first lap produces

289   positive prediction errors between the sharply peaked input from feature inputs learned at the

290   beginning of the trial, and the agent location estimate which is uncertain given the accumulation of PI

291   noise (Fig. 5B).

292   Decrease in structural error (the difference between the place field separations and their encoded

293   separations on the grid cell sheet) following *online+offline* inference was markedly larger than

294   following *online* learning alone (Fig. 5E). The inferential power of this 'one-shot' learning process

295   derives from consideration of the full covariance structure of the feature locations (captured by the

296   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).

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).

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).

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

14

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

15

377   translation of activity on the grid cell sheet is driven by cells with velocity-dependent firing rates[38,39].

378   Indeed, a recent study showed that velocity dependence in mEC firing is tied to environmental

379   manipulations[39].

380   We propose that *offline* structural inference events correspond to coordinated HPC/mEC replay[16–19,58–

381   60], which can be viewed as synchronizing predictions from *associative* (CA3) and *metric* (mEC)

382   generative models (Fig. S7). In this way, structural changes to an environment can be propagated to

383   non-local regions of the *metric* embedding, in contrast to models in which these states need to be

384   physically revisited[13], consistent with the observation that replays do not necessarily repeat

385   experienced trajectories[51]. Prediction errors between the two models may trigger replay events and

386   corresponding sharp-wave ripples[61,62]. To our knowledge, this is the first functional model of

387   coordinated place cell-grid cell replay[18] (although cf. Ref. [63]), and provides an alternative to reward-

388   based theories[64,65] (we note that rewards may themselves represent salient sensory features,

389   independent of their reward value).

390   Our model makes a number of experimentally testable predictions. Firstly, systematic manipulation of

391   the discriminability of sensory cues distributed within an environment should produce predictable

392   distortions to the grid pattern, observed with increasing experience of an environment. Secondly,

393   replay should be more frequent after structural changes such as shortcuts, blockages or gain

394   manipulations as in the experimental setup of Fig. 5. Thirdly, replay events triggered by specific

395   unexpected sensory observations should become less frequent (Fig. 5B) and smooth (Fig. 6E, F) with

396   continued experience, if the observations remain stable. Fourthly, multiple local replay events may

397   occur in inter-leaved fashion (Fig. 6C, *middle*, grey shading). Fifthly, we predict the existence of

398   travelling waves in grid cells (as a function of their spatial phase; see also[66,67], Fig. 7). Lastly, initial

399   messages propagating from the animal's current location may not cause subsequent messages in

400   remote regions of the graph which are already sufficiently converged (messages will not cause

401   changes in the beliefs of their recipients), although activity in mEC will continue to propagate. Thus

402   grid cell replay could thus be detectable in the absence of simultaneous place cell replay[19] (but place

403   cell replay requires the grid cell *transition* model and so depends on mEC[68]).

16

404   Our proposed structure learning framework can account for diverse phenomena observed in the HPC-

405   mEC system, and makes several novel, experimentally testable predictions.

Methods

Online recursive Bayesian estimation

408   The transition matrix T defines the probability of transitioning from agent location x' to location x,

409   and is a function of the perceived current velocity $\hat{u}$ and *transition* model gain $\alpha = [\alpha_x, 0; 0, \alpha_y]$.

410   Since our *metric* space is periodic, T accounts for cyclic transitions c_{mn}, with Gaussian noise

411   proportional to the perceived velocity $\hat{u} \sim u + N(0, \text{diag}(u)\Sigma_{\text{PI}})$:

$$T(x, x'|\hat{u}, \alpha) = \sum_{mn=-\infty}^{\infty} f(x - x'|\hat{u} + c_{mn}, \text{diag}(\hat{u})\alpha^{-1}\Sigma_{\text{PI}}) \tag{7}$$

413   where $f(x|\mu, \Sigma)$ is a multivariate Gaussian PDF, $c_{mn} = 2\alpha^{-1}(mv_1 + nv_2)$ and $v_1 =$

414    $[\cos(\phi), \sin(\phi)]$ and $v_2 = [\cos\left(\phi + \frac{\pi}{3}\right), \sin\left(\phi + \frac{\pi}{3}\right)]$ define the unit vectors of a hexagonal lattice[69]

415   with grid pattern orientation ϕ and $\text{diag}(\cdot)$ produces a diagonal matrix from a vector input. Since

416   most of the mass is associated with shorter transitions, in practice we approximate the full distribution

417   with a finite number of periodic summations (i.e. ignore the tails; 5 cycles in our simulations; Fig.

418   S1B).

419   The *observation* model defines the likelihood of the current environmental sensory inputs (i.e. the

420   population vector of place cell firing P, where $p_i(\overline{x}) = f(\overline{x}|\mu_i, \Sigma_{PC}^i)$ is the firing rate distribution of

421   place cell i over physical space $\overline{x}$) given the predicted *metric* location x, via a thresholded weighted

422   sum: $H(P|x) = [\sum_i B_i(x)P_i]^+$. Here, $B_i(x)$ is the location distribution of landmark i in *metric* space,

423   which would be encoded biophysically in the learned $[N_P \times N_G]$ matrix B of synaptic weights from

424   place to grid cells (i.e. the gth row and i^{th} column of B is the distribution $B_i(\cdot)$ evaluated at the

425   location of the gth grid cell). The normalization constant $K = \int G(x)dx$ in Eq. 2 simply sums over

426   the current grid cell activity and might biophysically be implemented by inhibitory interneurons.

427 *Online* learning of structural priors

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

429
$$\begin{aligned}
\boldsymbol{B} &\leftarrow & \boldsymbol{B} + \tau_{PG}\boldsymbol{P}^{\mathsf{T}} \otimes [\boldsymbol{G}' \odot (\boldsymbol{G}' - \boldsymbol{\theta})] &\quad (8a)\\
\boldsymbol{\theta} &\leftarrow & \boldsymbol{\theta} + \tau_{\theta}[(\boldsymbol{G} \odot \boldsymbol{G}) - \boldsymbol{\theta}] &\quad (8b)
\end{aligned}$$

430   where $\tau_{PG} = 1e - 4$ is the learning rate and $\boldsymbol{G}'$ and $\boldsymbol{H}$ are column vectors whose elements are the

431   *transition* and *observation* models estimated at the locations of a finite set of grid cell locations. $\odot$ is

432   the element-wise (Hadamard) product between two vectors and $\otimes$ is their outer product. The sliding

433   threshold $\boldsymbol{\theta} \in R^{1 \times N_G}$ provides adaptive synaptic normalization, where $\tau_{\theta} \approx 10\tau_{PG}$. Learning takes

434   place between the *apriori* distribution $\boldsymbol{G}'$ and the current sensory observation $\boldsymbol{H}$ (i.e. before the

435   observation correction to $\boldsymbol{G}'$).

The *offline* probabilistic graphical model

437   The pairwise potentials $\psi\big(\boldsymbol{b}_i - \boldsymbol{b}_j, \boldsymbol{\alpha}|\delta_{ij}\big) = \sum_{m,n=-\infty}^{\infty} \exp\big(-w_{ij}\big(\delta_{ij} - d_{mn}(\boldsymbol{b}_i - \boldsymbol{b}_j, \boldsymbol{\alpha})\big)^2\big)$ penalize

438   differences in the pairwise distances encoded by *association* δ_{ij}, and those that would be computed by

439   comparing their absolute encodings in *metric* space $d_{mn}(\cdot)$; i.e. they encourage a metric embedding

440   that reflects the associative distance. The *metric* distance function $d_{mn}^2(\boldsymbol{x}, \boldsymbol{\alpha}) = (\boldsymbol{x} + \boldsymbol{c}_{mn})^{\mathsf{T}} \boldsymbol{\alpha}(\boldsymbol{x} +$

441     $\boldsymbol{c}_{mn})$ defines the pairwise distance between the encoding of locations i and j in mEC in *metric* space,

442   and is dependent on the gain factor $\boldsymbol{\alpha}$ of the *transition model*. Pairwise measurements are assumed to

443   have confidence w_{ij} (inverse variance) that increases with decreasing inferred distance (i.e. $w_{ij} =$

444     $1/(\sigma_{PC} + \sigma_{PI}\delta_{ij})$). The *transition model* gain is assumed to have a Gaussian prior $p(\boldsymbol{\alpha}) =$

445     $\exp\big(-w_{\alpha}(\boldsymbol{\alpha} - \boldsymbol{\alpha_0})^{\mathsf{T}}(\boldsymbol{\alpha} - \boldsymbol{\alpha_0})\big)$, the w_{α} term representing the confidence in the prior gain value $\boldsymbol{\alpha_0}$.

446   The periodic offset term $\boldsymbol{c}_{mn}$ is the same as defined for the transition model.

Associative encoding in the hippocampus

448   The associative distances are recovered from the $[N_P \times N_p]$ synaptic weights in CA3 $\boldsymbol{A}.$ Under a

449   random-walk behavioural trajectory, the simple modified Hebbian learning rule:

450
$$\boldsymbol{A} \leftarrow \boldsymbol{A} + \tau_{PP}[\boldsymbol{P}\boldsymbol{P}^{\mathsf{T}} - \boldsymbol{A} \odot \boldsymbol{A}] \qquad\qquad (9)$$

18

451   where $\odot$ is the element-wise (Hadamard) product and τ_{PP} the learning rate. The synaptic weights can

452   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

453   rates of two PCs[70]. Where place fields have uniform receptive field widths (σ_{PC}) and peak firing rates,

454   the Euclidean distance between place fields i and j can be inferred via the simple transformation[46]:

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

456   The recovered distance is therefore scaled by the receptive fields' variance (the Bhattacharyya

457   distance[46]), and so relates to 'discriminability' (Fig. 1D). CA3 synapses effectively average over

458   multiple pairwise measurements. By assuming that noise in the pairwise distance measurements scale

459   linearly with distance, both the mean and variance of the Gaussian describing this distribution is

460   efficiently encoded in a single PC-PC synapse.

461   ## Simplified analysis of the probabilistic graphical model likelihood

462   To characterize model predictions in the environmental rescaling[71] and gain change[7] experiments, we

463   studied a reduced version of the full graphical model (Eq. 3; see Supplementary Methods for full

464   derivation). In 1D, given a linear *observation model* $x = H(x') = K_1 + Kx'$ and a large number of

465   evenly spaced place fields, Eq. 3 simplifies to (see Supplementary Methods):

$$-\log L = \ell \; \tilde{\propto} \; \left(1/\alpha^{Visual} - K/\alpha\right)^2 + P_C(K - K^0)^2 + P_C \cdot T_C(1/\alpha - 1/\alpha^0)^2 \qquad (11)$$

467   which can be solved analytically. A similar reduction was applied to the 2D case when considering

468   differential shifts in grid and place fields[23].

469   ## Belief propagation for o*ffline* inference

470   Belief propagation[48] is an iterative, two-stage local message-passing scheme in which, at each

471   iteration n, a feature node (i.e. a place cell) first updates its own belief (connections to the grid cells)

472     $B_i^{(n)}$ by integrating messages from connected nodes $j \in \Gamma_i$ with its own prior belief $B_i^{(0)}$ (Fig. 5C):

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

19

474 The message from node j to node i ($m_{j \to i}$) communicates its belief over the distribution of the

475 locations of place cell i in grid cell space, conditioned on its own location distribution:

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

477 where the pairwise potentials $\psi(\cdot)$ are the same as those described in the full likelihood function (Eq.

478 3). The graph converges when new messages cease to change the beliefs of their recipient nodes.

Scheduled message passing on the place cell graph

480 'Synchronous' belief propagation computes belief updates for each step before broadcasting all new

481 messages in the next step. In simulations, we demonstrated that scheduling message broadcasts based

482 on internal 'message tension' (divergence between previous and updated belief given new messages)

483 produced faster and more accurate convergence (Fig. 6B; see also Elidan et al.[49]). Message tension

484 between the node's previous $\boldsymbol{B}_i^{(n-1)}$ and updated $\boldsymbol{B}_i^{(n)}$ beliefs is defined as:

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

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

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

488 When the message tension is below a pre-defined threshold $\mathcal{T}_{min}$, a node has converged and ceases to

489 broadcast new messages. This mechanism is similar to the prediction error between *transition* and

490 *observation* models used to trigger *offline* inference, with the exception that it uses the symmetric

491 divergence measure $\mathbb{JS}$ rather than $\mathbb{KL}$.

Traveling waves in neural media

493 In simulations, the traveling waves in mEC are simulated explicitly by calculating the true messages

494 conditioned on the sending nodes' current beliefs at each time-step (Eq. 5B). In the 'neural model'

495 (Fig. 7), messages were approximated as waves propagating radially from an initial stimulation on the

496 mEC sheet using a modified mechanical wave model, as used to describe oscillations in water:

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

498 where c is the speed of wave propagation, $[\cdot]^+$ is a threshold linear activation function, and the

499 modified spatial Laplacian operator ∇' is a symmetric 2D Gaussian filter with variance equal to the PI

500 noise (see Supplementary Methods for extended discussion).

Acknowledgements

References

509 1. Burak, Y. & Fiete, I. R. Accurate path integration in continuous attractor network models of grid

510   cells. *PLoS computational biology* **5**, e1000291 (2009).

511 2. Fuhs, M. C. & Touretzky. A Spin Glass Model of Path Integration in Rat Medial Entorhinal

512   Cortex. *Journal of Neuroscience* **26**, 4266–4276 (2006).

513 3. McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I. & Moser, M.-B. Path integration and

514   the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience* **7**, 663–678 (2006).

515 4. Lever, C., Burton, S., Jeewajee, A., O'Keefe, J. & Burgess, N. Boundary Vector Cells in the

516   Subiculum of the Hippocampal Formation. *Journal of Neuroscience* **29**, 9771–9777 (2009).

517 5. Hartley, T., Burgess, N., Lever, C., Cacucci, F. & O'keefe, J. Modeling place fields in terms of

518   the cortical inputs to the hippocampus. *Hippocampus* **10**, 369–379 (2000).

519   6. O'Keefe, J. & Burgess, N. Dual phase and rate coding in hippocampal place cells: Theoretical

520   significance and relationship to entorhinal grid cells. *Hippocampus* **15**, 853–866 (2005).

521   7. Chen, G., Lu, Y., King, J. A., Cacucci, F. & Burgess, N. Differential influences of environment

522   and self-motion on place and grid cell firing. *Nature communications* **10**, 630 (2019).

523   8. Hardcastle, K., Ganguli, S. & Giocomo, L. M. Environmental Boundaries as an Error Correction

524   Mechanism for Grid Cells. *Neuron* (2015) doi:10.1016/j.neuron.2015.03.039.

525   9. Keinath, A. T., Epstein, R. A. & Balasubramanian, V. Environmental deformations dynamically

526   shift the grid cell spatial metric. *eLife* **7**, e38169 (2018).

527 10. Durrant-Whyte, H. & Bailey, T. Simultaneous localization and mapping: part I. *IEEE Robotics &*

528   *Automation Magazine* **13**, 99–110 (2006).

529 11. Milford, M. J., Wyeth, G. F. & Prasser, D. RatSLAM: a hippocampal model for simultaneous

530   localization and mapping. in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE*

531   *International Conference on* vol. 1 403–408 (IEEE, 2004).

532 12. Guo, W., Zhang, J. J., Newman, J. & Wilson, M. A. Latent learning drives sleep-dependent

533   plasticity in distinct CA1 subpopulations. *bioRxiv* (2020).

534 13. Ocko, S. A., Hardcastle, K., Giocomo, L. M. & Ganguli, S. Emergent elasticity in the neural code

535   for space. *Proceedings of the National Academy of Sciences* **115**, E11798–E11806 (2018).

536 14. Penny, W. D., Zeidman, P. & Burgess, N. Forward and backward inference in spatial cognition.

537   *PLoS computational biology* **9**, e1003383 (2013).

538 15. Evans, T. & Burgess, N. Coordinated hippocampal-entorhinal replay as structural inference. in

539   *Advances in Neural Information Processing Systems* (2019).

540 16. Diba, K. & Buzsáki, G. Forward and reverse hippocampal place-cell sequences during ripples.

541   *Nature Neuroscience* **10**, 1241–1242 (2007).

542 17. Foster, D. J. & Wilson, M. A. Reverse replay of behavioural sequences in hippocampal place

543   cells during the awake state. *Nature* **440**, 680–683 (2006).

544 18. Ólafsdóttir, H. F., Carpenter, F. & Barry, C. Coordinated grid and place cell replay during rest.

545   *Nature Neuroscience* **19**, 792–794 (2016).

546   19. O'Neill, J., Boccara, C., Stella, F., Schoenenberger, P. & Csicsvari, J. Superficial layers of the

547   medial entorhinal cortex replay independently of the hippocampus. *Science* **355**, 184–188 (2017).

548   20. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space.

549   *Science* **261**, 1055–1058 (1993).

550   21. Barry, C., Hayman, R., Burgess, N. & Jeffery, K. J. Experience-dependent rescaling of entorhinal

551   grids. *Nature Neuroscience* **10**, 682–684 (2007).

552   22. Barry, C., Ginzberg, L. L., O'Keefe, J. & Burgess, N. Grid cell firing patterns signal

553   environmental novelty by expansion. *Proceedings of the National Academy of Sciences* **109**,

554   17687–17692 (2012).

555   23. Krupic, J., Bauza, M., Burton, S. & O'Keefe, J. Local transformations of the hippocampal

556   cognitive map. *Science* **359**, 1143–1146 (2018).

557   24. Stensola, H. *et al.* The entorhinal grid map is discretized. *Nature* **492**, 72–78 (2012).

558   25. Butler, W. N., Hardcastle, K. & Giocomo, L. M. Remembered reward locations restructure

559   entorhinal spatial maps. *Science* **363**, 1447–1452 (2019).

560   26. Hägglund, M., Mørreaunet, M., Moser, M.-B. & Moser, E. I. Grid-Cell Distortion along

561   Geometric Borders. *Current Biology* (2019) doi:10.1016/j.cub.2019.01.074.

562   27. Hollup, S. A., Molden, S., Donnett, J. G., Moser, M.-B. & Moser, E. I. Accumulation of

563   hippocampal place fields at the goal location in an annular watermaze task. *Journal of*

564   *Neuroscience* **21**, 1635–1644 (2001).

565   28. Munn, R. G. K., Mallory, C. S., Hardcastle, K., Chetkovich, D. M. & Giocomo, L. M. Entorhinal

566   velocity signals reflect environmental geometry. *bioRxiv* (2019) doi:10.1101/671222.

567   29. Stensola, T., Stensola, H., Moser, M.-B. & Moser, E. I. Shearing-induced asymmetry in

568   entorhinal grid cells. *Nature* **518**, 207–212 (2015).

569   30. Hafting, T., Fyhn, M., Molden, S., Moser, M.-B. & Moser, E. I. Microstructure of a spatial map

570   in the entorhinal cortex. *Nature* **436**, 801–806 (2005).

571   31. Fiete, I. R., Burak, Y. & Brookings, T. What Grid Cells Convey about Rat Location. *Journal of*

572   *Neuroscience* **28**, 6858–6871 (2008).

573   32. Mathis, A., Herz, A. V. & Stemmler, M. Optimal population codes for space: grid cells

574   outperform place cells. *Neural computation* **24**, 2280–2317 (2012).

575   33. Stemmler, M., Mathis, A. & Herz, A. V. M. Connecting multiple spatial scales to decode the

576   population activity of grid cells. *Science Advances* **1**, e1500816–e1500816 (2015).

577   34. Guanella, A., Kiper, D. & Verschure, P. A model of grid cells based on a twisted torus topology.

578   *International journal of neural systems* **17**, 231–240 (2007).

579   35. Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. The hippocampus as a predictive map.

580   *Nature Neuroscience* (2017) doi:10.1038/nn.4650.

581   36. Whittington, J., Muller, T., Mark, S., Barry, C. & Behrens, T. Generalisation of structural

582   knowledge in the hippocampal-entorhinal system. in *Advances in Neural Information Processing*

583   *Systems 31* (eds. Bengio, S. et al.) 8484–8495 (Curran Associates, Inc., 2018).

584   37. Widloski, J. & Fiete, I. R. A Model of Grid Cell Development through Spatial Exploration and

585   Spike Time-Dependent Plasticity. *Neuron* **83**, 481–495 (2014).

586   38. Kropff, E., Carmichael, J. E., Moser, M.-B. & Moser, E. I. Speed cells in the medial entorhinal

587   cortex. *Nature* **523**, 419–424 (2015).

588   39. Munn, R. G. K., Mallory, C. S., Hardcastle, K., Chetkovich, D. M. & Giocomo, L. M. Entorhinal

589   velocity signals reflect environmental geometry. *Nat Neurosci* (2020) doi:10.1038/s41593-019-

590   0562-5.

591   40. Milford, M. J. & Wyeth, G. F. Mapping a suburb with a single camera using a biologically

592   inspired SLAM system. *IEEE Transactions on Robotics* **24**, 1038–1053 (2008).

593   41. Muller, R. U., Stead, M. & Pach, J. The hippocampus as a cognitive graph. *The Journal of*

594   *general physiology* **107**, 663–694 (1996).

595   42. Burgess, N. & O'Keefe, J. Neuronal computations underlying the firing of place cells and their

596   role in navigation. *Hippocampus* **6**, 749–762 (1996).

597   43. Campbell, M. G. *et al.* Principles governing the integration of landmark and self-motion cues in

598   entorhinal cortical codes for navigation. *Nature neuroscience* **21**, 1096 (2018).

599   44. Krupic, J., Bauza, M., Burton, S., Barry, C. & O'Keefe, J. Grid cell symmetry is shaped by

600   environmental geometry. *Nature* **518**, 232–235 (2015).

601   45. Hägglund, M. & Mørreaunet, M. Grid pattern development, distortions and topological defects

602   may depend on distributed anchoring. *BioRxiv* 830158 (2019).

603   46. Bhattacharyya, A. On a measure of divergence between two statistical populations defined by

604   their probability distributions. *Bull. Calcutta Math. Soc.* **35**, 99–109 (1943).

605   47. Tanni, S. & Barry, C. Relationship between place cell recruitment and the spatial scale of open-

606   field environments revealed using a novel scalable recording apparatus. in *Society for*

607   *Neuroscience* Program No. 689.18. (Society for Neuroscience, 2018).

608   48. Pearl, J. Reverend Bayes on inference engines: A distributed hierarchical approach. in *National*

609   *Conference on Artificial Intelligence* 133–-136 (AIII Press, 1982).

610   49. Elidan, G., McGraw, I. & Koller, D. Residual belief propagation: Informed scheduling for

611   asynchronous message passing. *arXiv preprint arXiv:1206.6837* (2012).

612   50. Davidson, T. J., Kloosterman, F. & Wilson, M. A. Hippocampal Replay of Extended Experience.

613   *Neuron* **63**, 497–507 (2009).

614   51. Stella, F., Baracskay, P., O'Neill, J. & Csicsvari, J. Hippocampal Reactivation of Random

615   Trajectories Resembling Brownian Diffusion. *Neuron* (2019) doi:10.1016/j.neuron.2019.01.052.

616   52. Bellmund, J. L. *et al.* Deforming the metric of cognitive maps distorts memory. *Nature human*

617   *behaviour* 1–12 (2019).

618   53. Boccara, C. N., Nardin, M., Stella, F., O'Neill, J. & Csicsvari, J. The entorhinal cognitive map is

619   attracted to goals. *Science* **363**, 1443–1447 (2019).

620   54. Nakazawa, K. *et al.* Requirement for hippocampal CA3 NMDA receptors in associative memory

621   recall. *Science* **297**, 211–218 (2002).

622   55. Dordek, Y., Soudry, D., Meir, R. & Derdikman, D. Extracting grid cell characteristics from place

623   cell inputs using non-negative principal component analysis. *eLife* **5**, (2016).

624   56. Banino, A. *et al.* Vector-based navigation using grid-like representations in artificial agents.

625   *Nature* **557**, 429 (2018).

626   57. Carpenter, F., Manson, D., Jeffery, K., Burgess, N. & Barry, C. Grid Cells Form a Global

627   Representation of Connected Environments. *Current Biology* 1–7 (2015)

628   doi:10.1016/j.cub.2015.02.037.

629 58. Lee, A. K. & Wilson, M. A. Memory of sequential experience in the hippocampus during slow
630   wave sleep. *Neuron* **36**, 1183–1194 (2002).

631 59. Skaggs, W. E. & McNaughton, B. L. Replay of neuronal firing sequences in rat hippocampus
632   during sleep following spatial experience. *Science* **271**, 1870–1873 (1996).

633 60. Wilson, M. A. & McNaughton, B. L. Reactivation of hippocampal ensemble memories during
634   sleep. *Science* **265**, 676–679 (1994).

635 61. Buzsáki, G., Vanderwolf, C. H. & others. Cellular bases of hippocampal EEG in the behaving rat.
636   *Brain Research Reviews* **6**, 139–171 (1983).

637 62. O'Keefe, J. Place units in the hippocampus of the freely moving rat. *Experimental neurology* **51**,
638   78–109 (1976).

639 63. Bicanski, A. & Burgess, N. A neural-level model of spatial memory and imagery. *eLife* **7**, e33752
640   (2018).

641 64. Mattar, M. G. & Daw, N. D. Prioritized memory access explains planning and hippocampal
642   replay. (2018) doi:10.1101/225664.

643 65. Momennejad, I., Otto, A. R., Daw, N. D. & Norman, K. A. Offline replay supports planning in
644   human reinforcement learning. *Elife* **7**, e32548 (2018).

645 66. Hasselmo, M. E. Neuronal rebound spiking, resonance frequency and theta cycle skipping may
646   contribute to grid cell firing in medial entorhinal cortex. *Philosophical Transactions of the Royal*
647   *Society of London B: Biological Sciences* **369**, 20120523 (2014).

648 67. Muller, L., Chavane, F., Reynolds, J. & Sejnowski, T. J. Cortical travelling waves: mechanisms
649   and computational principles. *Nature Reviews Neuroscience* (2018).

650 68. Yamamoto, J. & Tonegawa, S. Direct medial entorhinal cortex input to hippocampal CA1 is
651   crucial for extended quiet awake replay. *Neuron* **96**, 217–227 (2017).

652 69. Mathis, A., Stemmler, M. B. & Herz, A. V. Probable nature of higher-dimensional symmetries
653   underlying mammalian grid-cell activity patterns. *Elife* **4**, e05979 (2015).

654 70. Dayan, P. & Abbott, L. F. *Theoretical neuroscience*. (Cambridge, MA: MIT Press, 2001).

655 71. Barry, C. & Burgess, N. Learning in a geometric model of place cell firing. *Hippocampus* **17**,
656   786–800 (2007).

657

658 **Figure 1.** Online+offline localization and mapping. **A)** Illustration of recursive Bayesian

659 integration. A probability distribution over current location ($G(x)$), represented by grid cell

660 firing, is updated according to self-motion via the *transition model* ($\mathrm{T}(\mathbf{x}, \mathbf{x}'|\hat{\mathbf{u}}, \boldsymbol{\alpha})$) then refined

661 by environmental inputs via the *observation model* ($H(\boldsymbol{P}_t|\boldsymbol{x}_t)$). **C)** Estimates based on the

662 integration of noisy self-motion and environmental inputs may be stable, as shown in

663 simulated grid cell firing rate maps (*right*), despite instability when using only self-motion

664 (*left*) or environmental inputs after brief initial exploration (*middle*; numbers show gridness

665 score). **C)** Simulated grid cells exhibit spatially offset grid-like firing patterns, due to toroidal

666 connectivity, despite the absence of attractor dynamics. Right shows histogram of spatial

667 phases. **D)** Inferred pairwise distances D are a function of the 'overlap' between place fields.

668 **E)** Pairwise distances can be used to infer the structure of the world (the mapping of place

669 fields onto the grid map). **F)** Given noisy initial priors ("Initial"), structural encodings are

670 modified to reflect pairwise *associative* measurements. Inferred structure is sensitive to the

671 topology of the environment (cf. "Ring" and "Broken Ring"). **G)** Illustration of the prediction

672 error mechanism used to arbitrate between the *online* and *offline* systems (blue bars show

673 prediction error ε, ε$_0$ is the minimum prediction error needed to trigger *offline* inference).

674

675

676
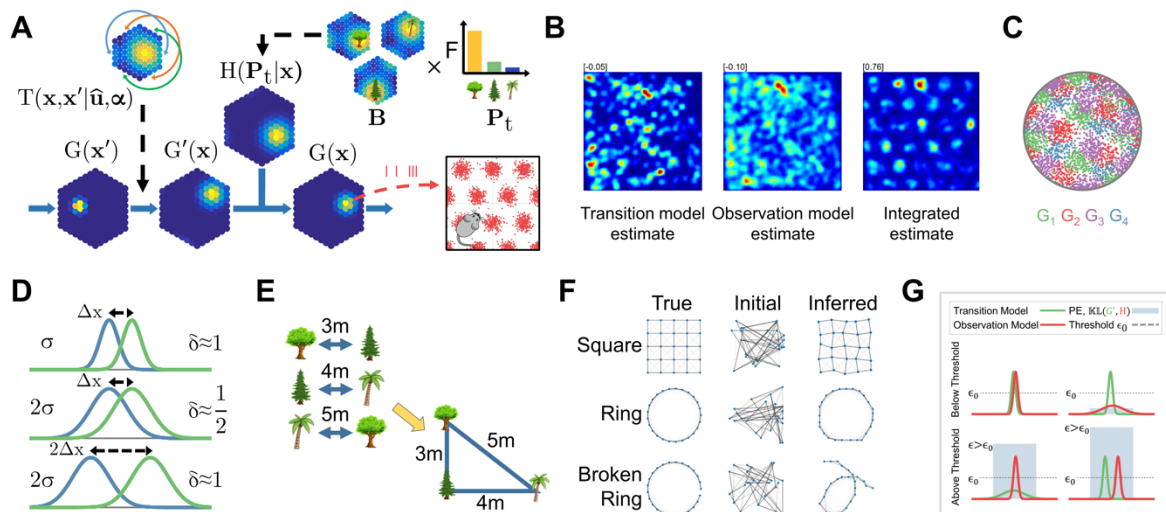
677 **Figure 2.** Minimizing prediction errors in the *offline* system: grid rescaling and direction-

678 dependent offsets under manipulations of environmental size or VR gain. **A)** During *online*

679 spatial localization, the *observation* model estimates location in metric space (activity on the

680 grid sheet) via inputs from place cells driven by environmental features (red curve), the

681 *transition* model updates the previous estimate (dashed blue curve) according to self-motion

682 (green curve), producing a combined estimate (blue curve). Manipulations of the

683 environment cause the predictions from both models to diverge. One way to minimize these

684 prediction errors is to modify the *observation* model by changing the connection weights

685 from place to grid cells (i.e. the mapping between environmental observations to estimated

686 location in grid space). An alternative is to modify the *transition* model to reflect the

687 *observation* model estimate (e.g. varying the 'gain' mapping self-motion to grid space). The

688 degree to which both are modified is controlled by the relative strength of their respective

689 prior confidences (T_C). **B)** Joint optimization of the *observation* and *transition models* predict

690 partial rescaling of grid patterns in response to increase in the VR gain (i.e. the rate of visual

691 movement in response to physical movement on the ball; below) or compression of a real

692 environment (above). When the system is confident in its self-motion ($T_C \to \infty$, green), the

693 *observation* model is modified to match the *transition* model (no change in grid scale plotted

694 in the real or visual VR environment). When the system is confident in its environmental

695 inputs ($T_C \to 0$, red), the *transition* model adapts and grid scale follows the environmental

696 change. Balanced model confidence produces intermediate rescaling (blue). **C)** Change in

697 observed grid scale (grid pattern plotted against self-motion) depends on the *transition*

698 confidence values T_C. X-axis shows VR gain change or environmental compression, where

699 $\alpha^{\text{Visual}} = 1$ is a control trial. Y-axis shows observed change in grid scale (see Methods). Data

700 points show corresponding values from Barry et al. (2007), Chen et al. (2019) and Campbell

701 et al., (2019), which suggest an approximately equal weighting of *transition* and *observation*

702 model priors for both gain decreases and increase trials ($T_C \approx 1$). **D)** Firing rate map of a grid

703 cell on the linear track in control (solid line) and VR gain decrease trials ($\alpha^{Visual} = 0.7$),

30

704   plotted in the visual VR environment. In the gain decrease condition, the grid fields are stably

705   shifted to fire earlier when running left (dashed line) or right (dotted line), i.e. towards the

706   location indicated by the transition model from that indicated by vision. **E)** The direction-

707   dependent shift in estimated location in the grid module in a VR gain decrease trial

708   ($\alpha^{Visual} = ⅔$) stabilises at a fixed distance. Sequence of eight updates of estimated location

709   on the grid module G(x) when running to the right in a VR visual gain decrease trial, colours

710   as in A. The *transition* model (green) predicts a location ahead of that from the *observation*

711   model on the grid sheet (red; driven by visual input) because of the visual gain decrease.

712   Combining these estimates produces an intermediate distribution (blue). At each new

713   update, the prediction from the *transition* model builds on the shift of the previous combined

714   estimate (not the previous transition model estimate) so that the distance between the

715   observation model estimate and the combined estimate stabilizes at a fixed value, producing

716   a fixed direction-dependent offset of the grid patterns in environmental coordinates (see the

717   differences in the location estimates between models, Inset).
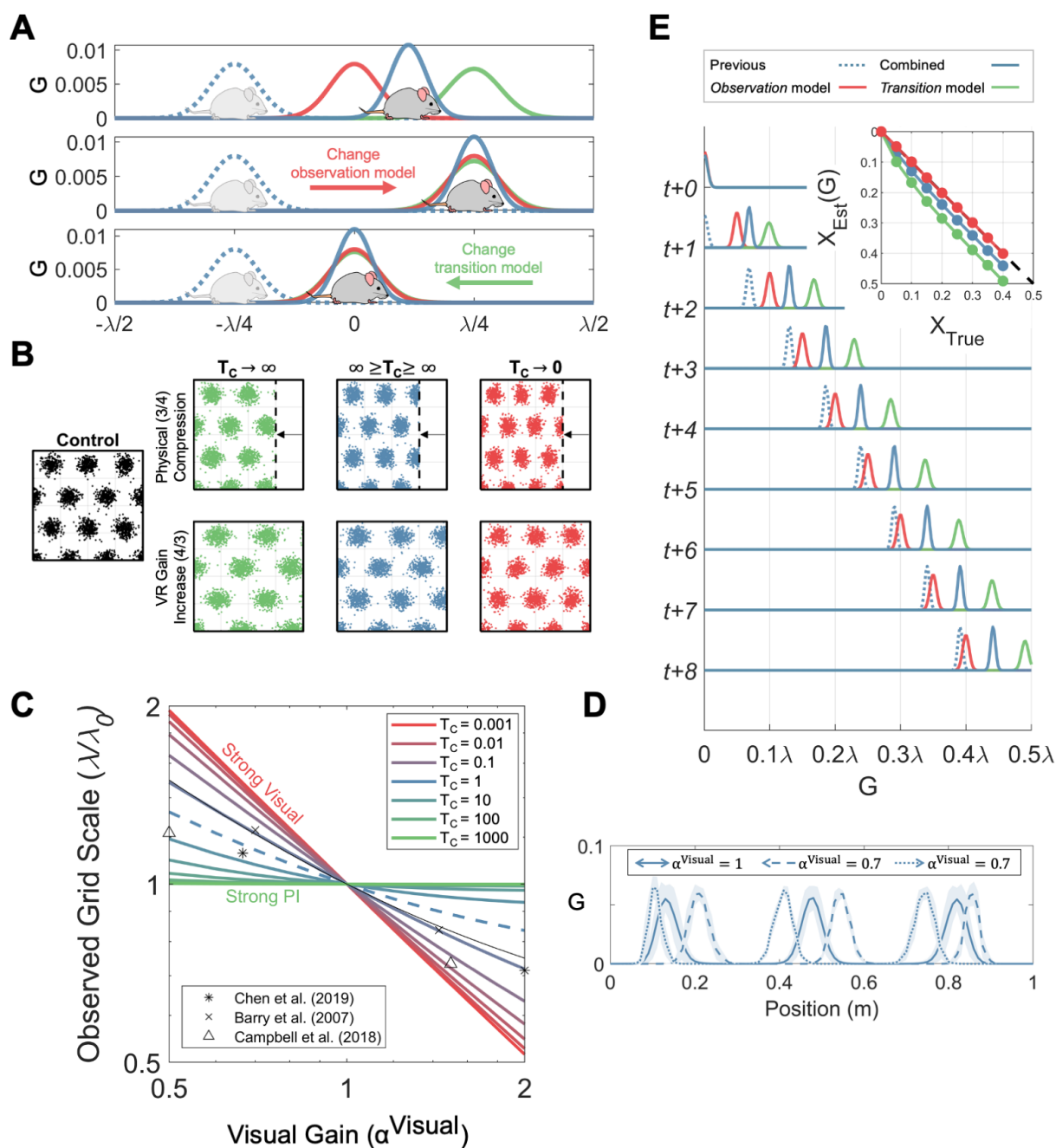
720 **Figure 3.** Effects of local deformation of a rectangular environment on grid patterns following

721 *offline* inference (Krupic et al., 2018). Dots indicate place/grid field locations, arrows indicate

722 shifts before/after *offline* structural optimization and coloured hexagons indicate magnitude

723 of shift vector. Grid patterns show estimate generated by the *observation* model (weighted

724 place cell activity). **A)** Place field shifts were measured from Krupic et al. (2018) and

725 interpolated and smoothed. **B-C)** Place field shifts cause immediate prediction errors (PE)

726 between the new pairwise place field distances and the distances between their

727 corresponding grid locations. If the *observation / transition* model priors are weak ($P_C \rightarrow 0$),

728 PEs are eliminated during *offline* inference by updating the *observation* model (top row;

729 unlike in Fig. 4B, modifying the global gain did not remove the effect of local distortions).

730 Complete adaptation in the *observation* model leads to an unperturbed grid pattern, i.e. grid

731 fields will not shift. Alternatively, strong model priors ($P_C \rightarrow \infty$) prevent adaptation to new

732 environmental inputs, leading to distorted grid patterns when driven purely by the

733 observation model, whose field shifts match those of the place cells (bottom row). When the

734 model priors are balanced against the new pairwise observations, the *observation* model is

735 partially adjusted, producing partial grid field shifts (which are smaller than those of the place

736 fields; middle row). Partial (middle) or no (bottom) adjustment to new observations preserve

737 mismatches between the *transition* and *observation* models, which would result in direction-

738 dependent offsets (see Figure 4E) and irregular firing patterns (see Figure S2D). NB the

739 confidence in the prior model may depend on location, e.g. if there is strong anchoring to the

740 wall prior connections from place cells with fields near the wall may be stronger than those

741 with fields further away. **D)** Comparison of experimentally observed place and grid field shifts

742 with simulation corresponding to the middle row of **B-C**.

743

744
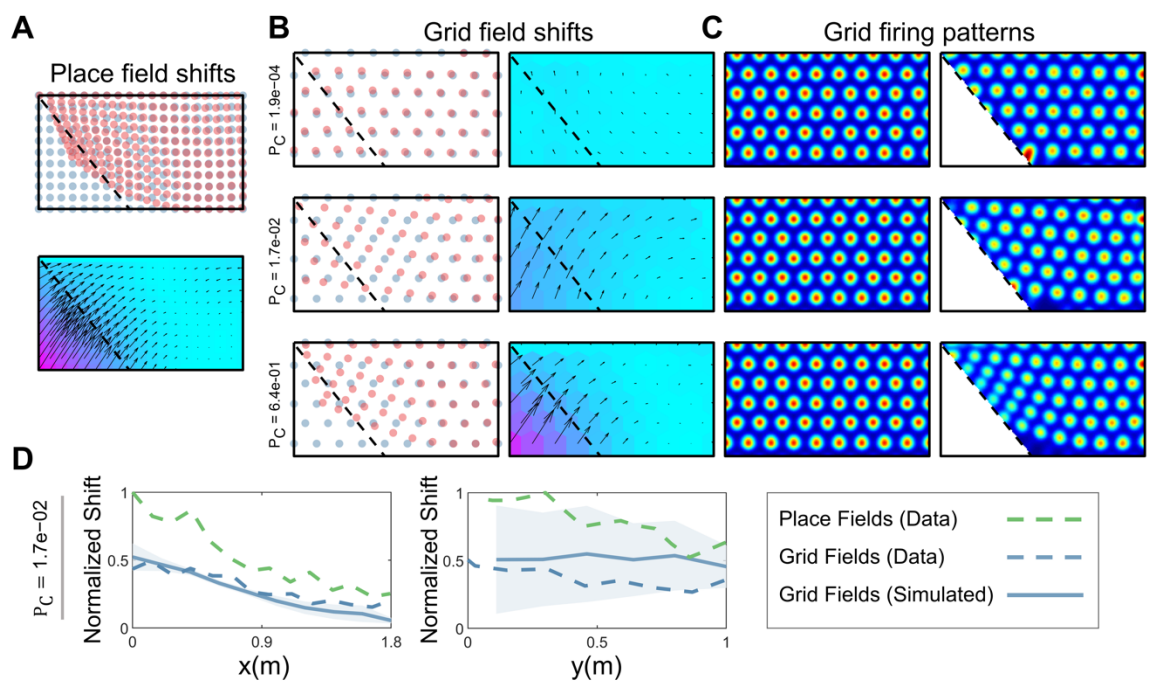
745 **Figure 4.** Distortions to grid patterns caused by inhomogeneous environmental input during

746 offline inference. **A)** Simulated variation in place field shape due to proximity to boundaries

747 (above), or inhomogeneous sampling of locations (Hagglund et al,. 2019; below) results in

748 distortions to the inferred pairwise distances in CA3 (**B**). These distortions lead to distortions

749 to grid scale due to adjustment of the observation model during offline inference (**C-D**).

750

751
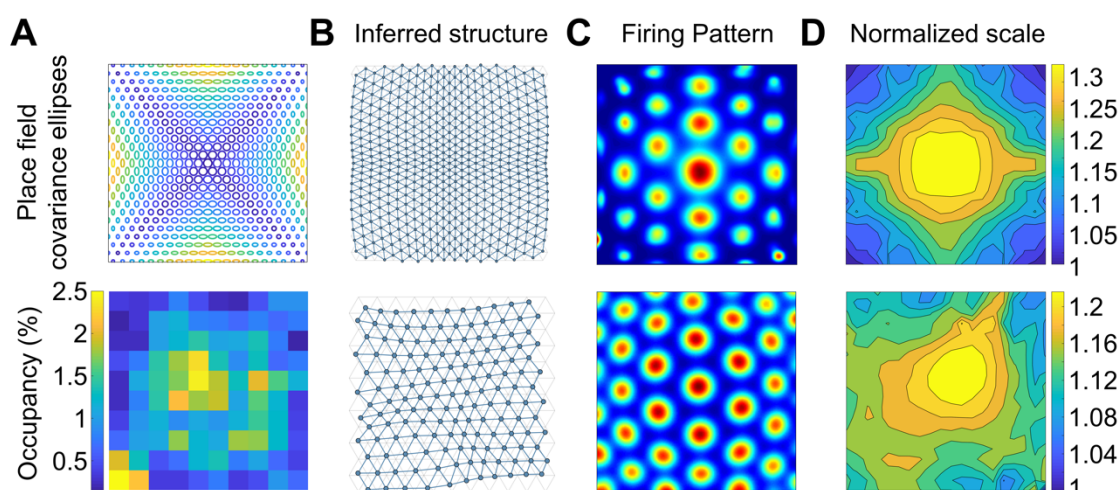
752 **Figure 5.** Illustration of the dual-systems model: prediction errors and replay in the loop-

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.

774

775
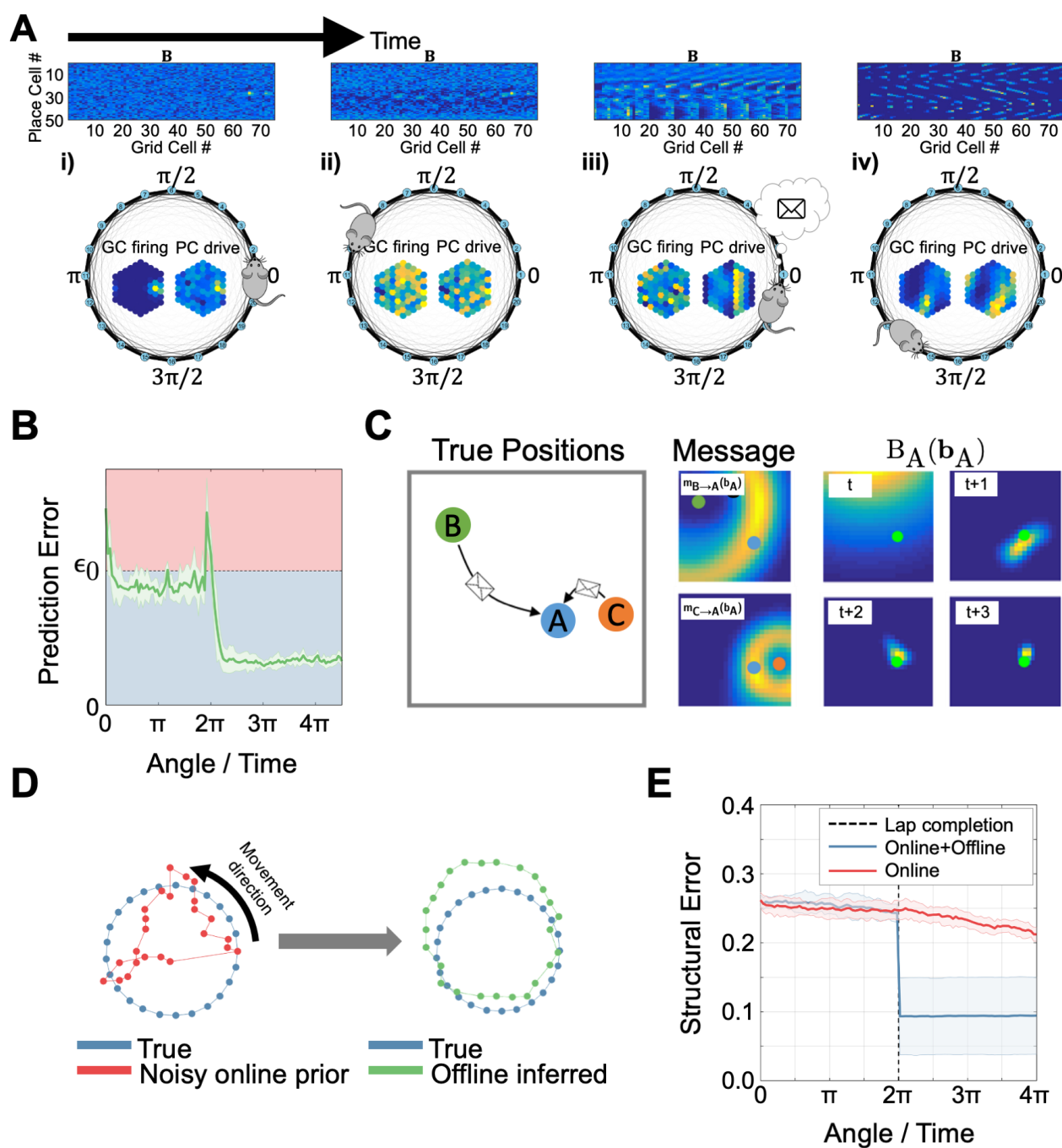
776   **Figure 6.** Principled message scheduling during *offline* inference generates sequences of

777   place cell activity. **A)** Place cells (i.e. graph nodes representing conjunctions of

778   environmental features) are connected via their pairwise potentials (ψ_{AB}, ψ_{BC}), which

779   penalize the mismatch between *associative* and *metric* pairwise distances (δ_{ij} and d_{ij}

780   respectively). ($t = 0$) Environmental sensory input ($m_{S \to A}(b_A)$ causes an update to the belief

781   of place cell A (i.e. updating its synaptic weights to the grid sheet) by making it fire in a new

782   location $b_A = 1$. ($t = 1$) Place cell A sends a message to B expressing its belief over the

783   location of B, given its own (new) location and the *associative* distance δ_{AB}, causing B to

784   update its belief. ($t = 2$) Messages from B to A and C only cause C to update its belief, so

785   only C broadcasts at the next time-step. **B)** Examples of PC reactivation sequences in loop-

786   closure task for different values of path integration noise σ_{PI}^2 (and therefore pairwise

787   measurement confidence, since $w_{ij} = 1/(\sigma_{PC}^2 + \sigma_{PI}^2 \delta_{ij})$; where $\sigma_{PC}^2 = 1e-4$). Multiple local

788   sequences can occur in interleaved fashion (Middle, grey shading) and become longer and

789   smoother when pairwise measurements are less confident (Right; also **E**, **F**). **C)** The 'Max-

790   Entropy' schedule (i.e., only the place cell with max entropy change fires in the next-step)

791   converges faster than when all PCs broadcast messages at each time-step. **D)** Forward and

792   reverse sequences occurred approximately equally often.

793

794
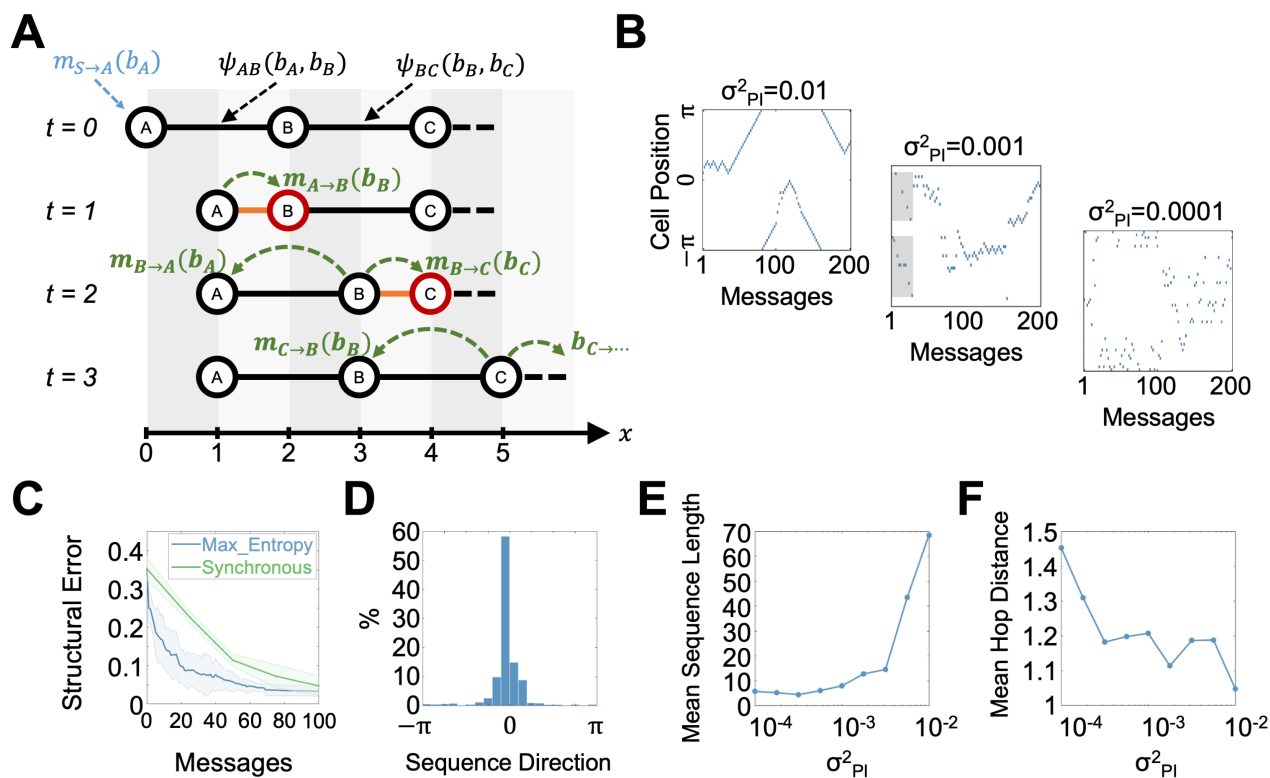
795   **Figure 7.** Schematic neural mechanism for probabilistic message passing (see also

796   Supplementary Video 2). **A)** Illustration in 1D. The broadcasting place cell P_B sends a spike

797   to receiving place cell P_R via recurrent connections in CA3, and also initiates a travelling

798   wave at the corresponding location on the grid cell module via their connections there. (Left)

799   No new learning occurs when the spike and travelling wave arrive at P_R and its

800   corresponding grid location (GC_7) at the same time, as P_R-GC_7 connection will already be

801   strong. (Middle) If the CA3 spike arrives at P_R ahead of the travelling wave reaching GC_7, the

802   synaptic associations of P_R are adjusted towards the currently active GC_5 (updating the

803   belief, by increasing P_R-GC_5 and decreasing P_R-GC_7, see *dW*, below). (Right) Same as

804   middle, except that GC wave reaches GC_7 before the CA3 spike reaches P_R. **B)** Propagating

805   messages as travelling waves in mEC. A neural simulation of travelling waves with a

806   modified Laplacian diffusion kernel (Wave) closely approximates the probabilistic

807   propagation of activity (True), reflecting the accumulation of self-motion noise in the

808   broadening of the wave front. **C)** Comparison of the algorithmic and neural belief

809   propagation performance.

810

811

812
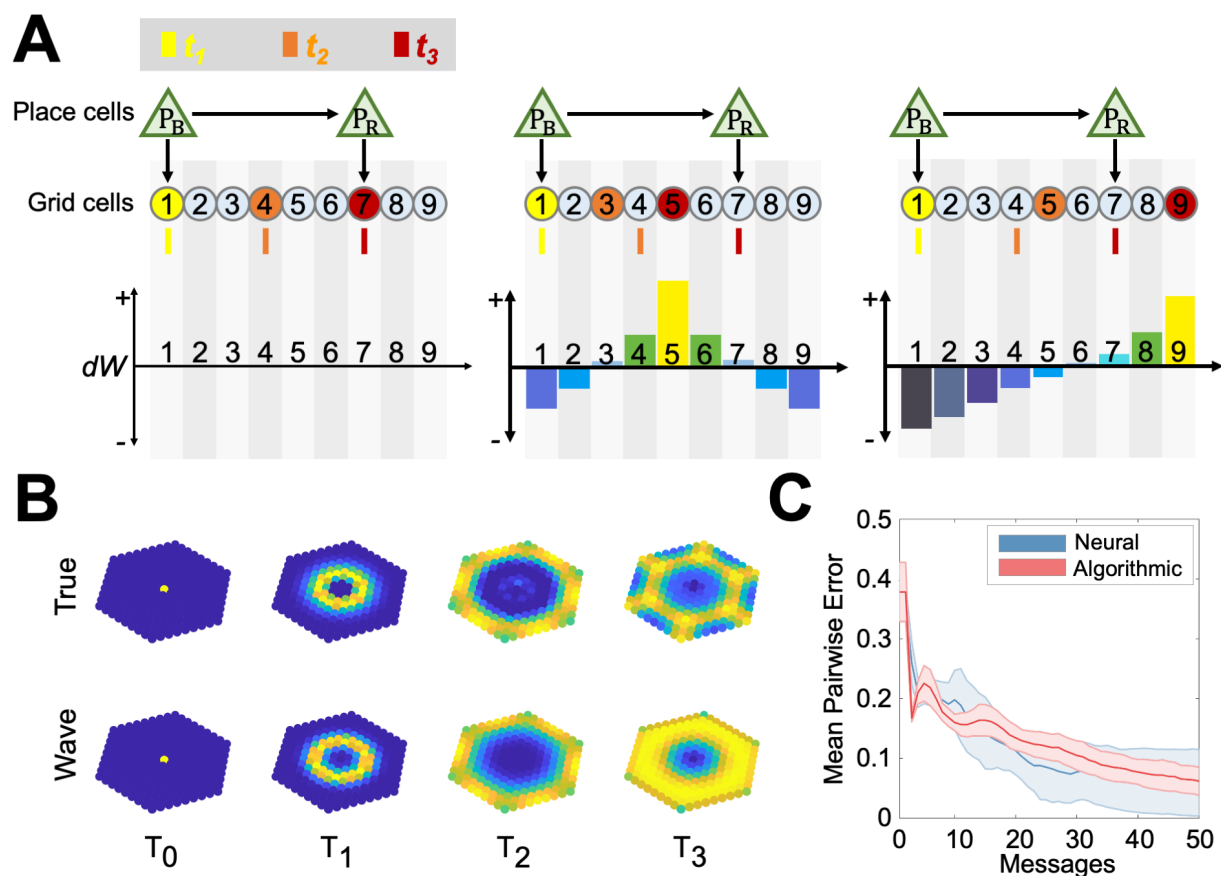
813   **Figure S1** Implementational details of the *online* system. **A** The mapping from real to grid

814   cell space can be considered as the subtraction of a mapping vector $[i \cdot \boldsymbol{v}_1 + j \cdot \boldsymbol{v}_2]$, where

815   each grid field can be described as a point on a 2D lattice with basis vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Each

816   grid field has associated with it a Voronoi domain, defined as the region within which any

817   point is closest to the corresponding grid field. When mapping from grid-to-real space, the

818   vector of the closest grid field is subtracted. **B** The wrapped Normal distribution is a

819   summation of the likelihoods of the current position estimate being at any one of an infinite

820   number of periodic tilings (here, five wrappings are shown). **C** The grid cells are connected

821   periodically to produce a 'twisted-torus' topology. **D** Illustration of the shifter cell mechanism.

822   Each 'readout' grid cell is connected to four 'shifter' grid cells and a single self-connecting

823   cell. **E** Illustration of the difference between wrapped and non-wrapped Gaussian

824   distributions. **F** Correspondence between the neural shifter cell mechanism and the

825   algorithmic transition function. **G** Pairwise distances between place fields can be inferred

826   from the strengths of Hebbian connections. **H** Mean entropy in the beliefs of each place cell

827   over their encoded location in grid space (encoded in place-grid cell connection weight

828   distributions decreases with iterations during *offline* inference.

829

830

831

832
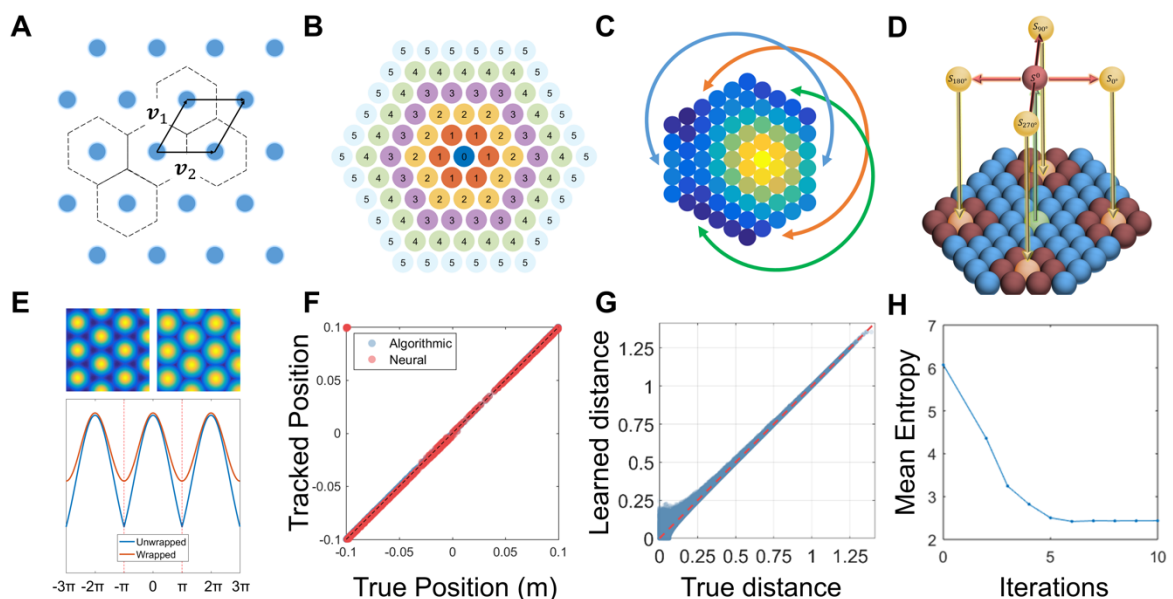
833 **Figure S2 A** Convergence of grid patterns and place-grid cell weights over three periods

834 (columns). The place-grid cell weights (*bottom row*, colour denotes normalized connection

835 strength) converge fastest near to the boundaries (as seen in development; Muessig et al.,

836 2015) and corners of an environment, since trajectories through are more directionally

837 constrained. Top row shows corresponding grid patterns (colour denotes normalized firing

838 rate). **B** Convergence of stable grid patterns over sensory noise (σ_{PC}) and path integration

839 noise (σ_{PI}). Colour denotes grid score (Sargolini et al., 2006). *P* indicates number of place

840 cells used in simulations. **C-D** Environmental deformations caused by prior structural beliefs.

841 **C** Place fields generated by the boundary vector cell model (Hartley et al., 2000). An

842 animal's perception of a trapezoidal environment (*right*) may be influenced by place-BVC

843 associations learned in a previous rectangular environment (*left*). In the trapezoid, simulated

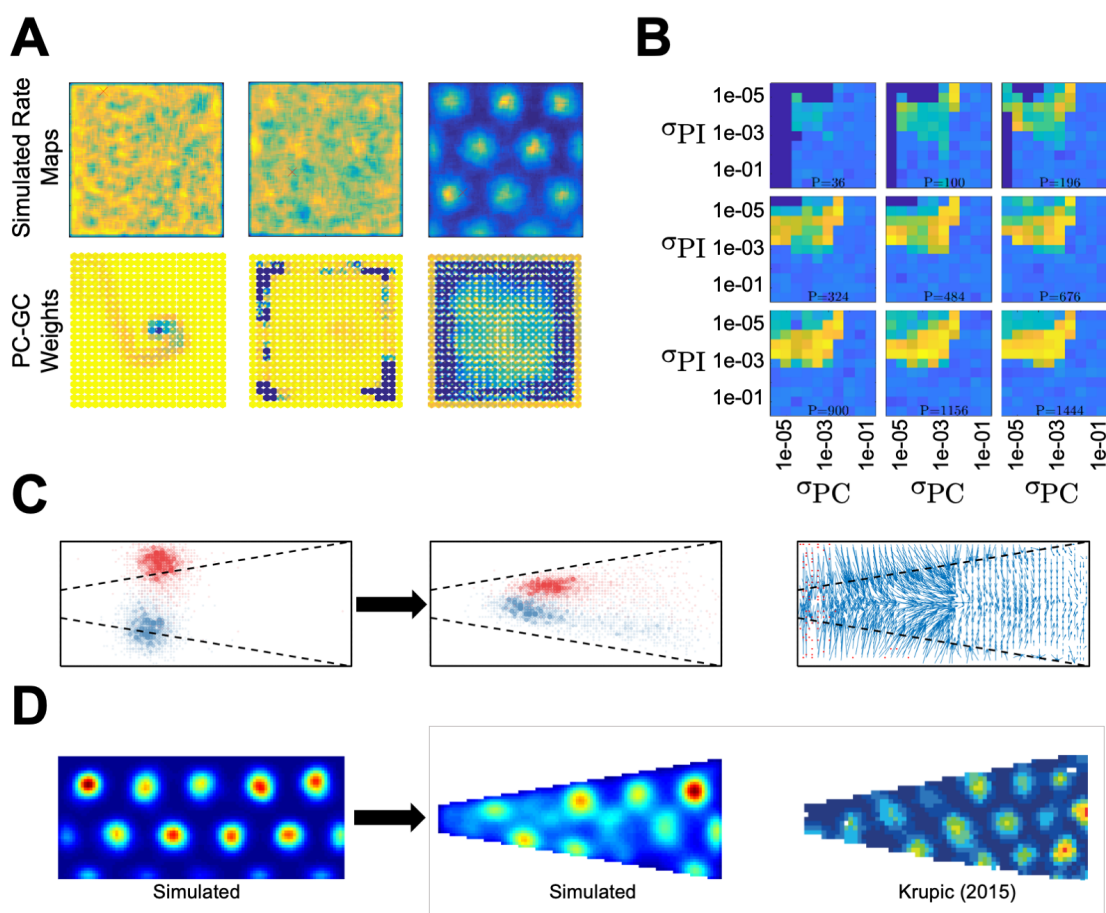844 place fields shift with the wall (**D**), causing similar distortion of the grid pattern.

845

44

846 **Figure S3.** Grid distortions due to inhomogeneous environmental inputs in the *online* model.

847 **A)** When environmental inputs are concentrated in one corner of the environment (*top-*

848 *middle panel*), the resultant grid cell firing rate maps undergo a shearing distortion which

849 produces an orientation offset (A_{min}; *bottom-left panels*). This offset increases with

850 experience, as grid cell firing becomes increasingly dependent on the maturing sensory

851 inputs (*bottom-middle panel*), matching experiments in which rats always entered the box at

852 the same corner (*top-left panels*; Stensola et al., 2015) The size of the experimental and

853 simulated offsets are similar (*right panels*). **B)** Simulated distortions based on an exponential

854 decay in place cell input from one corner as a function of the decay parameter (R). **C)**

855 Concentrated place field input along one wall and both corners (*bottom left panel*) causes

856 another distortion pattern (*bottom right panel*) also observed experimentally (*top* adapted

857 from Stensola et al., 2015).
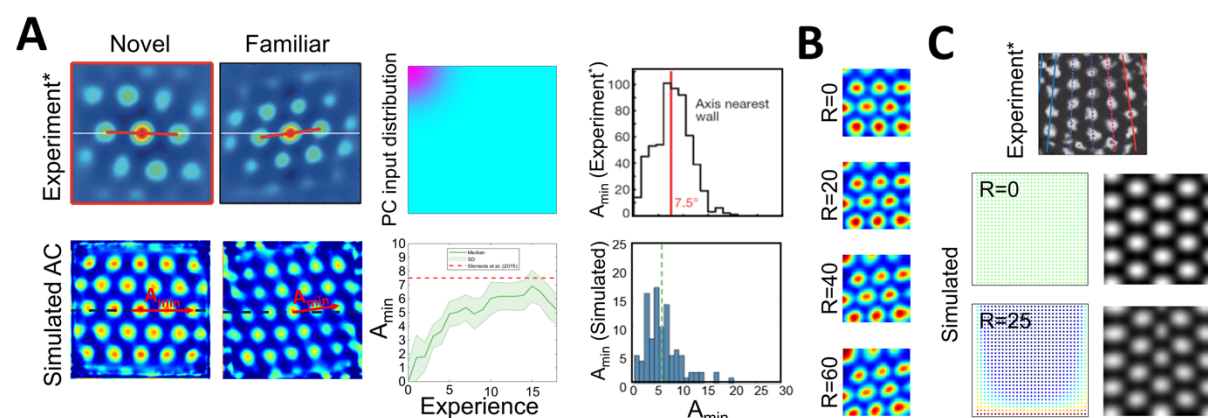

858

859

860   **Figure S4** Mathematical analysis of shearing of the grid pattern due to inhomogeneous

861   environmental inputs. **A** The orientation resulting from the shearing operation can be

862   calculated by analysing the angles of the sheared hexagonal lattice describing the centroids

863   of the grid fields. **B, C** The analysis predicts that the orientation offset should be dependent

864   on both the strength of the place fields' density / firing rate imbalance R (see Supplementary

865   Methods 1.7 and Fig. 2) (**B**) and their tuning widths (**C**).
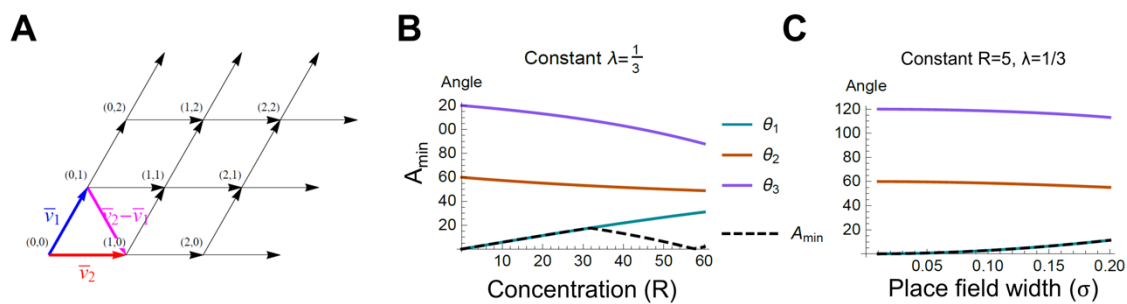

866

867

868   **Figure S5** Anatomy of a SLAM system. The joint location-map probability distribution (**A**) is

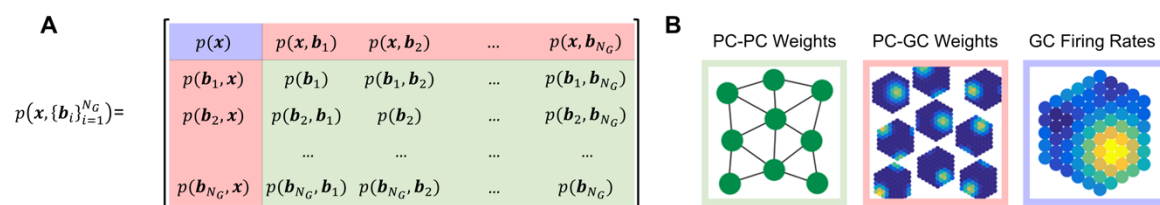869   represented in the firing rates and synaptic weights within the HPB-mEC system (**B**).

870

871

872

873 **Figure S6 A** Spring network analogy of the associative structure of an environment. Edge

874 'stiffness' is inversely proportional to the variance in the Gaussian observation. **B** 'Wrapping'

875 physical space to encoded location on the grid sheet. Each colour indicates the tiling of the

876 based grid sheet's domain in real space. **C** Pairwise distances near the edges of the

877 environment are overestimated due to under-sampling when the agent preferentially

878 explores the middle of an environment. Colours denote the distance of the pair of PCs i and

879 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 - $

880 $y_p))$. **D** Resulting local scale is proportional to the occupancy. C When the grid scale is

881 smaller than the size of structure being encoded, we can think of 'wrapping' the structure

882 onto the grid sheet. Here, colours denoted different tilings of the base metric tile (the Voronoi
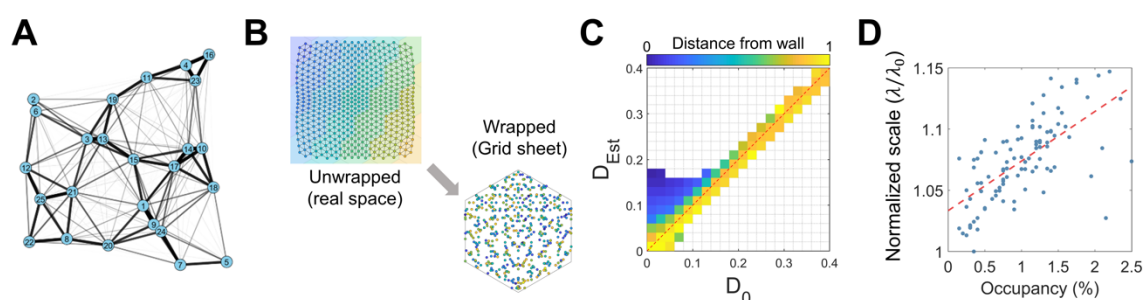
883 region of a given grid cell).

884

885

886

887

888   **Figure S7** An alternative view of the *online* and *offline* models. Place cells P are driven by

889   real-world stimulus S. During *online* exploration, the *associative* generative model is learned,

890   but does not generate predictions. During *offline* inference, the *metric* generative model is

891   corrected towards the predictions being generated by the *online* model, which becomes a

892   surrogate for sensory stimuli as would be generated by the real world model.

893

894

895

896

897

898

899

900

901

902

903

904

49
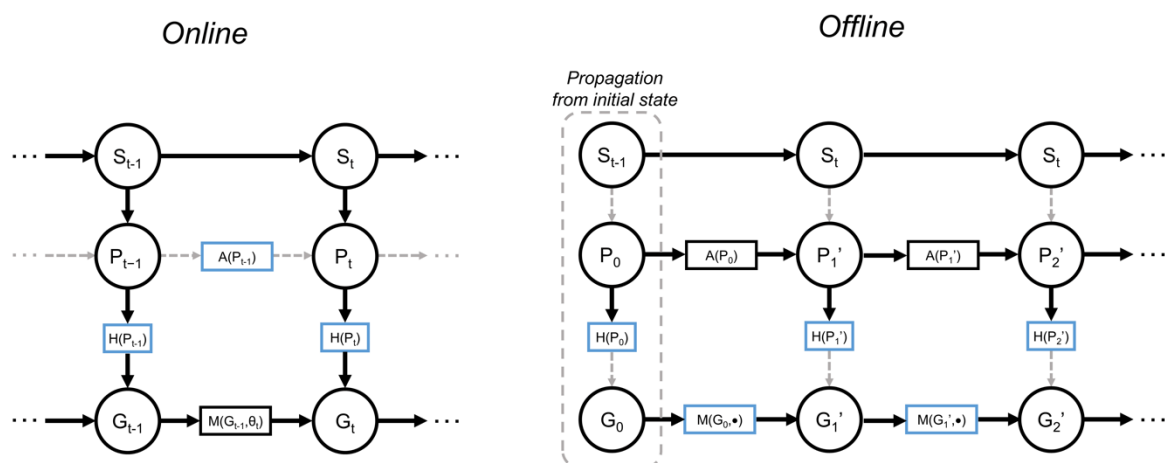
905

906

907

908

909 **Table S1** Algorithm detailing the overall dual-systems hypothesis of *online* and *offline*

910 localization and learning in the HPC/mEC.

Algorithm: *Online* localization and mapping with prediction error initiated *offline* inference	
1: initialize: $\mathbf{A} \sim \mathbb{U}(0, \Delta)$, $\mathbf{B} \sim \mathbb{U}(0, \Delta)$	% Initialize weights to small values $\sim \Delta$
2: while $k < K$ do:	% For the duration of the simulation
3: do_movement_update();	% Update state via path integration
4: make_observation();	% Compute PC firing
5: if $\epsilon_k > \epsilon_0$ do:	% If prediction error, do *offline* inference
6: while any $\mathcal{T}_i < \mathcal{T}_{min}$ do:	% Loop until cells below tension threshold
7: u=argmin($\boldsymbol{\mathcal{T}}$);	% Find node with max. message tension
8: compute_and_broadcast_message(u);	% Broadcast message to neighbours of u
9: for $t \in$ Neighbours(u) do:	% Loop over neighbours of u
10: update_belief(t);	% Update belief of node t
11: update_message_tensions(t);	% Compute change in belief of node t
12: update_PC_GC_weights();	% Do *associative* to *metric* map learning
13: update_PC_PC_weights();	% Update *associative* weights
14: do_measurement_update();	% Correct state estimate

911

912

913