Merging information in the entorhinal cortex: what can we learn from robotics experiments and modeling?

Philippe Gaussier, Jean Paul Banquet, Nicolas Cuperlier, Mathias Quoy, Lise
Aubin, Pierre-Yves Jacob, Francesca Sargolini, Etienne Save, Jeffrey
Krichmar, Bruno Poucet

REVIEW

Merging information in the entorhinal cortex: what can we learn from robotics experiments and modeling?

Philippe Gaussier[1],[*], Jean Paul Banquet[1], Nicolas Cuperlier[1], Mathias Quoy[1], Lise Aubin[1],[5], Pierre-Yves Jacob[2], Francesca Sargolini[2], Etienne Save[2], Jeffrey L. Krichmar[3],[4] and Bruno Poucet[2]

ABSTRACT

Place recognition is a complex process involving idiothetic and allothetic information. In mammals, evidence suggests that visual information stemming from the temporal and parietal cortical areas ('what' and 'where' information) is merged at the level of the entorhinal cortex (EC) to build a compact code of a place. Local views extracted from specific feature points can provide information important for view cells (in primates) and place cells (in rodents) even when the environment changes dramatically. Robotics experiments using conjunctive cells merging 'what' and 'where' information related to different local views show their important role for obtaining place cells with strong generalization capabilities. This convergence of information may also explain the formation of grid cells in the medial EC if we suppose that: (1) path integration information is computed outside the EC, (2) this information is compressed at the level of the EC owing to projection (which follows a modulo principle) of cortical activities associated with discretized vector fields representing angles and/or path integration, and (3) conjunctive cells merge the projections of different modalities to build grid cell activities. Applying modulo projection to visual information allows an interesting compression of information and could explain more recent results on grid cells related to visual exploration. In conclusion, the EC could be dedicated to the build-up of a robust yet compact code of cortical activity whereas the hippocampus proper recognizes these complex codes and learns to predict the transition from one state to another.

KEY WORDS: Hippocampus, Place cells, Grid cells, View cells, Navigation, Path integration, Modulo projection, Robot

Introduction

Navigation is a critical task for most species, and it requires complex representations (target objects, landmarks, places, etc.) and strategies. Many species use path integration (PI) or dead reckoning (a navigation strategy relying on idiothetic information useful in the dark or in a poor visual environment) in conjunction with allothetic information-based navigation (vision, odors, sounds, touch). Starting from a robotics and system modeling point of view, we investigate how mammals can use visual and proprioceptive information to perform homing behaviors or to reach specific goal places.

Many studies related to mammals focus on the hippocampal system (HS), which is composed of the hippocampus and the entorhinal cortex (EC), two brain structures highlighted by the discovery of place cells (O'Keefe and Nadel, 1978) and grid cells, respectively (Hafting et al., 2005), as well as by their involvement in episodic memory (Eichenbaum et al., 1992). Indeed, hippocampal destruction results in a severe anterograde amnesia in humans (Scoville and Milner, 1957). The HS is connected to the associative cortical areas and performs multimodal fusion, as well as fast detection and recall of complex events (McClelland et al., 1995). The HS can be also be thought of as a generic indexing tool for episodic memories (Teyler and DiScenna, 1986; McClelland et al., 1995; Cohen and Eichenbaum, 1993). The main inputs to the hippocampus come from the EC, which contains grid cells. Grid cells have been recorded in the dorsal medial EC (dMEC) in rodents (Hafting et al., 2005), bats (Yartsev et al., 2011; Geva-Sagiv et al., 2015) and even in humans (Doeller et al., 2010; Killian et al., 2012; Jacobs et al., 2013). In contrast, the lateral EC (LEC) shows little spatial selectivity (Hargreaves et al., 2005) but is hypothesized to perform joint processing of spatial and non-spatial information (Yoganarasimha et al., 2011; Deshmukh and Knierim, 2011; Van Cauter et al., 2012; Tsao et al., 2013; Knierim et al., 2014; Save and Sargolini, 2017). The linear combination of dMEC grid cells with different spacings and orientations provides an efficient way to build place cell activities (Rolls et al., 2006; Solstad et al., 2006). Yet, how grid cell activity is generated (and what grid cells code for) is intensely debated (Giocomo et al., 2011; Zilli, 2012; Moser et al., 2014). The hexagonal firing patterns of grid cells can be modeled either by oscillatory interference models (Burgess et al., 2007) or by continuous attractor models (Wan et al., 1994; Touretzky and Redish, 1996; Redish and Touretzky, 1997; Samsonovich and McNaughton, 1997). These models have been instrumental for designing neurobiological experiments to either validate their theoretical predictions or modify their assumptions (Giocomo et al., 2011; Zilli, 2012). Prior robotics experiments and simulations showed that: (1) visual place cells can be easily obtained from the merging of 'what' and 'where' information [we refer to the 'what' and 'where' visual pathways (Ungerleider and Haxby, 1994) for primates, but it looks as if rodents could also perform some detailed visual tasks (Lashley, 1938; Kolb and Tees, 1990)] and (2) homing behavior can be achieved through the competition between a small number of place–action associations (Gaussier and Zrehen, 1995; Gaussier et al., 1997, 2000) in a way similar to what has been shown for insects (Wittmann and Schwegler, 1995; Collett and Baron, 1995; Etienne, 1998; Schwarz et al., 2017). In addition, our model of PI shows that it can be performed based on a one-dimensional (1D) field of neurons in a way also quite similar to insects. In mammals, this computation might be performed in different brain

¹ETIS - UMR 8051, Université Paris-Seine, Université de Cergy-Pontoise, ENSEA, CNRS, Cergy-Pontoise 95302, France. ²Laboratory of Cognitive Neuroscience (LNC - UMR 7291), Aix-Marseille Université, Centre National de la Recherche Scientifique, Marseille 13331, France. ³Department of Cognitive Sciences, University of California, Irvine, Irvine, CA 92697, USA. ⁴Department of Computer Science, University of California, Irvine, Irvine, CA 92697, USA. ⁵Euromov, Université de Montpellier, Montpellier 34090, France.

*Author for correspondence (gaussier@ensea.fr)

P.G., 0000-0003-2028-2044; F.S., 0000-0002-5038-3746; J.L.K., 0000-0003-0739-2468; B.P., 0000-0003-2704-804X

List of abbreviations

1D	one-dimensional
dMEC	dorsal medial EC
EC	entorhinal cortex
HD	head direction (cells)
HS	hippocampal system
LEC	lateral EC
MEC	medial EC
mod	modulo operation (i mod M is the remainder of the Euclidian division of i by M)
PI	path integration
proj	projection according to a modulo value
RSC	retrosplenial cortex

structures, including the retrosplenial cortex (RSC) or the parietal cortex (Cooper and Mizumori, 2001; Save et al., 2001; Parron and Save, 2004; Vann et al., 2009; Elduayen and Save, 2014; P.G and J.L.K., unpublished data).

In this paper, we first summarize results showing how place cells can be built from visual information or PI. Second, we show how grid cell activity can be explained as a general 'modulo projection' property of the cortical activity (Gaussier et al., 2007) if we suppose the existence of PI traces in the RSC. Third, we show how to obtain visual grid cells again using a modulo operator on the 'where' component of visual information. We conclude that spatial specificity of the rodent HS reflects a more general role in memory formation through its capability to detect novelty and to learn transitions between complex multimodal states that cannot not be easily detected at the cortical level. Hence, we question whether the hippocampus is the brain area where PI is computed (so as to build a Cartesian map of the environment as proposed by O'Keefe and Nadel, 1978) or, rather, where the results of different spatial computations are projected.

Significance statement

In this paper, we show how a simple robotics model of mammal navigation is useful to interpret neurobiological recordings. We question the current models of the dMEC as a path integrator. Instead, we propose that the EC is a generic merging tool that builds a compact representation of the cortical activity (a kind of hash table). We summarize experiments and simulations showing that grid cells related to PI could be explained as a modulo projection of cortical activity computed in the RSC, where PI could take place. Furthermore, we suggest that the visual grid cells recorded in the human EC could also be explained by the same mechanism.

Place cells from visual information

A place can be characterized by the identity of the landmarks ('what'), and their azimuth, distance or elevation ('where'). It can be a direct visual snapshot from a given position (as proposed for insects). Yet, if the image is decomposed into local views related to 'important' visual features or landmarks, then conjunctive cells (i.e. cells combining some different input types on common postsynaptic targets to produce 'conjunctive' activities) can be used to build a place code merging 'what' and 'where' information with better generalization capabilities over long distances (Insausti et al., 1987; Zola-Morgan et al., 1989; Gaussier and Zrehen, 1995; Gaussier et al., 1997). This increases the robustness of the visual system to occlusions, and the displacement of objects or displacement of the animal by making it possible to recognize unaltered local areas (landmarks) and to measure their apparent displacement in the

visual scene for place recognition and navigation. The merging of 'what' and 'where' information stemming from the perirhinal and parahippocampal cortex, respectively, is a good candidate for such a conjunctive function. If we assume the animal has access to an allocentric reference frame corresponding to a kind of internal compass (built from vestibular, tactile or visual information, for instance), it may easily obtain information about the angular displacement of a given landmark while moving in the environment. The multimodal head direction (HD) cells can represent such a visual compass (Lepretre et al., 2000; Delarboulas et al., 2014).

Fig. 1 shows a simplified view of the neural network, which was tested on a robot that learned visual places. The robot performed a serial exploration of the visual scene focusing on the local maxima of off-center cells (difference of Gaussians) applied on the gradient of the input image. The robot's visual scan path was controlled by the intensity of the feature points (a winner-takes-all mechanism with an inhibition of return allows the selection of feature points such as corners, end of line, etc., of decreasing activity). This simulated ocular saccades and an attentional spotlight mechanism. One local view is extracted around each saccade after a log/polar transformation of the image, mimicking the projection from the retinal to the primary visual areas (Schwartz, 1980). This transformation provides a robust signature to small-scale variation and rotation. We proposed that the azimuth and elevation of the focus point are coded as a spatial activity bump on two different 1D maps of neurons representing the associated angles. Next, conjunctive cells could be triggered when a local view was recognized under a specific azimuth or elevation (Fig. 1). As a result, if the landmark was perceived from a different position, its azimuth would change. The activity bump on the 'where' field would move and would provide lower activity on the associated connection. This mechanism provided a direct way to measure the angular variation of a given landmark between its learned angular position and its current position in the visual field. When the local views or landmarks were explored sequentially (focus of attention or ocular saccades), adding a short-term memory on the conjunctive activities resulted in an activity pattern insensitive to the exploration order of the panorama. At the end of exploration, the short-term memory had a copy of all the activated conjunctive cells building a spatial code (a constellation of landmark×azimuth×elevation) easy to manage for pattern matching (i.e. place recognition).

To test the performance of this model, the robot was put in different locations in a room (7×5 m) to learn 5×5=25 places regularly spaced every 1 m. For each location, the learning of one place cell was triggered. Next, the robot was moved everywhere in the room and the activity of each cell was recorded.

As shown in Fig. 2A, we obtained place fields that generalize over long distances in the room. Each rectangle represents the recording of one cell according to the location of the robot in the room. We can see, for instance, that the neuron that has learned the upper left corner of the room (upper-left rectangle) responds more in that part of the environment. Similarly, the neuron associated with the position in the middle of the room responds more in that location. The neuronal activity is related to the different places. It decreases slowly and spreads over more than 2 m in the experimental room. This interesting result means that we can create large attraction basins around a given location that allow very good generalization capabilities to unvisited places, and the capability to learn a path from place–action association and competition (Giovannangeli et al., 2006). Yet, it is also a puzzling finding, because place fields recorded in the dorsal hippocampus (in the CA3 or CA1 regions) are clearly sharper than the model's place fields. The diameter of the
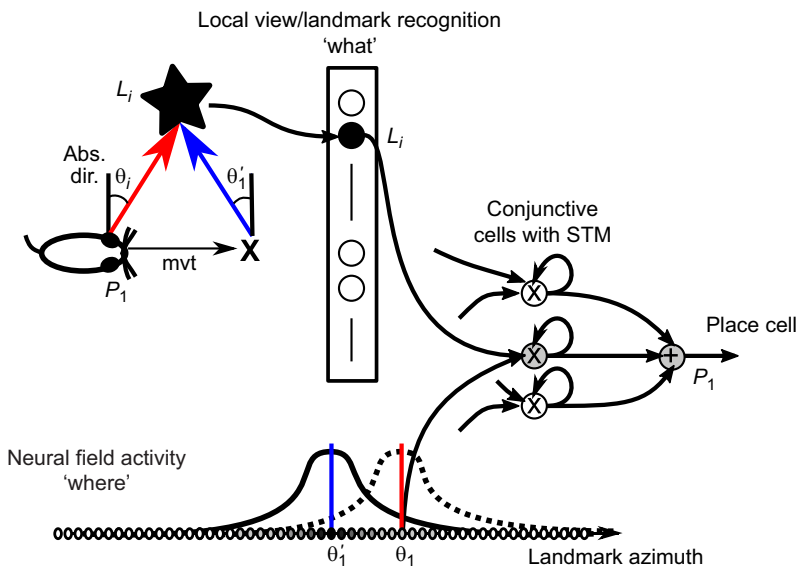
Fig. 1. Visual place cell from the merging of 'what' and 'where' information. Landmark angular variation induces a decrease in the landmark recognition. L_i, neuron associated with the recognition of landmark i; P_1, classical neuron performing the visual place recognition from the conjunctive (product) units; θ_i, azimuth of landmark i (here, the black star) according to an absolute referential; θ_i', landmark azimuth after a movement shown by the vector mvt; STM, short-term memory.

real fields in rodents is at least 10 times smaller (10 to 20 cm) (Mizuseki et al., 2012). This discrepancy may be resolved by adding a competition between place cells, thus leading to sharper place fields as shown in Fig. 2B) or by supposing that our cells correspond more to the place cells recorded in the ventral hippocampus (Jung et al., 1994; Poucet et al., 1994). A second question is whether the cells we recorded in the robot would be classified as place cells by neurobiologists. The answer is clearly no, because the cells are somewhat active everywhere in the environment. The neurobiologists would term these 'diffuse place cells' (Quirk et al., 1992; Song et al., 2012), much like some of the unclassified medial EC (MEC) cells [e.g. see Savelli et al. (2008) and also the discussion in Poucet et al. (2014) on ventral MEC 'place-like cells'].

Is it possible that place cells useful for navigation in an open environment (with long-distance generalization capabilities) exist outside the hippocampus proper? If we suppose that entorhinal cells merge multimodal information, they could be a good candidate to explain why rats with a lesioned hippocampus can still perform some navigation tasks. Even though such large visual place cells could be located in the LEC, results from Deshmukh and Knierim (2011) and Tsao et al. (2013) do not support their existence. Yet, many cells react to objects and their locations, and their weak spatial correlates could correspond to our large place cells. Place-related activity in the LEC in the presence of objects or goals could also correspond to the kind of cells proposed in our model because their learning would be related to the presence of specific distant landmarks and their azimuth or apparent distance. The cells recorded in the ventral MEC could also correspond to some low-resolution grid cells, which we will discuss in the next section, or to the parahippocampal place area, where neurons were found to react to the visual recognition of places (Epstein and Kanwisher, 1998; Epstein, 2008). A last question related to our simple model of visual place cell is what happens if the field of view is limited to 180 deg, as in primates who have a smaller field of view than rodents? By using the same neural network as that used for place cells with a field of view of 180 deg, we obtained an activity pattern similar to the one recorded in the hippocampus of monkeys by Rolls and O'Mara (1995) (Fig. 2). It should be noted that the landmarks were tables, chairs, windows, cabinets, etc., in the room. Yet, no parallax issue related the apparent displacement of proximal landmarks can be seen in the results because the robot was not using two or three

landmarks but more than 20 local views for the learning and recognition of the view cells. This provides an averaging effect, reducing the parallax effect regarding any particular landmark. Moreover, landmarks perceived from two different points of view are not recognized and do not provide activity to the view cells, thus also reducing parallax issues in this case. In more complex cases, the solution proposed by Bicanski and Burgess (2016) may apply as a robust input to our network.

Grid cells from path integration

Although place cells occur in the HS and grid cells have been observed in the EC, we propose that PI in a two-dimensional environment is primarily performed outside the EC and the HS. PI can be performed on a 1D field of neurons on which activity bumps related to the current direction and speed of the animal are integrated. The summation of different bumps of activity having a cosine shape results in a single bump of activity in which the most active neuron corresponds to the direction of the global motion (PI vector) and the amplitude of its activity represents the distance traveled in a straight line. This model, inspired by insect navigation (Mittelstaedt and Mittelstaedt, 1980; Wehner and Srinivasan, 1981; Hartmann and Wehner, 1995; Collett et al., 1996; Heinze and Homberg, 2007; Stone et al., 2017), can be generalized to mammals if the cosine shape is replaced by a Gaussian-like shape (or a Von Mises function) representing the activity of the HD cells (built from a 1D attractor model using the angular speed as an input; see McNaughton et al., 1996). Even if the HD cell activity bump is limited to 120 deg, it is sufficient to perform a good approximation of PI, and a leaky integrator or synaptic learning is sufficient to perform PI (P.G. and J.L.K., unpublished data). Field activity can be recalibrated according to egocentric (the entrance of a maze) or allocentric information (a distant landmark). As a result, we can obtain allocentric information related to either turns or the route based on the time constant used in the leaky integrators. The resulting activity looks similar to cell activity recorded in the rat's RSC (Alexander and Nitz, 2015, 2017) and parietal cortex (Nitz, 2012), two of the main inputs to the EC.

At the level of the RSC and posterior parietal cortex, several 1D fields associated with different preferred directions may be maintained and used to perform navigation in the direction of one target according to some population code. The interesting point is that
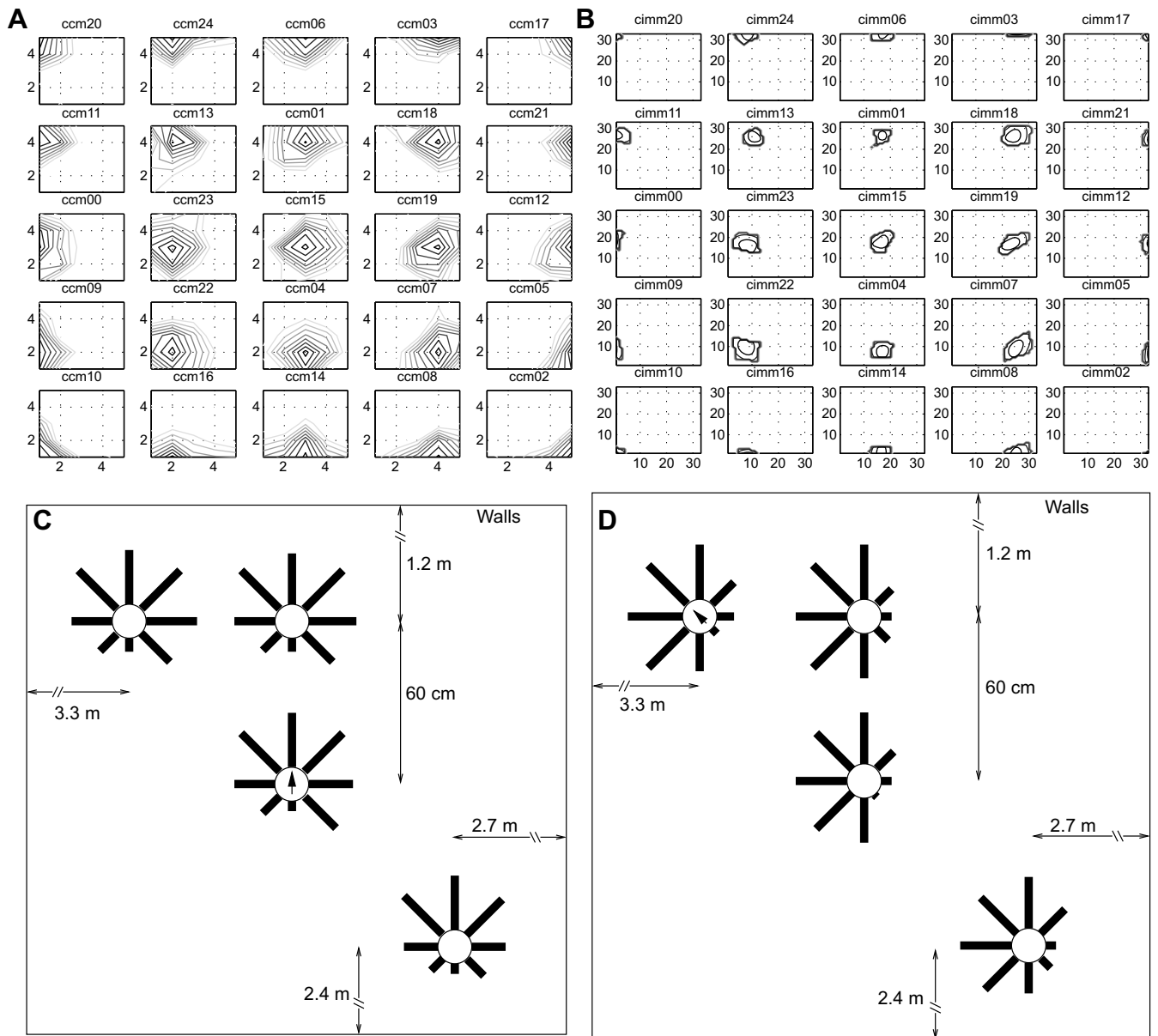
Fig. 2. Robot place cells from vision (Gaussier et al., 2002). (A) Each box represents the contour level of one cell labelled ccmXX, where XX is the neuron number. The cells react preferentially to their associated place in the room. (C,D) Two examples of robot view cells obtained in the same environment when the visual field of view is reduced from 300 to 180 deg (Gaussier et al., 2002). The circle with an arrow represents the position and orientation of the robot during the learning. All the other oriented bars represent the activity of the same cell for different positions while the robot is in the orientation of the bar. The length of the bars is proportional to their discharge frequency.

the coding of the direction on these maps would be similar to the one used in the motor cortex (Georgopoulos, 1988) allowing direct control of the action from those fields (Hasson and Gaussier, 2010). Next, if we assume that the activity of some neurons on the field is discretized and folded onto EC cells using a modulo projection, then we can obtain grid cells (Gaussier et al., 2007) (as shown in Fig. 3A).

On a field of neurons indexed with the variable i from 1 to N, we define a modulo projection (based on the mathematical modulo operator) as a parallel projection of N neurons on a group of M neurons. Every M neurons, a folding is performed so that the neuron j in the output group receives the projection of the activities of all the neurons i such that i mod $M=j$, where mod is the modulo operation and j is the remainder of the Euclidian division of i by M. In our case, grid cell activity results from a projection of PI and should not be seen as an interference model (Giocomo et al., 2011; Burgess et al.,

2007), even if the process is somehow similar to a spatial interference. The difference in relative direction preference of two neurons on the PI field defines the grid orientation whereas the modulo value determines the grid spacing.

The drift of PI over time requires us to introduce a recalibration procedure to allow spatial stabilization of the grid cells according to the animal position in its environment (Gaussier et al., 2007). The recalibration can be performed by visual place cells or from tactile information such as that stemming from environmental boundaries (Hardcastle et al., 2015). Robotics experiments (Jauffret et al., 2015) confirm that the sharpness of grid maps is better if the recalibration takes place in a corner of the arena rather than in the middle of the arena because the walls stop the robot and provide a more reliable positioning for the reset or recalibration of the PI field (see Fig. 3B for an example of grid cells obtained from the
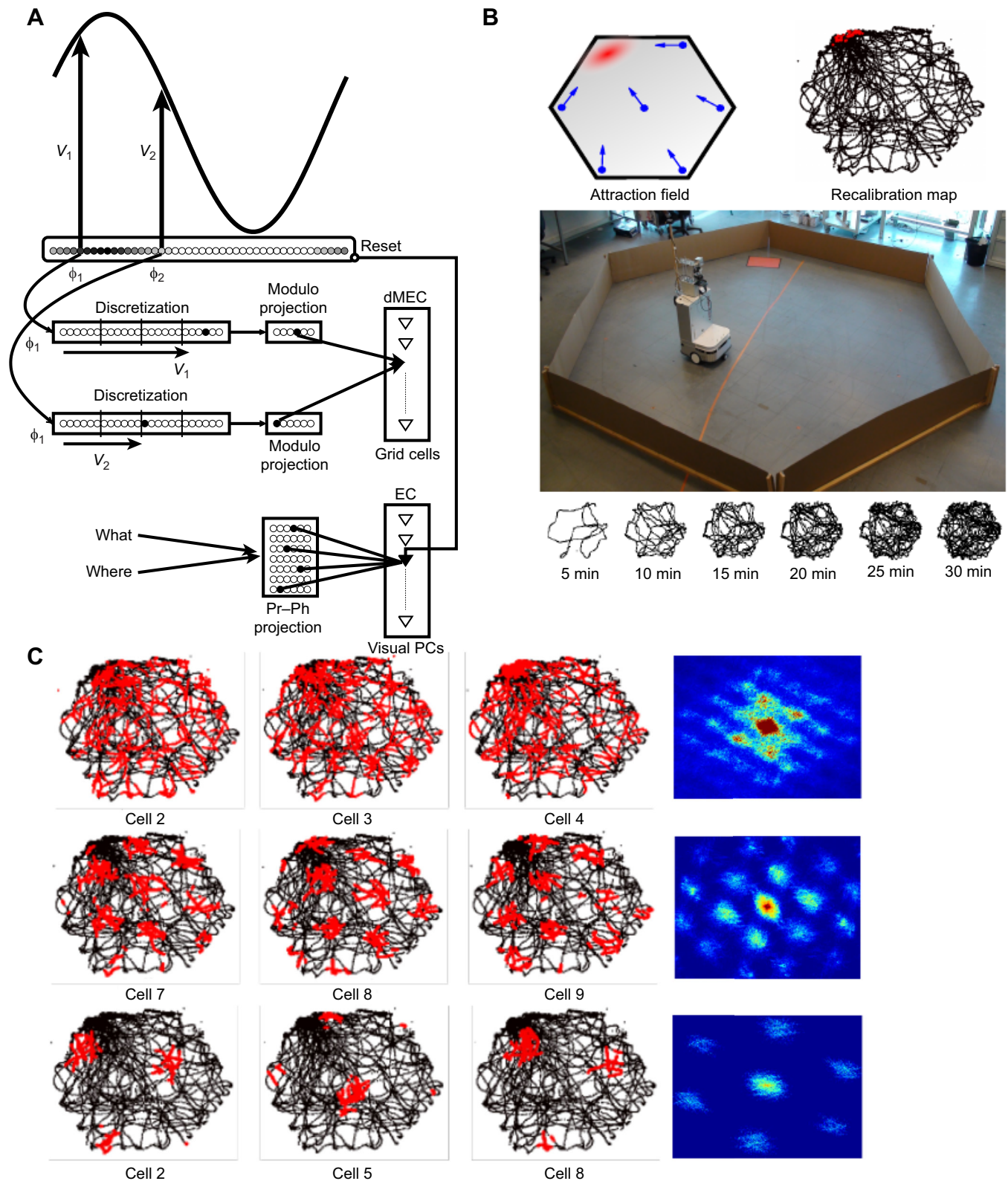
Fig. 3. Robot grid cells from path integration and visual recalibration. (A) Architecture for path integration and building grid code. The input is a one-dimensional path integration (PI) field. The activity of two neurons is discretized and projected with a modulo operator onto dorsal medial entorhinal cortex (dMEC) cells to build grid cells. Recalibration of the PI field is performed thanks to the visual recognition of one specific place (usually in the corner of the environment for better performances). V_i, activity of neuron ϕ_i on the PI field; EC, entorhinal cortex; PC, place cell; Pr, perirhinal; Ph, parahippocampal. (B) Example of grid cells recorded in a robotic experiment. Blue arrows represent the place/action association used by the robot to come back to the home base in the hexagonal environment. (C) The activity of nine grid cells associated with three resolutions or spacings and their respective auto-corellogram.

projection of the discretized activity of a 1D PI field using a modulo projection). From a spectrum of grid cells with different spatial frequencies, it becomes possible to build place cells (Solstad et al., 2006; Gaussier et al., 2007). Yet to obtain useful place recognition for navigation, it is important to avoid the merging of too sharp grid cells. As a matter of fact, the use of binary grid codes provides place cells with a weak generalization property. If the contribution of each grid is the same for the recognition of a place, this means that a change in the smallest grid resolution has the same effect as a change in the largest grid resolution, resulting in very poor navigation (and generalization) capabilities. To overcome this difficulty, a simple technical solution would be to smooth the grid activity pattern using positive lateral interaction between the neurons having the same grid spacing. Hence, there is a need to control the size of the peak activity of grid cells to obtain a good generalization. The need for both a competition mechanism, to learn a sparse code, and positive lateral interactions (or at least a positive effect of interaction), to allow better generalization capabilities, is reminiscent of the properties of the attractor networks. Yet, there is no particular constraint on the lateral interaction weights in terms of regularity of the connectivity or homogeneity in the weight values (e.g. to avoid bias in a given direction).

For these reasons, the grid cell system appears to be a better candidate to represent and categorize the displacements than a system for the computation of PI. This seems to be confirmed by the recordings of grid-like neuronal activity during human spatial navigation in a virtual environment (Jacobs et al., 2013) and in functional magnetic resonance imaging studies (Doeller et al., 2010; Stangl et al., 2018). It must be emphasized that in a multi-compartment environment, where the open field is broken up by barriers, the grid cell firing pattern is fragmented [the grid map is reset when the animal enters a new alley (Derdikman et al., 2009)]. Moreover, grid cells mature more slowly than place cells (Langston et al., 2010; Wills et al., 2010). This could support our initial intuition of a PI performed outside the HS for our grid cell model in the MEC (Gaussier et al., 2007).

Finally, it is interesting to note that grid cell firing is also observed when animals are not moving and when people are performing ocular saccades on a visual scene (Killian et al., 2012; Wilming et al., 2018), a task with no obvious relationship to navigation and PI. Some direction-sensitive cells may also relate to viewing particular landmarks (Rolls and O'Mara, 1995; Rolls, 1999; Ekstrom et al., 2003). In the next section, we examine whether grid cells can be built from visual information only (without the need of idiothetic information and PI).

Grid cells from still images
In the second section ('Place cells from visual information'), 'what' and 'where' information were merged in the EC to obtain large visual place fields, whereas in the third section ('Grid cells from path integration'), different projections of PI were merged after a modulo operation to obtain grid fields. What would happen if the modulo projection was applied to the azimuth or to the apparent distance of different landmarks connected to the same conjunctive cell used for building visual place cells?

We show here that simulated grid cell activity can be obtained from visual information only (similar to Killian et al., 2012; Wilming et al., 2018). For the sake of simplicity, we propose that the modulo operation is only performed on the angular or distance information related to some visual landmark (the 'where' information). If a cell learns at a given position the conjunction of the modulo projection of two distinct landmark fields (see Fig. 4A),

it will be reactivated for all the angular positions associated with the conjunctive cell [i.e. for all the angles θ_i such that $proj(\theta_i)=proj(\theta_{learned})$, where proj is a projection following the modulo principle (see Fig. 4A) and $\theta_{learned}$ is the azimuth associated with the neuron activated for the learned position]. The places where the two modulo patterns coincide correspond to grid activity as shown in Fig. 4B–D.

First, we consider the minimal case of an environment with two different landmarks (i.e. landmarks associated with the recognition of different visual features) L_1 and L_2 located at the coordinates (0,0) and (0,50) in a simulated environment. The azimuths are computed from landmark angular positions relative to an absolute direction obtained from idiothetic information (vestibular compass or even a visual compass; Delarboulas et al., 2014) measured at the center of the environment. We assume here that one conjunctive cell in the EC has learned the landmark configuration at the center of the environment [for the coordinates (25,25)]. Fig. 4C shows that grid patterns are directly obtained from the conjunction of the modulo of the azimuth (according to Eqn A1 in the Appendix). A high thresholding of cell activity, in our case set to 0.995, is necessary for contrast enhancement. Neurons with an activity level lower than the average activity are set to zero and a high gain is applied so that the maximum of activity is set to 1. Without this strong competition, all neurons would be very active.

Fig. 4E shows the results with four different identifiable landmarks located in the four corners of the visible environment. The grid distribution is much more regular in this case. Using the apparent size of the landmarks instead of their azimuth produces slightly different results as the apparent size varies as the arc tangent of the distance and induces non-homogeneous deformations (see Appendix for more details).

Yet, to obtain hexagonal grid maps, several constraints must be verified. (1) The 'where' information must be preserved when animal's displacements induces view changes. Either the recognition of 'what' information has to be powerful enough to be invariant to the animal's displacements, or each piece of 'where' information has to be stored in an attractor network, allowing a smooth update of the landmark heading when the animal is moving. (2) The conjunctive cells merge the activity of few landmarks. If the number of landmarks used to build one grid cell becomes too high (for instance, using 30 landmarks), the probability of obtaining a grid cell decreases because the landmarks allow a complete place discrimination in spite of the folding (sharp place cells are obtained instead of grid cells). (3) It is very important to merge landmarks associated with very different orientations relative to the position where the learning was triggered. The best results will be obtained for an angular distance approximately 60 deg between two landmarks. This value does not need to be precise, as shown in our simulations (e.g. see Fig. 4E), but still the use of the cell activity will be more relevant if the angular distance between the landmarks allows a good differentiation of the grid pattern. If we suppose that the 'where' information comes from an activity bump in the RSC or the parietal cortex, its shape could be similar to that of HD cells. If their activity covers 90 deg and we consider that a given conjunctive cell tries to maximize the independence between its inputs, then two landmarks with angular differences less than 45 deg will present an overlap and will have a lower independence score for learning (so their conjunction should not be selected for the learning of a grid cell). However, an angular landmark distance of approximately 60 deg will present the best independence score. If a modulo operator is used, then there will be coincidence for larger angular differences. Hence, if we assume the cells are conjunctive cells
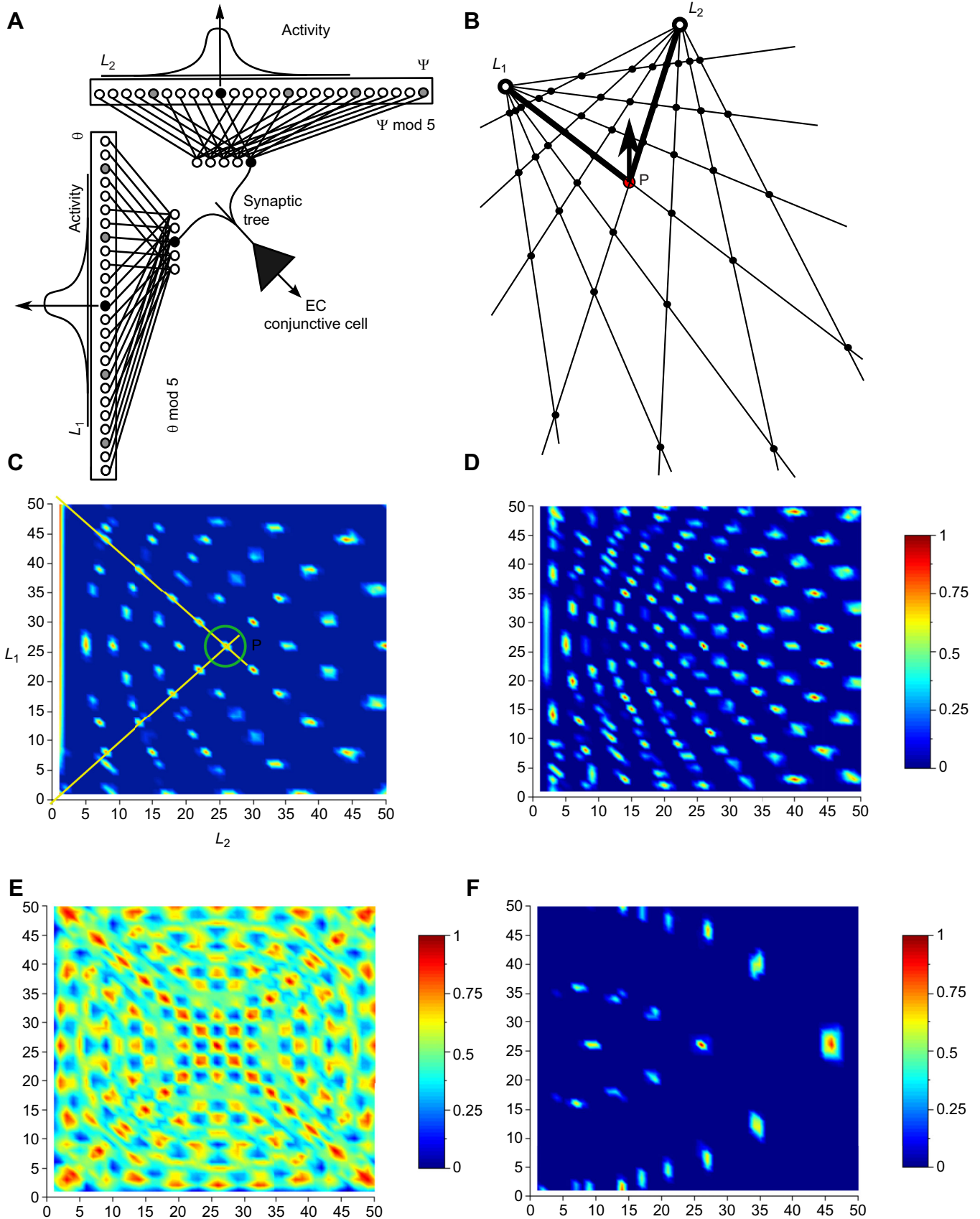
Fig. 4. See next page for legend.

Fig. 4. Building grid cells from visual information. (A) Neural network architecture for building grid cells based on the merging of at least two different fields (i.e. two azimuths, distance, etc.). θ and Ψ: discretized field showing the azimuth of landmark L_1 and L_2, respectively; θ mod 5 and Ψ mod 5: projections of θ and Ψ, respectively, with a modulo of 5. (B) Visual grid cells obtained from the conjunction of two landmark×azimuth couples with a modulo projection on the azimuth field. P, current position of the animal. (C,D) Activity of two visual grid cells according to the position of the simulated animal in a square environment (size 50×50 in simulation units). The activity is displayed after thresholding and contrast enhancement of the conjunction of two landmark×azimuth couples with a modulo projection on the azimuth field. (C) Modulo=180/20 and (D) modulo=180/30. (E) Visual grid cells obtained from the conjunction of four landmark×azimuth couples with a modulo projection on the azimuth field 180/30 and threshold=0.9. (F) Visual grid cell obtained from the conjunction of two landmark×visual distance couples with a modulo projection on the distance field size/6. Neuron threshold is 0.995.

using some kind of maximum of independence to build the modulo operator and to select landmarks, then grid cell patterns should be obtained. (4) To obtain effective grid activity, an important thresholding needs to be performed as well as a strong competition between the cells, especially if the bump activity on the 'where' field is large, allowing high activity for a distant azimuth, especially after the modulo projection.

Even if azimuth information seems more reliable than visual distance information, integration of both types of information in visual grid cells can provide more accurate discrimination (conjunction of the landmark identity×azimuth×elevation). In the brain, clearly the projection of the cortical areas coding for the landmark identity or features could also be compressed using the same folding mechanism proposed for azimuth and elevation without changing the main properties at study here.

Discussion and conclusions

Our work supports the hypothesis that the EC is able to build compact representations of cortical activity. Assuming that the HS tries to detect spatio-temporal configurations of the brain activity, it is important that the EC builds a compact code representing the complex instantaneous configuration of the brain activity (including sensory-motor activities, internal body state, emotions, motivations, etc.). To differentiate a huge number of states presumably largely superior to the number of neurons in the EC, each LEC or MEC neuron must receive projections from a large number of cortical neurons through the postrhinal and perirhinal cortices. If neighbor neurons in the cortex code for neighbor states and are projected onto the same EC neuron, then the capability to discriminate between neighbor states in the hippocampus would be very poor. A modulo projection (cyclic projection using a modulo operator) of cortical information on a set of EC neurons minimizes the risk that the same EC neuron will be activated from neighbor situations. More precisely, the modulo projections using prime numbers are a good way to build a hash code of the cortical activity. This projection minimizes collisions similar to how a correspondence table, used in computer science, represents a complex state (e.g. words in a dictionary) with a short reference code. Although in computer science, collisions are an important issue that need to be avoided, in our model, local collisions can be accepted for states belonging to different environments because other projections may provide a way to avoid confusion. The modulo projection could be explained by biological self-organization mechanisms following the Turing diffusion equation as proposed by McNaughton et al. (2006), by the competition between cells in an attractor network (Giocomo et al., 2011) or by the hexagonal tiling of place cells in a competitive

structure [if the recruitment of visual place cells is performed when the activity of the most activated place cell is below a given threshold, then the result is a hexagonal tiling of the environment (Gaussier et al., 2007) that could be used as a way to modulate EC learning]. Another solution could be based on synaptic learning rules using the hypothesis that neighboring neurons in the cortex cannot project onto the same neuron in the EC, or that EC neurons try to maximize the diversity of their inputs across their dendritic tree. If the input comes from some cortical HD cells with bumps of activity having a radius of approximately 60 deg, it would be possible that selecting landmarks with as little as possible overlap on the conjunctive cells would end in triangular grid activity of approximately 60 deg. Hence, it looks as if several different mechanisms could contribute alone or in mixed way to explain the necessary properties of the modulo mapping necessary in our model.

Starting from homing behaviors and detour problems in navigation (Gaussier et al., 2000), we conclude as others (Spiers and Gilbert, 2015) that 'path integration' or 'dead reckoning' (Etienne and Jeffery, 2004) may (Worsley et al., 2001; Wolbers et al., 2007; Stangl et al., 2018) or may not (Shrager et al., 2008; Kim et al., 2013) involve hippocampal–parahippocampal structures, and that this involvement perhaps depends on the distance navigated (Gaussier et al., 2007; Arnold et al., 2014). Our different studies, as well as work on insect modeling (Stone et al., 2017; Schwarz et al., 2017; Goldschmidt et al., 2017), support the hypothesis that PI can be obtained without the need of EC grid cells. HD cells or any kind of internal compass associated with local memories (to support temporal integration) can be used to perform PI in different brain structures [different kinds of reset/preset and different time constants are sufficient to explain many different cells found in the RSC and the parietal cortex (P.G. and J.L.K., unpublished data)]. Grid cells from PI can be explained from the projection of the discretized activity of a 1D field presumably located in the RSC without the need of a two-dimensional attractor network in the EC (Gaussier et al., 2007; P.G. and J.L.K., unpublished data). Some EC cells would be dedicated to representing in a concise way different cortical PI vectors thanks to a modulo projection. Yet, the PI mechanism required recalibration based on recognition of some specific places or boundaries/corners in the environment. This could explain why the disruption of the MEC seems to induce an incapacity to perform PI (Gil et al., 2018) and why young rats can recognize places and navigate correctly despite their immature grid cells (Bjerknes et al., 2018). Indeed, our model suggests that grid-like activity in the MEC is not necessary to build place cells, but place cells are important for the recalibration of PI.

Applying the modulo projection to the 'where' information in an image (i.e. relative azimuth and elevation) can explain other forms of grid activity found in the human brain (Killian et al., 2012, 2015; Killian and Buffalo, 2018). Our results on grid cells also bear a relation with results from Krupic et al. (2015) showing that grid cell symmetry is shaped by environmental geometry. In the same way, non-spatial information would also take advantage of the modulo projection and conjunctive cells could maximize the independence between input codes. They could correspond to some stripe activity or some grid activity in complex sensory space, explaining why the activity of these neurons is not easy to decode or associate with a specific situation. From our model, we predict that the activities related to non-spatial modalities should also provide some periodic folding. Projection of tactile information at the cortical level could generate interesting grid patterns. For audio information, or the perception of time, the information being mainly unidimensional,
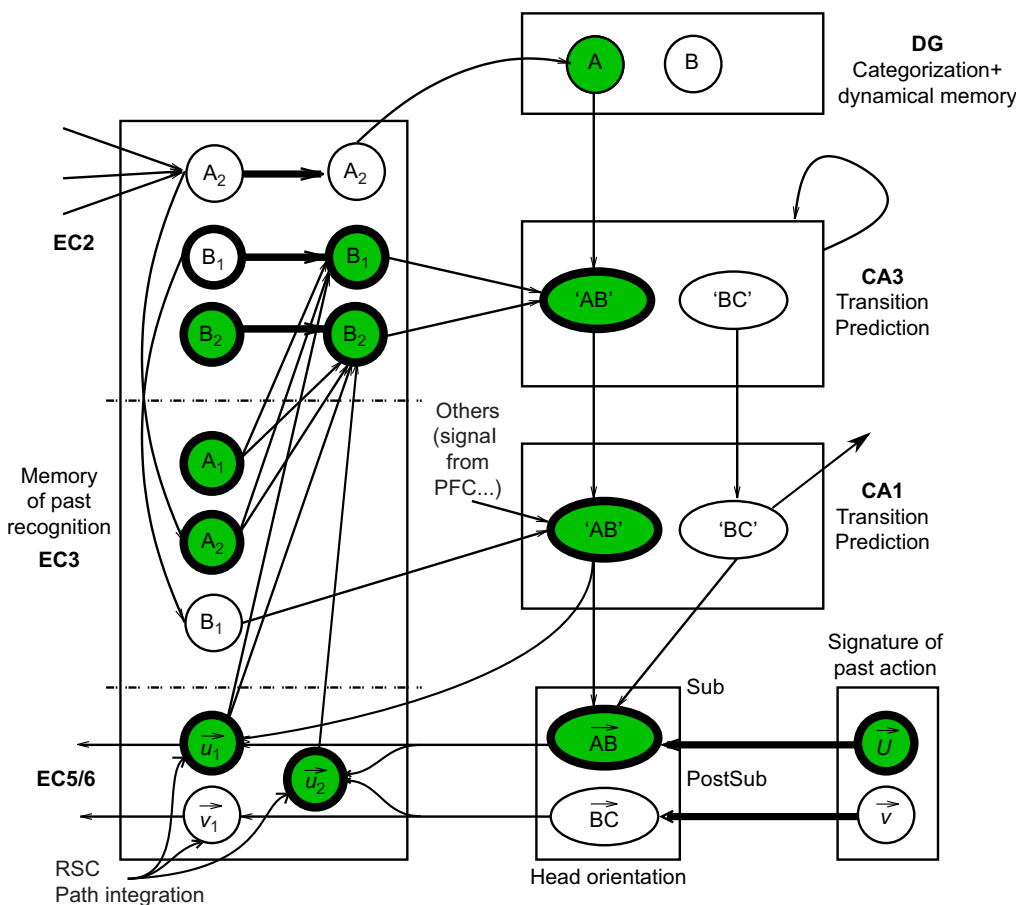
Fig. 5. Hippocampal loop showing how the prediction of transition in CA1 can be associated with one action in the subiculum (Sub) and come back in EC5 (deep layer) to improve the recognition of the current grid pattern at the level of EC2. In green, we represent some activated cells and how they should propagate their activity in the hippocampal system loop. A_i represent the neurons that were active at t–dt in EC2/3 (grid cells and others); B_i represent the active neurons at time t in EC2/3. A_i and B_i can be grid cells or any kind of other activity found in EC2/3, 'AB' and 'BC' represent two predicted transitions and $\overrightarrow{AB}$ and $\overrightarrow{BC}$ represent the direction associated with the selected transition in the subiculum. $\overrightarrow{u_1}$, $\overrightarrow{u_2}$ and $\overrightarrow{v_1}$ represent the activity of conjunction cells in the deep EC. $\overrightarrow{U}$ and $\overrightarrow{v}$ represent the direction of previous movements. DG, dentate gyrus; PFC, prefrontal cortex; RSC, retrosplenial cortex; PostSub, post subiculum.

the modulo folding should induce some stripe-like activity where the same neuron would be activated for increasing delays (perhaps non-periodic because the time perception is not linear with the clock time). Even some trace of temporal activities (if discretized) could activate periodically the same EC conjunctive cell.

To obtain grid patterns as predicted by our model, several constraints need to be verified. For visual grid cells, the competition in the EC superficial layer (EC2) has to be strong enough to help isolate the maxima associated with the grid activity. Yet, if the number of coherent landmark azimuths or distances used by a given conjunctive cell is too large, our simulations show that this cell becomes a place cell because a high response can only be obtained for a single location. This is not a strong constraint if we suppose that the time window for the temporal integration on conjunctive cells in the dMEC is limited or that information related to visual local views is merged before it reaches the EC, so that the EC receives only a limited number of signatures characterizing the visual environment (conjunction of visual signatures distant of an angle of ~60 deg). Next, because the modulo value controls the scale of the grid, the modulo value should be related to the position on the dorso-ventral axis of the EC. More work needs to be performed to determine whether the $\sqrt{2}$ scale variation found by Stensola et al. (2012) in the field spacing of grid cells could be related to the distance integration properties on the PI fields (leaky integrators) used as input. Our robotic experiments show that the projections of visual information without the modulo operator produce place cells or view cells with large fields. The modulo operator could help in building precise place recognition, and this would be useful for the hippocampus to recognize transitions from

one location to another. Yet this modulo mapping can induce the loss of *a priori* generalization capabilities to neighbor locations and the incapacity to perform a shortcut or a detour to reach a goal when starting from a place never visited but belonging to a known visual environment. If the projection on the ventral MEC is not performed using a modulo projection or with a very large modulo, then the ventral MEC could code for the large visual place cells used in our robots for visual navigation in an open environment (large place cells allow some competition mechanism to find a shortcut and reach a homing location from a long distance). These cells could correspond to those found previously (Quirk et al., 1992; Song et al., 2012) or the unclassified MEC cells (Savelli et al., 2008).

Finally, in this paper, we focused only on the feedforward integration of information coming from associative cortical areas onto the EC with the hippocampus proper as a merging output to obtain state categorization. In previous work (Gaussier et al., 2002; Banquet et al., 2005), we proposed that the hippocampus learned to predict transitions between places or any state characterizing the current behavior (and its timing) for the building of cognitive graphs. In the fronto-parietal network, the feedback loop from neurons in CA1 to the subiculum and the deep layers of EC (EC5/6) can be used as a way to predict the next states (see Fig. 5). Our modulo projection model could take advantage of any feedback from the hippocampal activity to stabilize the grid activity. Because our model shows that grid cells may be obtained from different modalities, we face the issue of the alignment of these different grids. In navigation tasks, if grid cells are primarily obtained from the modulo projection of a discretized PI code (or landmarks) built outside the HS (presumably in the RSC), maintaining the alignment

of the grid pattern when one goes from one environment to another and then back again requires recalibrating the PI and/or the head orientation always in the same way as in the original environment. Assuming the recognition of specific places or transitions between places is performed in the hippocampus proper, impairment at the hippocampal level should suppress the coherent recalibration of the different PI fields inducing drift in the grid patterns from one experiment to the next, as shown by Stangl et al. (2018) for Alzheimer's disease.

However, this PI outside the HS does not suppress the capabilities of the hippocampal loop to maintain and update grid patterns in the EC. As seen on Fig. 5, movement integration activity coming from the subiculum and its thalamic input the anterodorsal nucleus of the thalamus (where HD cells are found) can be easily associated with the transition between 'important' places to build conjunctive grid×HD cells (Sargolini et al., 2006). It would be interesting to classify CA1 neurons as 'transitions cells' instead of classical 'place cells' because one transition can trigger a unique action whereas a place cell can be associated with multiple actions. Hence, while selecting an action from a winning place cell is not direct because several actions may be associated with the same place, a unique action will be associated with a winning transition (Gaussier et al., 2002; Banquet et al., 2005; Hirel et al., 2013). Moreover, the transitions should predict the future state of the system and can be very useful for self-assessment [novelty detection, detection of deadlock, etc. (Jauffret et al., 2013)] and the building of a cognitive map in the fronto-parietal network. As a result, the hippocampus can be seen as a predictor of complex spatio-temporal and multimodal events allowing the detection of novelty and the coding of future transitions (and timing).

Results showing the importance of theta oscillations (Brandon et al., 2011; Koenig et al., 2011) and the excitatory drive from the hippocampus (Bonnevie et al., 2013) or the existence of grid cells without theta oscillations in the EC of bats (Yartsev et al., 2011) could be interpreted in light of the hippocampal loop and the interaction between the hippocampus and the septum (Hasselmo and Schnell, 1994; Hasselmo, 2006). Our model would be also consistent with Sanders et al.'s (2015) questioning whether PI occurs in the EC and their proposal about mind-travel in the hippocampus (see also Gorchetchnikov and Grossberg, 2007; Mhatre et al., 2012; Grossberg and Pilly, 2014). The formation of grid cells from an extra-hippocampal PI does not exclude the capacity of the EC to combine their 'stripe cells' (note that the proposed 'stripe cells' could correspond to the result of the modulo projection of the discretized activity of our one-dimensional PI field in a given direction (Gaussier et al., 2007)] to form grid cells and/or to use the hippocampal loop to predict the evolution of the grid pattern, especially if we suppose that EC neurons learn the correlations between different input signals. One important point in the future will be to determine which pathway(s) control the formation of the grid cells and their shape. It is likely that some pathways are very important for the maturation and organization of grid cells and that other sources of information could provide sufficient information to maintain the grid patterns (and to exploit them) in the absence of the first ones. These are very challenging questions both for cognitive sciences and for the development of robust navigation systems.

Appendix

In the simulations, the animal is located according to its (x,y) floating-point coordinates, its head direction and instantaneous speed (mass and other inertial constraints are neglected because

navigation issues are not our main topic in this paper). There is no discretization for the agent location. The landmarks are recognized by a group of neurons. The output of these neurons can be represented by a vector L such that L_i is the recognition level of the neuron i. We assume either a sequential exploration of the landmarks or a simultaneous access to fields of neurons associated with azimuths or some reference point. At each time step, a landmark can be recognized and the gaze direction can be coded in the azimuth field for horizontal information [see Giovannangeli et al. (2006) for more detail of the neural network taking into account both azimuth and site information to recognize a place in navigation tasks in robots]. The vertical position of the landmark could be associated with an apparent distance field. To avoid writing the discretized equation of the neural fields and its projection, we can use the variable θ to represent both the angle and the neuron associated with this angle and the same for its modulo. In this case, the activity of one grid cell at position p can be written as follows:

$$G(p) = \prod_{i=1}^{n} \left(1 - h_i \left\| (\theta_i \bmod m_i) - W_i^{\theta_{m_i}} \right\| \right), \qquad (A1)$$

where $W_i^{\theta_{m_i}}$ represents the ith link associated with θ_i after the modulo operation with m_i, and h_i is a scaling factor defining the maximum decrease of the conjunction according to the variation of projection after the modulo operator. In a more general form, the grid cell equation can also be written as follows:

$$G_{k_1,k_2,\ldots,k_n}(p) = \prod_{k=k_1}^{k_n} \sum_{i=1}^{N_k} f_k(F^k(p,i) \times ((i \bmod m_k) = W_k)),$$

$$(A2)$$

where i is the index of the neuron on the F^k field, the kth input field, and $F^k(p,i)$ is the activity of the neuron i on the field when the animal is at position p. Each field is composed of N_k neurons (the size of the field). m_k is the value of the modulo operator applied on the projection of the field, and $f_k()$ is a function representing the activity bump on the input field k.

Place cell and not grid cell activity

If the number of landmarks used increases, the probability of obtaining grid cells decreases. Yet, place discrimination increases owing to the increase of independent information. In the following example (Fig. S1), with 10 random landmarks located on the four different sides of the environment, we have still some kind of grid or wave activity. However, with 30 random landmarks, there is a peak of activity at the center of the arena: the location used to build the code. We can also obtain cells that look like place cells with modulo $\pi/4$ when using 30 'landmarks' randomly distributed around the simulated room (see Fig. S2). In this case, the 'place field' is much larger because of the choice of a higher value for the modulo.

Visual grid cells from apparent landmark size

Distances can be obtained from the apparent size of a given landmark in the camera image. This apparent size corresponds to the elevation of the landmark. Hence, for a landmark of height h located at a distance d from the animal, the visual angle is $\phi=\arctan(h/d)$. This information has to be scaled according to the number of neurons or pixels in the vertical dimension of the camera. In our case, for the sake of comparison, we consider that an angle of $\pi/2$ corresponds to the maximal distance $d_{\max}$ to allow comparisons

between azimuth and apparent size (or elevation). The visual distance is computed as $d_v=d_{max}\cdot(\phi/2\pi)$, where ϕ is the elevation. In our case, we simulate a room of 5×5 m with $d_{max}=2\sqrt{5}$ and visual landmark of size $d=1$ m. We see the activity decreases less than from classical distance. With a threshold of 0.995, there are still plenty of activated locations.

For comparison purposes, in the following simulations, the distance is used explicitly as it can be provided by a regular PI starting from each landmark. With two landmarks in the corners and a modulo value equal to the size of the environment divided by 6, we have still grid cells has shown in Fig. S3A,B. With 30 landmarks and a modulo equal to the size of the environment divided by 10, the grid pattern disappears as with the apparent size information (Fig. S3C,D). With four landmarks taken at the corners, we obtain quite regular grid cells, as shown Fig. S4. The activity pattern looks very regular because of the specific geometrical properties of this configuration. The distortions are not visible in this case.

References
Alexander, A. S. and Nitz, D. A. (2015). Retrosplenial cortex maps the conjunction of internal and external spaces. *Nat. Neurosci.* **18**, 1143-1151.

Alexander, A. S. and Nitz, D. A. (2017). Spatially periodic activation patterns of retrosplenial cortex encode route sub-spaces and distance traveled. *Curr. Biol.* **27**, 1551-1560.

Arnold, A. E., Burles, F., Bray, S., Levy, R. M. and Iaria, G. (2014). Differential neural network configuration during human path integration. *Front. Hum. Neurosci.* **8**, 263.

Banquet, J.-P., Gaussier, P., Quoy, M., Revel, A. and Burnod, Y. (2005). A hierarchy of associations in hippocampo-cortical systems: cognitive maps and navigation strategies. *Neural Comput.* **17**, 1339-1384.

Bicanski, A. and Burgess, N. (2016). Environmental anchoring of head direction in a computational model of retrosplenial cortex. *J. Neurosci.* **36**, 11601-11618.

Bjerknes, T. L., Dagslott, N. C., Moser, E. I. and Moser, M.-B. (2018). Path integration in place cells of developing rats. *Proc. Natl Acad. Sci. USA*, 201719054.

Bonnevie, T., Dunn, B., Fyhn, M., Hafting, T., Derdikman, D., Kubie, J. L., Roudi, Y., Moser, E. I. and Moser, M.-B. (2013). Grid cells require excitatory drive from the hippocampus. *Nat. Neurosci.* **16**, 309.

Brandon, M. P., Bogaard, A. R., Libby, C. P., Connerney, M. A., Gupta, K. and Hasselmo, M. E. (2011). Reduction of theta rhythm dissociates grid cell spatial periodicity from directional tuning. *Science* **332**, 595-599.

Burgess, N., Barry, C. and O'keefe, J. (2007). An oscillatory interference model of grid cell firing. *Hippocampus* **17**, 801-812.

Cohen, N. and Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System*. MIT Press.

Collett, T. S. and Baron, J. (1995). Learnt sensori-motor mappings in honeybees: interpolation and its possible relevance to navigation. *J. Comp. Physiol. A* **177**, 287-298.

Collett, T. S., Baron, J. and Sellen, K. (1996). On the encoding of movement vectors by honeybees. are distance and direction represented independently? *J. Comp. Physiol. A* **179**, 395-406.

Cooper, B. G. and Mizumori, S. J. (2001). Temporary inactivation of the retrosplenial cortex causes a transient reorganization of spatial coding in the hippocampus. *J. Neurosci.* **21**, 3986-4001.

Delarboulas, P., Gaussier, P., Caussy, R. and Quoy, M. (2014). Robustness study of a multimodal compass inspired from hd-cells and dynamic neural fields. International Conference on Simulation of Adaptive Behavior, pp. 132-143. Springer.

Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M.-B. and Moser, E. I. (2009). Fragmentation of grid cell maps in a multicompartment environment. *Nat. Neurosci.* **12**, 1325-1332.

Deshmukh, S. S. and Knierim, J. J. (2011). Representation of non-spatial and spatial information in the lateral entorhinal cortex. *Front. Behav. Neurosci.* **5**, 69.

Doeller, C. F., Barry, C. and Burgess, N. (2010). Evidence for grid cells in a human memory network. *Nature* **463**, 657.

Eichenbaum, H., Otto, T. and Cohen, N. J. (1992). The hippocampus—what does it do? *Behav. Neural. Biol.* **57**, 2-36.

Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L. and Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature* **425**, 184.

Elduayen, C. and Save, E. (2014). The retrosplenial cortex is necessary for path integration in the dark. *Behav. Brain Res.* **272**, 303-307.

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn. Sci.* **12**, 388-396.

Epstein, R. and Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature* **392**, 598.

Etienne, A. (1998). Mammalian navigation, neural models and biorobotics. *Connect. Sci.* **10**, 271-289.

Etienne, A. S. and Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus* **14**, 180-192.

Gaussier, P. and Zrehen, S. (1995). Perac: a neural architecture to control artificial animals. *Robotics and Autonomous System* **16**, 291-320.

Gaussier, P., Joulain, C., Zrehen, S., Banquet, J. and Revel, A. (1997). Visual navigation in an open environement without map. In International Conference on Intelligent Robots and Systems - IROS'97, Grenoble, France, pp. 545-550. IEEE/RSJ.

Gaussier, P., Joulain, C., Banquet, J., Leprêtre, S. and Revel, A. (2000). The visual homing problem: an example of robotics/biology cross fertilization. *Rob. Auton. Syst.* **30**, 155-180.

Gaussier, P., Revel, A., Banquet, J.-P. and Babeau, V. (2002). From view cells and place cells to cognitive map learning: processing stages of the hippocampal system. *Biol. Cybern.* **86**, 15-28.

Gaussier, P., Banquet, J., Sargolini, F., Giovannangeli, C., Save, E. and Poucet, B. (2007). A model of grid cells involving extra hippocampal path integration, and the hippocampal loop. *J. Integr. Neurosci.* **6**, 447-476.

Georgopoulos, A. (1988). Neural interpretation of movement: role of motor cortex in reaching. *FASEB J.* **13**, 2846-2857.

Geva-Sagiv, M., Las, L., Yovel, Y. and Ulanovsky, N. (2015). Spatial cognition in bats and rats: from sensory acquisition to multiscale maps and navigation. *Nat. Rev. Neurosci.* **16**, 94.

Gil, M., Ancau, M., Schlesiger, M. I., Neitz, A., Allen, K., De Marco, R. J. and Monyer, H. (2018). Impaired path integration in mice with disrupted grid cell firing. *Nat. Neurosci.* **21**, 81.

Giocomo, L. M., Moser, M.-B. and Moser, E. I. (2011). Computational models of grid cells. *Neuron* **71**, 589-603.

Giovannangeli, C., Gaussier, P. and Banquet, J. P. (2006). Robustness of visual place cells in dynamic indoor and outdoor environment. *Int. J. Adv. Robot. Syst.* **3**, 115-124.

Goldschmidt, D., Manoonpong, P. and Dasgupta, S. (2017). A neurocomputational model of goal-directed navigation in insect-inspired artificial agents. *Front. Neurorobot.* **11**, 20.

Gorchetchnikov, A. and Grossberg, S. (2007). Space, time and learning in the hippocampus: How fine spatial and temporal scales are expanded into population codes for behavioral control. *Neural Netw.* **20**, 182-193.

Grossberg, S. and Pilly, P. K. (2014). Coordinated learning of grid cell and place cell spatial and temporal properties: multiple scales, attention and oscillations. *Philos. Trans. R. Soc. B* **369**, 20120524.

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

Hardcastle, K., Ganguli, S. and Giocomo, L. M. (2015). Environmental boundaries as an error correction mechanism for grid cells. *Neuron* **86**, 827-839.

Hargreaves, E. L., Rao, G., Lee, I. and Knierim, J. J. (2005). Major dissociation between medial and lateral entorhinal input to dorsal hippocampus. *Science* **308**, 1792-1794.

Hartmann, G. and Wehner, R. (1995). The ants path integration system: a neural architecture. *Biol. Cybern.* **73**, 483-497.

Hasselmo, M. E. (2006). The role of acetylcholine in learning and memory. *Curr. Opin. Neurobiol.* **16**, 710-715.

Hasselmo, M. E. and Schnell, E. (1994). Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region ca1: computational modeling and brain slice physiology. *J. Neurosci.* **14**, 3898-3914.

Hasson, C. and Gaussier, P. (2010). Path integration working memory for multi tasks dead reckoning and visual navigation. In *From Animals to Animats 11*. SAB 2010. Lecture Notes in Computer Science, vol 6226 (ed. S. Doncieux, B. Girard, A. Guillot, J. Hallam, J. A. Meyer and J. B. Mouret), pp. 380-389. Springer.

Heinze, S. and Homberg, U. (2007). Maplike representation of celestial e-vector orientations in the brain of an insect. *Science* **315**, 995-997.

Hirel, J., Gaussier, P., Quoy, M., Banquet, J.-P., Save, E. and Poucet, B. (2013). The hippocampo-cortical loop: spatio-temporal learning and goal-oriented planning in navigation. *Neural Netw.* **43**, 8-21.

Insausti, R., Amaral, D. and Cowan, W. (1987). The entorhinal cortex of the monkey: Iii. subcortical afferents. *J. Comp. Neurol.* **264**, 396-408.

Jacobs, J., Weidemann, C. T., Miller, J. F., Solway, A., Burke, J. F., Wei, X.-X., Suthana, N., Sperling, M. R., Sharan, A. D., Fried, I. et al. (2013). Direct recordings of grid-like neuronal activity in human spatial navigation. *Nat. Neurosci.* **16**, 1188.

Jauffret, A., Cuperlier, N., Gaussier, P. Tarroux, P. (2013). From self-assessment to frustration, a small step toward autonomy in robotic navigation. *Front. Neurorobot.* **7**, 16.

Jauffret, A., Cuperlier, N. and Gaussier, P. (2015). From grid cells and visual place cells to multimodal place cell: a new robotic architecture. *Front. Neurorobot.* **9**, 1.

Jung, M. W., Wiener, S. I. and McNaughton, B. L. (1994). Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *J. Neurosci.* **14**, 7347-7356.

Killian, N. J. and Buffalo, E. A. (2018). Grid cells map the visual world. *Nat. Neurosci.* **21**, 161.

Killian, N. J., Jutras, M. J. and Buffalo, E. A. (2012). A map of visual space in the primate entorhinal cortex. *Nature* **491**, 761.

Killian, N. J., Potter, S. M. and Bualo, E. A. (2015). Saccade direction encoding in the primate entorhinal cortex during visual exploration. *Proc. Natl Acad. Sci. USA* **112**, 15743-15748.

Kim, S., Sapiurka, M., Clark, R. E. and Squire, L. R. (2013). Contrasting effects on path integration after hippocampal damage in humans and rats. *Proc. Natl Acad. Sci. USA* **110**, 4732-4737.

Knierim, J. J., Neunuebel, J. P. and Deshmukh, S. S. (2014). Functional correlates of the lateral and medial entorhinal cortex: objects, path integration and local–global reference frames. *Phil. Trans. R. Soc. B* **369**, 20130369.

Koenig, J., Linder, A. N., Leutgeb, J. K. and Leutgeb, S. (2011). The spatial periodicity of grid cells is not sustained during reduced theta oscillations. *Science* **332**, 592-595.

Kolb, B. and Tees, R. (1990). *The Cerebral Cortex of the Rat*. MIT Press.

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

Langston, R. F., Ainge, J. A., Couey, J. J., Canto, C. B., Bjerknes, T. L., Witter, M. P., Moser, E. I. and Moser, M.-B. (2010). Development of the spatial representation system in the rat. *Science* **328**, 1576-1580.

Lashley, K. (1938). The mechanism of vision XV: preliminary studies of the rat's capacity for detail vision. *J. Genral Psychol.* **18**, 123-193.

Lepretre, S., Gaussier, P. and Co, J. (2000). From navigation to active object recognition. In The Sixth International Conference on Simulation for Adaptive Behavior SAB'2000, Paris, pp. 266-275. MIT Press.

McClelland, J. L., McNaughton, B. L. and O'reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419.

McNaughton, B., Barnes, C., Gerrard, J., Gothard, K., Jung, M., Knierim, J., Kudrimoti, H., Qin, Y., Skaggs, W., Suster, M. et al. (1996). Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J. Exp. Biol.* **199**, 173-185.

McNaughton, B., Battaglia, F. P., Jensen, O., Moser, E. I. and Moser, M.-B. (2006). Path integration and the neural basis of the 'cognitive map'. *Nat. Rev. Neurosci.* **7**, 663-678.

Mhatre, H., Gorchetchnikov, A. and Grossberg, S. (2012). Grid cell hexagonal patterns formed by fast selforganized learning within entorhinal cortex. *Hippocampus* **22**, 320-334.

Mittelstaedt, M.-L. and Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften* **67**, 566-567.

Mizuseki, K., Royer, S., Diba, K. and Buzsáki, G. (2012). Activity dynamics and behavioral correlates of ca3 and ca1 hippocampal pyramidal neurons. *Hippocampus* **22**, 1659-1680.

Moser, E. I., Moser, M.-B. and Roudi, Y. (2014). Network mechanisms of grid cells. *Phil. Trans. R. Soc. B* **369**, 20120511.

Nitz, D. A. (2012). Spaces within spaces: rat parietal cortex neurons register position across three reference frames. *Nat. Neurosci.* **15**, 1365-1367.

O'Keefe, J. and Nadel, N. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press.

Parron, C. and Save, E. (2004). Evidence for entorhinal and parietal cortices involvement in path integration in the rat. *Exp. Brain Res.* **159**, 349-359.

Poucet, B., Thinus-Blanc, C. and Muller, R. U. (1994). Place cells in the ventral hippocampus of rats. *Neuroreport* **5**, 2045-2048.

Poucet, B., Sargolini, F., Song, E. Y., Hangya, B., Fox, S. and Muller, R. U. (2014). Independence of landmark and self-motion-guided navigation: a different role for grid cells. *Phil. Trans. R. Soc. B* **369**, 20130370.

Quirk, G. J., Muller, R. U., Kubie, J. L. and Ranck, J. (1992). The positional firing properties of medial entorhinal neurons: description and comparison with hippocampal place cells. *J. Neurosci.* **12**, 1945-1963.

Redish, A. D. and Touretzky, D. S. (1997). Cognitive maps beyond the hippocampus. *Hippocampus* **7**, 15-35.

Rolls, E. T. (1999). Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus* **9**, 467-480.

Rolls, E. and O'Mara, S. (1995). View-responsive neurons in the primate hippocampal complex. *Hippocampus* **5**, 409-424.

Rolls, E. T., Stringer, S. M. and Elliot, T. (2006). Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning. *Netw. Comput. Neural Syst.* **17**, 447-465.

Samsonovich, A. and McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *J. Neurosci.* **17**, 5900-5920.

Sanders, H., Rennó-Costa, C., Idiart, M. and Lisman, J. (2015). Grid cells and place cells: an integrated view of their navigational and memory function. *Trends Neurosci.* **38**, 763-775.

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

Save, E. and Sargolini, F. (2017). Disentangling the role of the MEC and LEC in the processing of spatial and non-spatial information: contribution of lesion studies. *Front. Syst. Neurosci.* **11**, 81.

Save, E., Guazzelli, A. and Poucet, B. (2001). Dissociation of the effects of bilateral lesions of the dorsal hippocampus and parietal cortex on path integration in the rat. *Behav. Neurosci.* **115**, 1212-1223.

Savelli, F., Yoganarasimha, D. and Knierim, J. J. (2008). Influence of boundary removal on the spatial representations of the medial entorhinal cortex. *Hippocampus* **18**, 1270-1282.

Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Res.* **20**, 645-669.

Schwarz, S., Mangan, M., Zeil, J., Webb, B. and Wystrach, A. (2017). How ants use vision when homing backward. *Curr. Biol.* **27**, 401-407.

Scoville, W. and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* **20**, 11-21.

Shrager, Y., Kirwan, C. B. and Squire, L. R. (2008). Neural basis of the cognitive map: path integration does not require hippocampus or entorhinal cortex. *Proc. Natl Acad. Sci. USA* **105**, 12034-12038.

Solstad, T., Moser, E. I. and Einevoll, G. T. (2006). From grid cells to place cells: a mathematical model. *Hippocampus* **16**, 1026-1031.

Song, E. Y., Rivard, S. and Muller, B. (2012). Neuronal representations of two visual stimuli modifications on the network level of the spatial cognition in rat brain. *New Orleans LA Soc. Neurosci.* **38**, 1.

Spiers, H. J. and Gilbert, S. J. (2015). Solving the detour problem in navigation: a model of prefrontal and hippocampal interactions. *Front. Hum. Neurosci.* **9**, 125.

Stangl, M., Achtzehn, J., Huber, K., Dietrich, C., Tempelmann, C. and Wolbers, T. (2018). Compromised grid cell-like representations in old age as a key mechanism to explain age-related navigational deficits. *Curr. Biol.* **28**, 1108-1115.

Stensola, H., Stensola, T., Solstad, T., Frøland, K., Moser, M.-B. and Moser, E. I. (2012). The entorhinal grid map is discretized. *Nature* **492**, 72.

Stone, T., Webb, B., Adden, A., Weddig, N. B., Honkanen, A., Templin, R., Wcislo, W., Scimeca, L., Warrant, E. and Heinze, S. (2017). An anatomically constrained model for path integration in the bee brain. *Curr. Biol.* **27**, 3069-3085.

Teyler, T. J. and DiScenna, P. (1986). The hippocampal memory indexing theory. *Behav. Neurosci.* **100**, 147.

Touretzky, D. S. and Redish, A. D. (1996). Theory of rodent navigation based on interacting representations of space. *Hippocampus* **6**, 247-270.

Tsao, A., Moser, M.-B. and Moser, E. I. (2013). Traces of experience in the lateral entorhinal cortex. *Curr. Biol.* **23**, 399-405.

Ungerleider, L. G. and Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Curr. Opin. Neurobiol.* **4**, 157-165.

Van Cauter, T., Camon, J., Alvernhe, A., Elduayen, C., Sargolini, F. and Save, E. (2012). Distinct roles of medial and lateral entorhinal cortex in spatial cognition. *Cereb. Cortex* **23**, 451-459.

Vann, S. D., Aggleton, J. P. and Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* **10**, 792-802.

Wan, H., Touretzky, D. and Redish, A. (1994). Towards a computational theory of rat navigation. In *Proc. of the 1993 Connectionist Models Summer School* (ed. M. Mozer, P. Smolensky, D. Touretzky, J. Elman and A. Weigend), pp. 11-19. Lawrence Erlbaum Associates.

Wehner, R. and Srinivasan, M. V. (1981). Searching behaviour of deserts ants, genus *Cataglyphis* (Formicidae, Hymenoptera). *J. Comp. Physiol. A* **142**, 315-338.

Wills, T. J., Cacucci, F., Burgess, N. and O'keefe, J. (2010). Development of the hippocampal cognitive map in preweanling rats. *Science* **328**, 1573-1576.

Wilming, N., König, P., König, S. and Buffalo, E. A. (2018). Entorhinal cortex receptive fields are modulated by spatial attention, even without movement. *eLife* **7**, e31745.

Wittmann, T. and Schwegler, H. (1995). Path integration – a network model. *Biol. Cybern.* **12**, 569-575.

Wolbers, T., Wiener, J. M., Mallot, H. A. and Büchel, C. (2007). Differential recruitment of the hippocampus, medial prefrontal cortex, and the human motion complex during path integration in humans. *J. Neurosci.* **27**, 9408-9416.

Worsley, C. L., Recce, M., Spiers, H. J., Marley, J., Polkey, C. E. and Morris, R. G. (2001). Path integration following temporal lobectomy in humans. *Neuropsychologia* **39**, 452-464.

Yartsev, M. M., Witter, M. P. and Ulanovsky, N. (2011). Grid cells without theta oscillations in the entorhinal cortex of bats. *Nature* **479**, 103.

Yoganarasimha, D., Rao, G. and Knierim, J. J. (2011). Lateral entorhinal neurons are not spatially selective in cue-rich environments. *Hippocampus* **21**, 1363-1374.

Zilli, E. A. (2012). Models of grid cell spatial firing published 2005–2011. *Front. Neural Circuits* **6**, 16.

Zola-Morgan, S., Squire, L., Amaral, D. and Suzuki, W. (1989). Lesions of perirhinal and parahippocampal cortex that spare the amygdala and hippocampal formation produce severe memory impairment. *J. Neurosciences* **9**, 4355-4370.
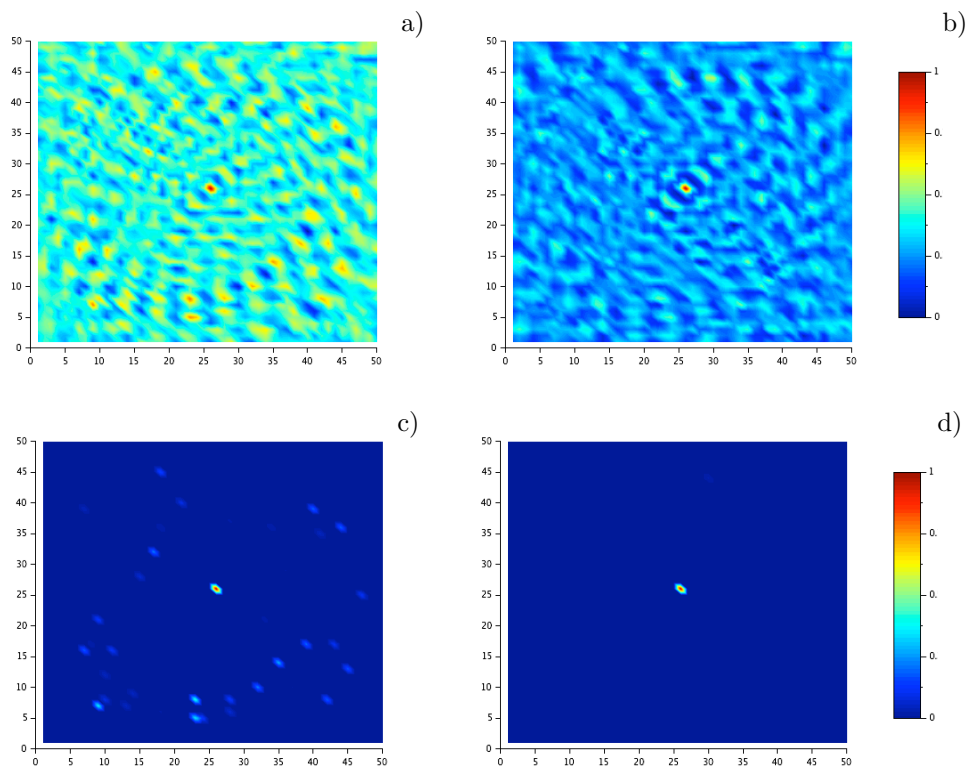
Figure S1: Visual grid cells obtained from the conjunction of respectively 10 (a) and 30 (b) random landmarks x azimuths couples with a modulo projection on the azimuth field $\pi/30$ and threshold=0. (c) and (d) again 10 and 30 landmarks with a threshold of 0.995.
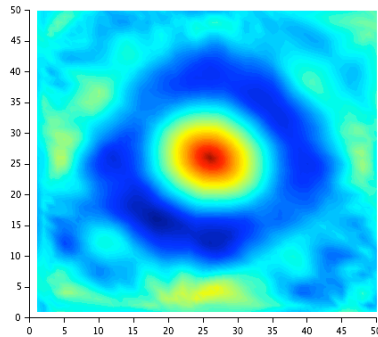
Figure S2: Large visual place cells obtained from the conjunction of 30 random landmarks x azimuths couples with a modulo projection on the azimuth field $\pi/4$ and threshold=0.
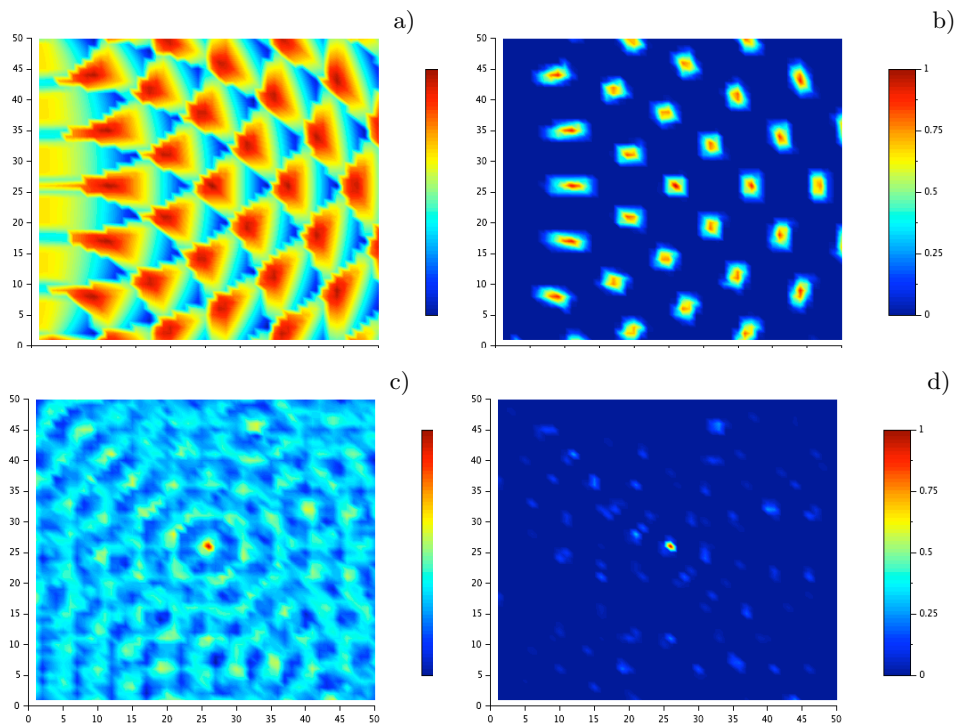

Figure S3: a) and b) Grid cells obtained from the conjunction of 2 landmarks x distance couples with a modulo projection on the distance field $size/6$ with threshold=0. and 0.95 respectively. c) and d) Grid cells obtained from the conjunction of 30 random landmarks x distance couples with a modulo projection on the distance field $size/10$ with increasing threshold=0. and 0.98
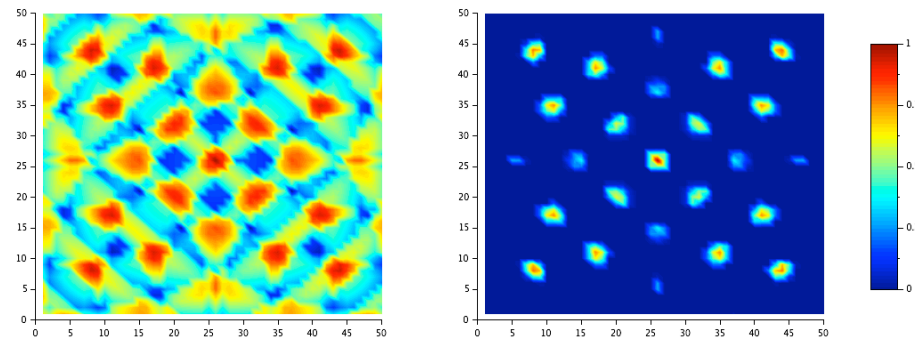
Figure S4: Visual place cells obtained from the conjonction of 4 landmarks x distance couples with a modulo projection on the distance field $size/6$ with threshold=0. and 0.98