A model for the differentiation between grid and conjunctive units in medial entorhinal cortex

A Model for the Differentiation between Grid and Conjunctive Units in Medial Entorhinal Cortex[*]

Bailu Si[1] and Alessandro Treves[2, 3]

[1]*Department of Neurobiology, Weizmann Institute, 234 Herzl St, Rehovot 76100, Israel*[†]
[2]*Sector of Cognitive Neuroscience, International School for Advanced Studies, via Bonomea 265, 34136 Trieste, Italy*
[3]*Kavli Institute for Systems Neuroscience and Center for the Biology of Memory,*
Norwegian University of Science and Technology, 7489 Trondheim, Norway[‡]

The multiple layers of medial entorhinal cortex (mEC) contain cells that differ in selectivity, connectivity, and cellular properties. Grid cells in layer II and in the deeper layers express triangular grid patterns in the environment. The firing rate of the conjunctive cells found in layer III and below, on the other hand, show grid-by-head direction tuning. In this study, we model the differentiation between grid and conjunctive cells in a network with self-organized connections. Arranged into distinct "layers", the model grid units and conjunctive units develop, with a similar time course, grid fields resulting from firing rate adaptation and competitive learning. Grid alignment in both layers is delayed with respect to the formation of triangular grids. A common grid orientation among conjunctive units is produced, in the model, by head-direction modulated collateral interactions, while the grids of grid units inherit the same orientation through connections from conjunctive units. Grid units as well as conjunctive units share a similar spacing but show a random distribution of spatial phases. Grid units however carry more spatial information than conjunctive units, thus providing better inputs for the hippocampus to form spatial memories.

Keywords: Grid cells; Conjunctive grid-by-head direction cells; Firing rate adaptation; Self-organization; Lamination

Introduction

Spatial memory is one of the fundamental functions for an animal to survive successfully in its environment. In mammals, the neural basis of spatial memory has been thought to largely reside in the hippocampus and related cortices. Place cells in the rat hippocampus show elevated firing activity whenever the rat enters a specific portion of the environment, the place field[42]. Head direction (HD) cells in the rat postsubiculum as well as in many other regions are characterized by steady firing when the animal points its head towards a specific direction in the environment[13,45,50,52,53]. In recent years, place-modulated cells were discovered also in the medial entorhinal cortex (mEC), a region just one synapse upstream from the hippocampus, and observed also in the pre- and parasubiculum of rats as well as in the entorhinal cortex of other species [5,26,57], possibly even in humans[18]. The multiple firing fields of a layer II mEC grid cell collectively form a remarkably regular triangular grid spanning the environment which the animal explores[26]. In the deeper layers of mEC, conjunctive grid-by-head-direction cells show firing selectivity to head direction in addition to the same spatial tuning as grid cells[45]. Why do rodents, and bats, and possibly other species, have both grids and conjunctive grid-by-HD units: would one type not suffice? And how do they form (presumably together, as they share the same substrate)?

A series of modeling studies have proposed mechanisms for the expression of grid firing patterns[24,60]. They may be roughly classified into two main categories, namely path-integration models and self-organization models. In path-integration models, these units serve to accumulate the velocity of the animal, in order to track its location internally, either by the collective state of many units in the continuous attractor network sub-variant[8,21,37,41], or by the phases of velocity-controlled oscillators at the single unit level, in the interference sub-variant[9,10,23,27,61]. The triangular grid pattern is imposed ab initio by structured collateral connections or by the summation of multiple oscillators with preferred running directions separated by multiples of 60 degrees. In self-organization models[55], on the other hand, once formed and wired together through recurrent connections, the units may also subserve path-integration, but their spatial responses first emerge spontaneously, at the single unit level. The periodicity of the grid pattern is a result of firing rate adaptation in isotropic exploration[30,48] and is fixated gradually by means of synaptic plasticity in the feedforward connections, which convey broad spatial inputs, for example but not necessarily from "place units"[30]. In a recent variant, grid units receive inputs from periodic "stripe cells"[39], and they inherit the periodicity of these one-dimensional stripe-shaped inputs.

These models describe either grid units or conjunctive units or are compatible with both, and sometimes critically depend on a feature of either cell type[32], but they do not really relate to the striking phenomenon that both cell types are present, nor do they try to explain how their differential properties may emerge.

In layer II, all grid cells found are purely positional. In layer V, the majority of grid cells are conjunctive cells. In layer III and VI, there is a mixture of pure
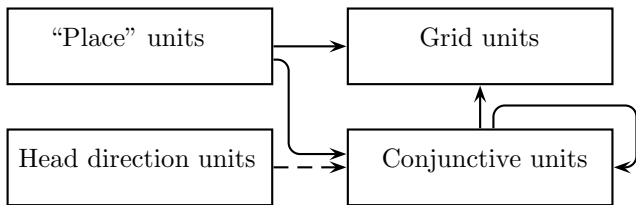
Figure 1: Network model of grid cells and conjunctive cells in mEC. The network is comprised of 320 place units, $N = 256$ conjunctive units and 256 grid units. Place units are fully connected to grid units and to conjunctive units. Each conjunctive unit is modulated by one head direction unit. Conjunctive units are randomly connected to $M = 154$ other conjunctive units (i.e. 60% connectivity, without self connections). The connectivity of the projections from the conjunctive layer to the grid layer is 60%. There are no collateral connections within the grid layer.

grid cells and conjunctive cells[5,45] . This differential laminar localization of grid and conjunctive cells in mEC has not been touched upon by existing models of grid cells, and is a first motivation for this study, following up on previous analyses of cortical lamination[54]. Second, we adapt the self-organizing adaptation model, where unlike our previous study also the recurrent collateral connections, assumed to exist between conjunctive units, self-organize their weights, and we focus particularly on the time course of self-organization, a process burdened with shaping at least three distinct major sets of connections within the same tissue. Note that our model is intended for the development of mEC circuitry during postnatal development, in particular at about P14-P30 in the rat[1]. Third, to better appreciate the contribution of the two cell types, we quantify their self-organized spatial codes with information-theoretic measures, thus providing quantitative support for the functional characterization of the peculiar type of quasi-laminar organization observed in mEC.

Results

Network model of the mEC layers

Grid cells and conjunctive cells are represented by the units in two different layers in the network model (Fig. 1). These layers are abstractions of the real ones, and are meant to capture solely the hypothesis that, given a distinct laminar arrangement in the tissue, a simple genetic instruction, such as "no recurrent connections in layer II", can lead to distinct self-organization trajectories. Both grid units and conjunctive units receive inputs from place units. This assumption is consistent with experimental observations:[6] show that inputs from the hippocampus are necessary for grid cells to maintain their grid firing pattern; during postnatal development, place

cells form adult-like spatial fields earlier than grid cells do[33,56]. The assumption is not strictly necessary, however, as place unit inputs can be replaced with broadly modulated inputs[30]. In the model, conjunctive units are interconnected through collaterals[17], while grid units receive connections from conjunctive units, but have no collaterals among themselves[16], consistent with existing experimental evidence. The connections from the grid layer to the conjunctive layer are not considered in the model, since anatomically connections from superficial to deep layers of mEC appear weaker[25,29].

In the model, conjunctive units develop aligned grids with the help of recurrent collaterals and of the head direction modulation. The grid units develop their own grids, but inherit that common grid orientation through the connections from the conjunctive unit layer to the grid layer.

Self-organization in the conjunctive layer

A virtual rat is simulated to randomly explore a square environment with variable speed, while all the weights in the network are initialized at random values in the beginning, and adapted in the course of the exploration by Hebbian-type rules (ref. *Methods*).

Coherent grids form together with self-organized collateral connections

We first investigate whether conjunctive units can form grids with a coherent orientation while the collateral connections develop. The key difference from our previous model is that here the collateral connections are learned.

In the beginning of the development, the firing maps of the units in the network have multiple fields at irregular locations (Fig.2A, odd rows). Shown below each firing map is the corresponding autocorrelogram, which is the correlation of a firing map to its shifted version in 2-dimensional space. We use the so-called gridness score, which is in the range [-2, 2], to measure the 6-fold spatial periodicity of an autocorrelogram, as in[45]. Conjunctive units develop reasonably regular grids very fast.

It takes about 10^4 equivalent rat seconds (about 2.8 hours, or 10^6 simulation steps) of continuous exploration for conjunctive units to develop good grids (e.g. the unit shown in the first two rows of Fig.2A). Note that each step of the virtual rat corresponds to 10 ms in real time. While grids appear relatively early, grids mutually align to each other at a later stage. For example, unit 93 and unit 111 orient toward 25 degrees after 1.6×10^5 seconds (about 44.4 hours, or 1.6×10^7 simulation steps) of continuous exploration, before finally aligning to a common orientation of about 30 degrees. Unit 149 (last two rows in Fig.2A), however, does not form a good grid even by the end of simulation, rather it diverges to a stripe-like pattern as observed in some of the real mEC cells[31].
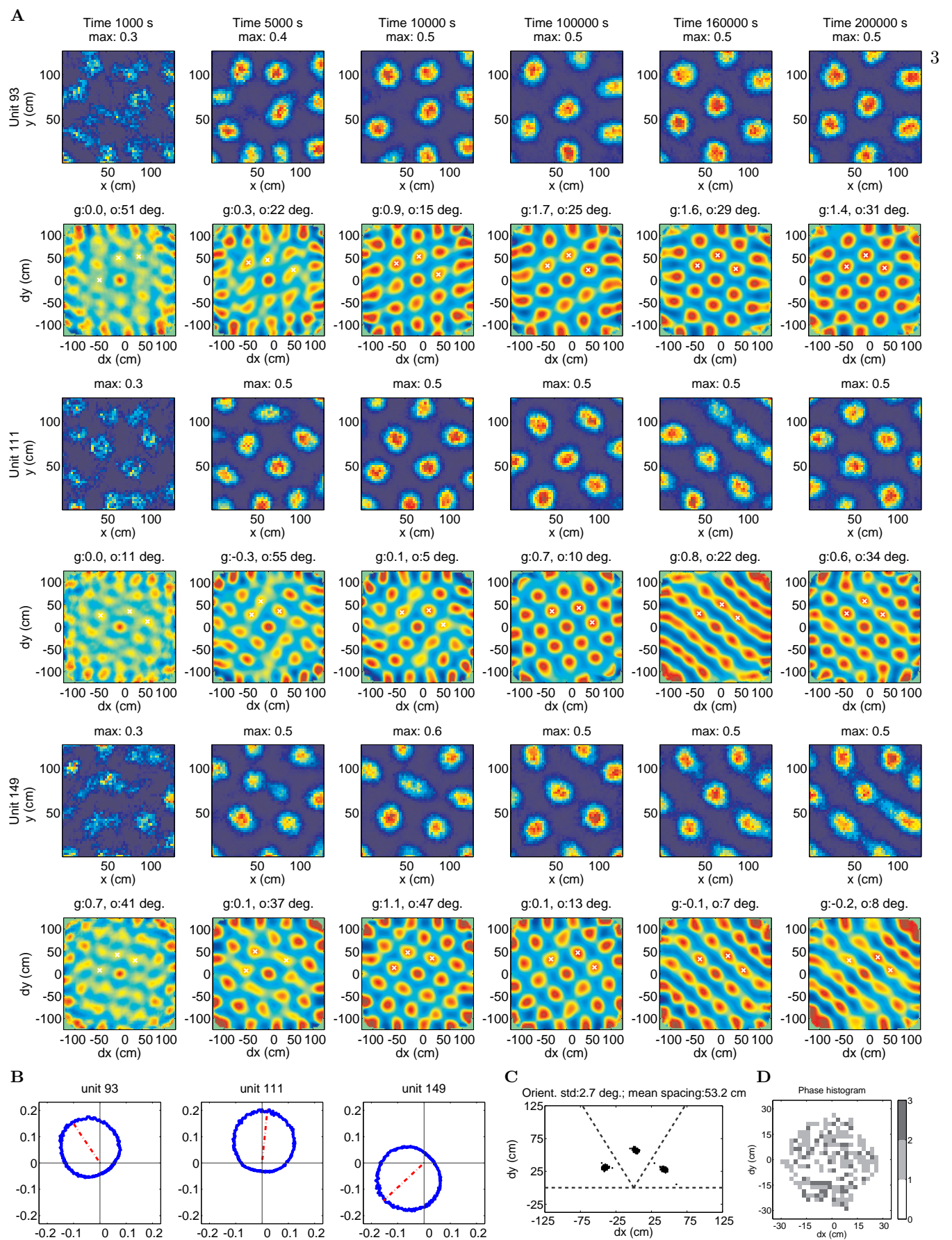
Figure 2: Units in the conjunctive layer develop triangular spatial firing maps over time. (A) The spatial firing rate maps of three example conjunctive units are shown in odd rows, together with their autocorrelograms in even rows, ordered in columns by time. Small values are color-coded by blue, progressing to higher values as red. The maximal firing rate (in arbitrary units) is indicated above each rate map. The gridness score and grid orientation (in degrees) are noted above the corresponding autocorrelogram. The white markers show the three peaks above the x-axis closest to the center of an autocorrelogram, indicating the spacing of a grid as well as the orientations of the three grid axes; (B) Head direction firing rate maps of the three example units in (A) are shown in polar coordinates, with radial coordinate representing firing rate. Dash-dot lines indicate the preferred head directions of the units; (C) Locations of the three peaks found in the autocorrelograms at the end of learning (as shown by the white markers in the right column of (A)) are overlaid for units with gridness score > 0; The average standard deviation in grid axis orientation is 2.1 degrees. The mean spacing is 53.2 cm; (D) Two-dimensional histogram of the spatial phases, relative to the best grid, of the units shown in (C).

Only about 2% of the units (5 out of the 256 conjunctive units) develop stripe-like firing patterns, much less than estimated from experimental data[31]. The formation of a stripe-like pattern is due to firing rate modulation by head direction. Along non-preferred head directions, the firing rate is adapted less since the firing rate is smaller. As plotted in Fg.2B, the firing rates of the units are indeed much higher in their preferred head direction. The stripe-like firing maps show less periodicity along directions roughly perpendicular to the preferred head directions (last two rows in Fig.2A and right panel of Fig.2B).

At the end of the simulation, most conjunctive units in the network develop grids with similar spacing and orientation. The three maxima above the x-axis in the autocorrelograms appear at similar distances and orientations from the origin (Fig.2C). The spatial phases of the grids are randomly distributed (Fig.2D), in accordance with experimental data[26].

The common grid orientation is a result of collateral interactions between units. During development, collateral weights between conjunctive units gain structure through Hebbian learning. At the end of learning, the collateral weights of a conjunctive unit are a function of both the preferred head directions of the presynaptic units and the correlation between the fields of the connected units (Fig.3). The learned collateral weights approach maxima at zero difference between the preferred head directions of the pre- and post-synaptic units, distributing loosely under a Gaussian-like envelope (Fig.3A). The collateral weights also show an increasing trend with respect to the spatial correlation between the maps of the connected units, till moderately high correlation (Fig.3B).

Self-organized collateral weights differ from ad-hoc collateral weights

Although the learned collateral weights have the same function in aligning grids as the ad-hoc collateral weights that we considered in a previous study[48], the learned collateral weights better reflect the correlation between the activity of units. To see the difference, we simulate a network that has already formed grids with fixed ad-hoc collateral weights (ref. *Methods*), and see how the weights change while the network is subject to learning.

Compared to ad-hoc weights, the learned collateral weights fall in a lower range (diamond vs. circle markers in Fig.4A), indicating that the ad-hoc procedure effectively tends to overestimate the larger weights. The learned collateral weights are smoother functions of the preferred head directions of pre-synaptic units, as can be seen from the mean in head direction bins (solid lines in Fig.4A).

While the collateral weights are being learned, conjunctive units develop new maps. Therefore the correlations between the grid maps of unit pairs are remapped. The learned collateral weights coarsely increase only with respect to the new correlation of the corresponding spa-

tial firing maps, but not to the previous one (Fig.4C vs. Fig.4B).

Averaging the collateral weights across conjunctive units reveals that the average learned collateral weights are a sharper function of the difference between the preferred head directions of the pre- and post-synaptic units (Fig.4D). The average learned collateral weights increase with respect to the corresponding mean spatial correlation, but not to the previous correlation (Fig.4F vs. Fig.4E).

Self-organization in the grid layer

How may grid cells in layer II of mEC form coherent triangular grid maps without head direction modulation and without, as recent evidence suggests[16], strong excitatory recurrent connections? It is possible that the common grid orientation results from the excitatory connections from the deeper layers to layer II.

In the model network shown in Fig.1, the spatial responses of the units in the grid layer also evolve into triangular grids as learning proceeds on the connections from place units to grid units. After 10^5 rat equivalent seconds (27.8 hours) of continuous exploration, the three example units in Fig.5A all show fairly good grid (odd rows), as indicated by their median gridness scores (even rows). By about 1.6×10^5 seconds (44.4 hours) of learning, these grids become very triangular and mutually align to about 32 degrees. The activity of the grid units does not show any preference in head direction, due to the absence of head direction modulation (Fig.5B).

The grids of the grid units share the same spacing (52.6 cm), and are aligned to a common orientation (32 degrees), as manifested by the concentrated scattering of the peaks from autocorrelograms (Fig.5C). The grids of the units in the grid layer are aligned to the same orientation as the conjunctive units. The grids of the grid units are also randomly shifted relative to each other, resulting in distributed spatial phases (Fig.5D).

Different from the collateral weights in the conjunctive layer, the self-organized connections from the conjunctive layer to the grid layer do not show a clear pattern dependent on the head direction of the presynaptic units (Fig.6A). Grid units therefore receive balanced inputs from the conjunctive layer, at all possible head directions, producing isotropic firing in head direction. Grid units however are preferentially excited by conjunctive units with similar spatial phases (Fig.6B). The conjunctive units that send strong connections to the same units in the grid layer cluster in space as a result of Hebbian learning.
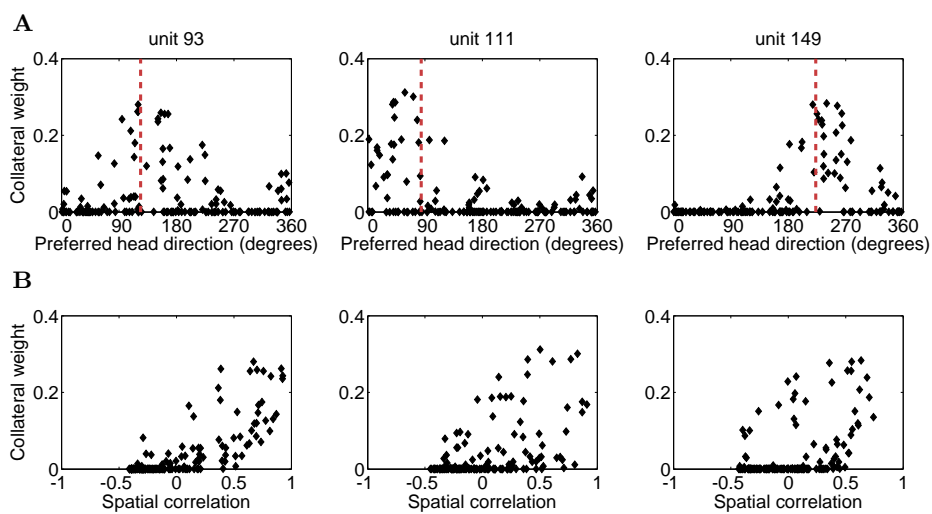
Figure 3: Learned collateral weights between conjunctive units. (A) The collateral weights of three example conjunctive units are plotted against the preferred head directions of pre-synaptic units. In each plot, there are 154 data points, each for one connection. Large weights scatter around the preferred head directions of the post-synaptic units, indicated by the broken lines; (B) The same collateral weights as in (A) are shown with respect to the correlation of the spatial maps between connected units.
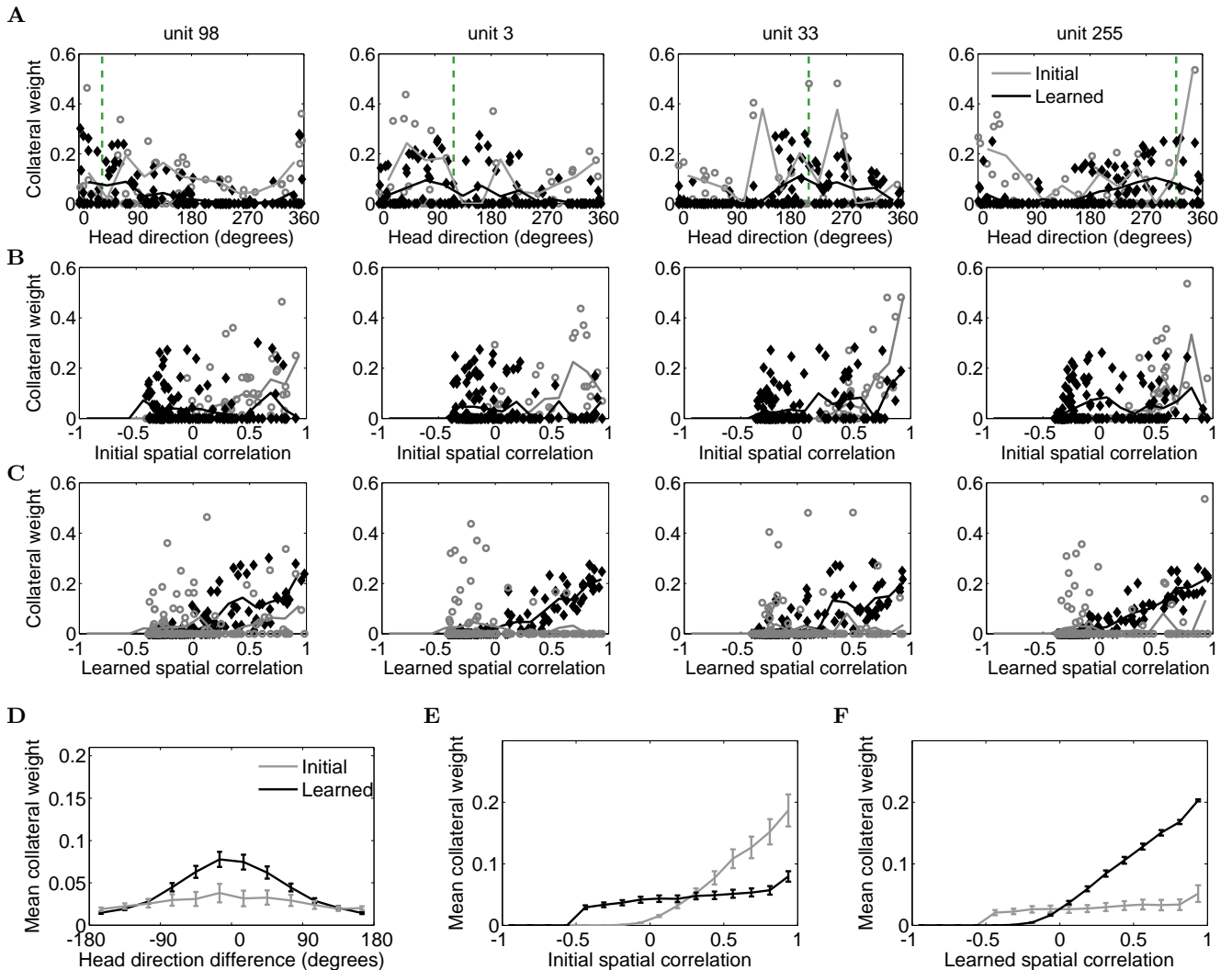
Figure 4: Redistribution of ad hoc collateral weights by the learning rule. (A) The initial ad hoc (gray circles) and the learned collateral weights (dark diamonds) are plotted with respect to the preferred head directions of pre synaptic units. The dotted lines denote the preferred head directions of the post-synaptic units. The average of the weights within head direction bins (bin width 30 degrees) is plotted as a solid line in gray, and in black for ad hoc and learned weights respectively; (B-C) The ad hoc and learned weight vectors of the units in (A) are plotted against the correlation between the spatial firing maps formed with ad hoc weights (B) and with learning weights (C). Solid lines are the average weights in correlation bins (bin width 0.125). Legends are the same as in (A); (D-F) Average ad hoc and learned collateral weights across units are compared as functions of the preferred head directions of pre synaptic units (D), correlation between the spatial firing maps formed with ad hoc weights (E), and correlation between the spatial firing maps formed with learning weights (F). Error bars show $\pm$ standard deviations. The same legend is used in E-F as in D.

Similar time scales for grid development in both layers

Depending on the input from the conjunctive layer, do grid units develop grids later than conjunctive units? We have quantified the gridness, grid alignment and grid spacing of both layers during development.

Averaged across 20 independent trials, the mean gridness scores of the units in both layers increase with a similar time course (Fig.7A). At 4×10^4 seconds (dash-dot line in Fig.7A) , both layers already form grids with

mean gridness scores as high as 0.8. However, mutual grid alignment is reached much later. The standard deviation in grid orientation of the grid units is below 5 degrees at 1.5×10^5 seconds (41.7 hours, Fig.7B) of continuous exploration. The conjunctive layer aligns grids on a similar time scale. For conjunctive units, the standard deviation in grid orientation is below 5 degrees at abut 1.7×10^5 seconds (47.2hours). Considering the time real rats spend in rest and sleeping, the time needed by our model to develop grid maps is comparable with the developmental time required for real rats to express sta-
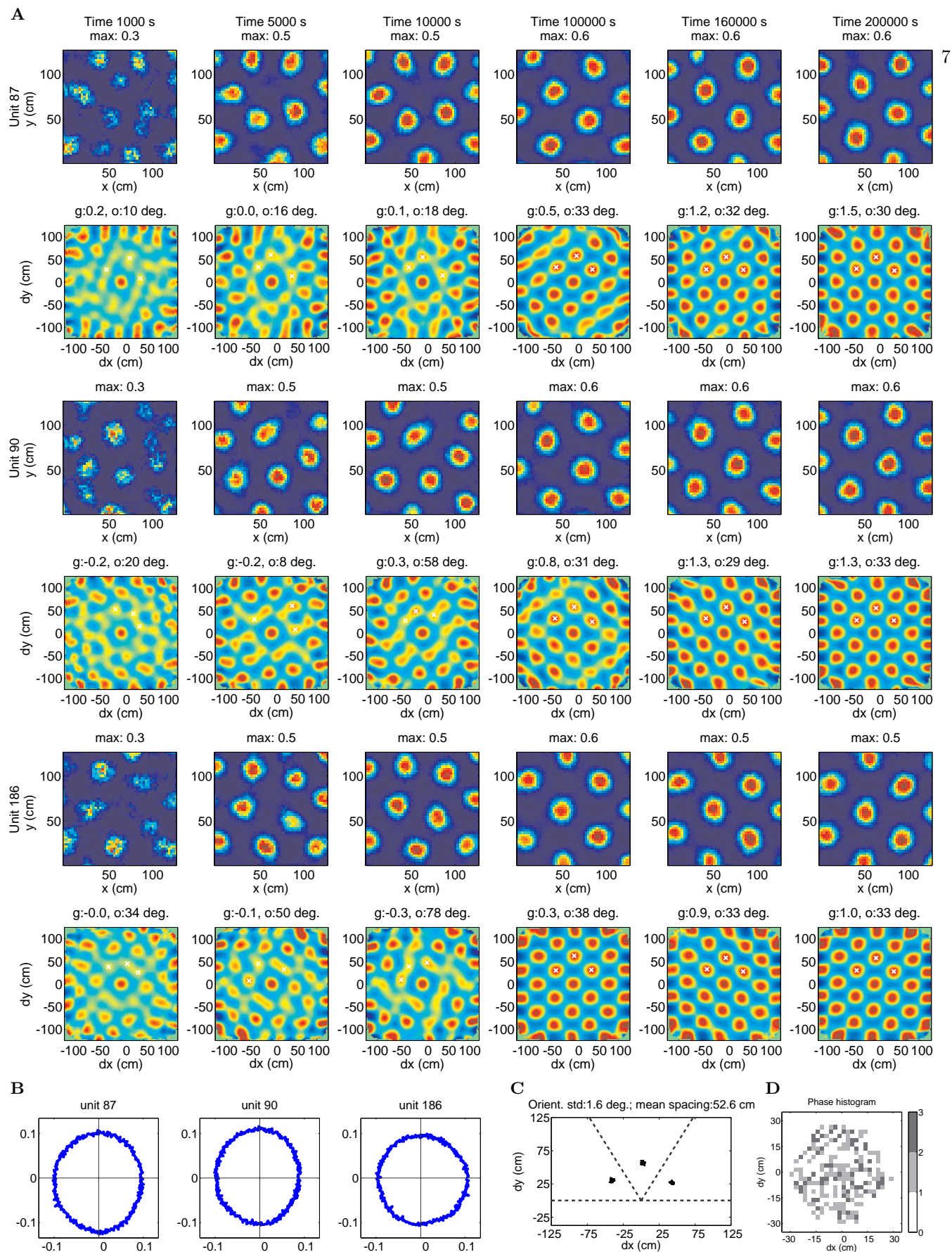
Figure 5: Emergence of grids in the spatial responses of grid units in the network. (A) Evolution of the spatial firing maps and the corresponding autocorrelograms are visualized in odd and even rows, respectively, for the same unit, with respect to developmental time, in columns. A heat color map is used to represent small values in blue and increasing larger values in warmer colors till red. Maximal firing rate (in arbitrary units), gridness score and grid orientation (in degrees) are noted on the top of the panels. The locations of the three peaks above the x-axis closest to the center of an autocorrelogram are marked by white crosses; (B) Polar plots of the head direction firing rate maps of the three example units in (A) are isotropic; (C) At the end of learning, the locations of the three peaks found in the autocorrelograms (white markers in the right column of (A)) are plotted for units with gridness score > 0; The average standard deviation in grid axis orientation is 1.6 degrees. The mean spacing is 52.6 cm; (D) Two-dimensional histogram of the spatial phases, relative to the best grid, of the units shown in (C).
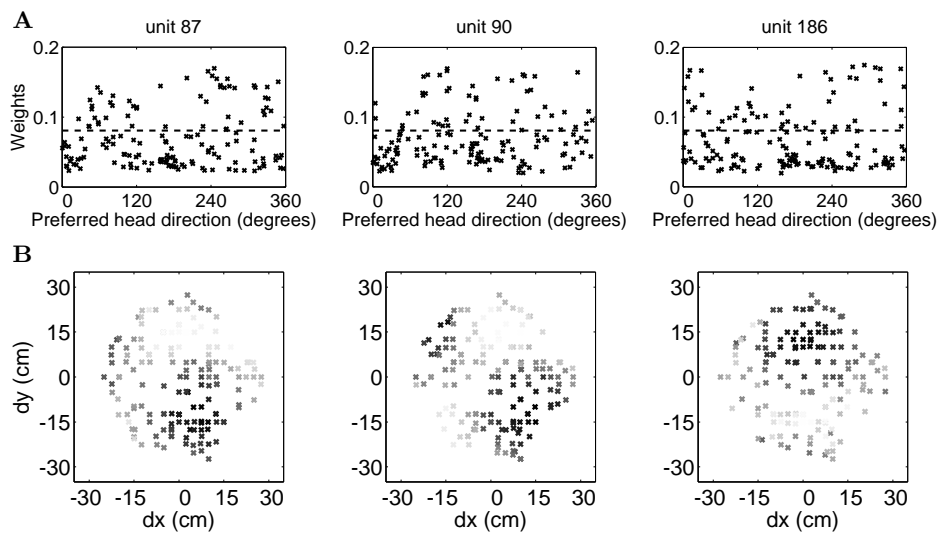
Figure 6: Weights from the conjunctive layer to the grid layer. (A) The inter-layer weights of three example units in the grid layer are plotted with respect to the preferred head directions of the connected conjunctive units. Each marker represents one of the 154 connections. Broken lines show the uniform weight $1/\sqrt{M}$; (B) The inter-layer weights of three example grid units are plotted at the spatial phases of the pre-synaptic units in the conjunctive layer, color-coded by gray for small weights and black for large weights.
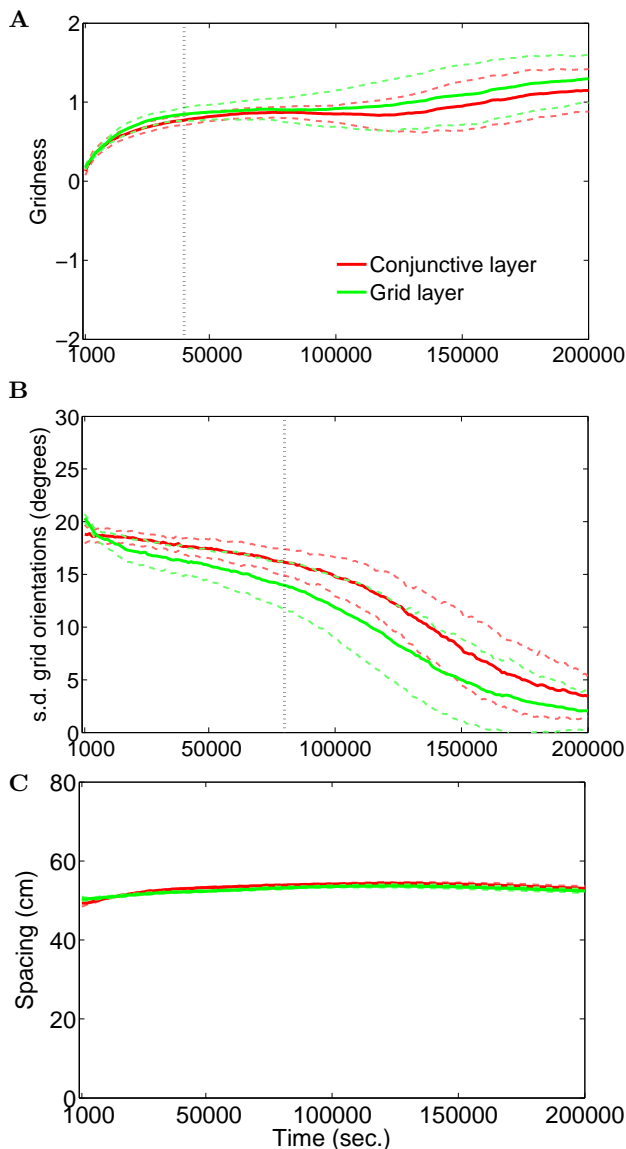
Figure 7: The grid layer and the conjunctive layer develop grids together. We run twenty independent simulations, and plot the cross-trial mean gridness (A), the grid alignment (B) and the spacing (C) of all the units, as solid lines in green for the grid layer, in red for the conjunctive layer. The broken lines show $\pm$ standard deviations. Each panel shares the same legend as in (A). By 4×10^4 seconds (dash-dot line in A), both layers already form grids with mean gridness scores as high as 0.8. The mutual grid alignment in both layers is reached much later, at about 1.7×10^5 seconds, which is a much longer time as compared with the time needed for mutual grid alignment in a network with fixed ad-hoc collateral connections (8×10^4 seconds, dash-dot line in B[48]).

ble grid maps in the environment[1,33,56].

The spacing of the grids in both layers does not change during development (Fig.7C). This is not surprising since the adaptation constants do not change during development.

We have also investigated whether the time scale of grid alignment depends on the learning rates. We have increased by a factor 10 the learning rate for the collateral connections between conjunctive units and the connections from conjunctive units to grid units. The mean gridness scores of 20 independent trials increase with a similar time course as in the reference case with smaller learning rate for collateral and inter-layer connections (Fig.8A top vs. Fig.7A). This is because grid formation depends on the feedforward connections from place units, and the learning rate for feedforward connections is kept the same. The major effect is that the grid alignment in both layers occurs earlier (Fig.8A bottom vs. Fig.7B), though still later than grid formation. By 1.2×10^5 seconds (33.3 hours, Fig.8A bottom) both grid layer and conjunctive layer achieve tight grid alignment.

Another issue is whether the grid layer really needs the direct inputs from the place units, considering that grid units already get inputs from the conjunctive layer. We therefore switched off the connections from place units to grid units. Without direct input from the place units, it takes grid units longer time to develop grid maps (10^5 seconds in Fig.8A top vs. 4×10^4 seconds in Fig.7A), while the grid alignment in the grid layer is not influenced by the removal of the direct input from place units, demonstrating a primary role of the connections from the conjunctive layer in aligning the orientation of grid units. To which extent grid cells in superficial layers of mEC rely on hippocampal inputs should be determined by experiments[6,38].

Information in the spatial codes

In the model, units in both the grid layer and the conjunctive layer form grids with similar gridness, orientation and spacing. Anatomically, layer II and also layer III, where most grid cells reside, differ from layers V and VI, where conjunctive cells are the major cell type, with the strong projections (perforant path) the superficial layers send to the dentate gyrus and hippocampus. Grid cells are therefore presumed to provide accurate spatial information to the dentate gyrus and hippocampus, to form context dependent spatial memory[19,36].

We quantify the information about spatial position each layer conveys, according to a maximal likelihood decoding method (ref. *Methods*). Averaged over trials, the information decoded from the activity of all the units in each layer shows a very different time course. In the grid layer, the information is high from the very beginning of learning (green bars in Fig. 9A). This is because without learning the grid units in the network express random fields due to random initial weights. Random spatial fields are capable of distinguishing locations in the environment as well as the grid fields developed by learning rules. In contrast, the information in the conjunctive layer sharply increases during the early stages of learning (red bars in Fig. 9A). The transition is caused
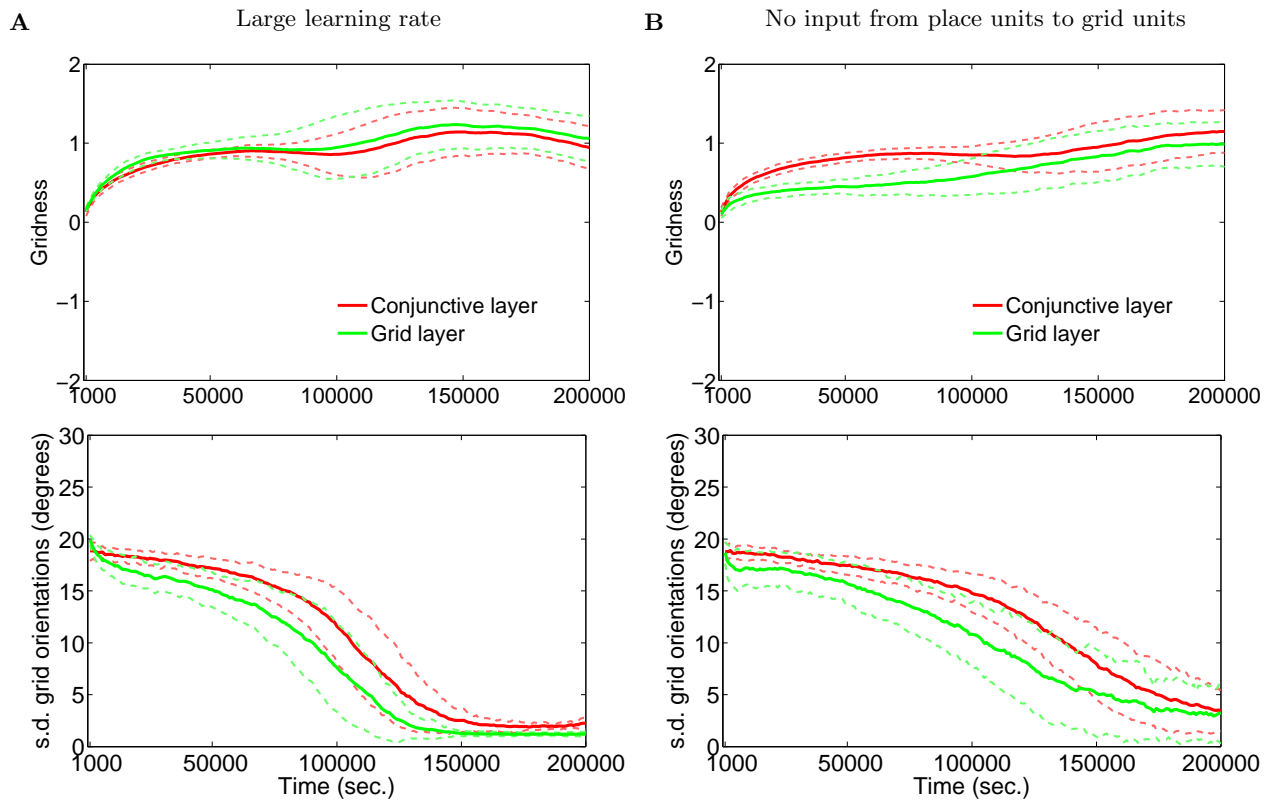
Figure 8: Grid alignment and grid formation are separated. (A) Grid maps are aligned to a common orientation earlier with a larger learning rate, but still well after grid formation ($\zeta = 2 \times 10^4$). Units in both layers develop grid maps with a similar time course as in the reference case (ref. Fig.7A). A coherent orientation of grid maps is developed earlier than in the reference case (bottom panel vs. Fig.7B); (B) Without direct input from place units, it takes longer time for grid units to develop grid maps with similar mean gridness score as in the case with direct place unit inputs (top panel vs. Fig. 7A). A coherent grid orientation is developed in a similar time as in the case with direct inputs (bottom panel vs. Fig. 7B).
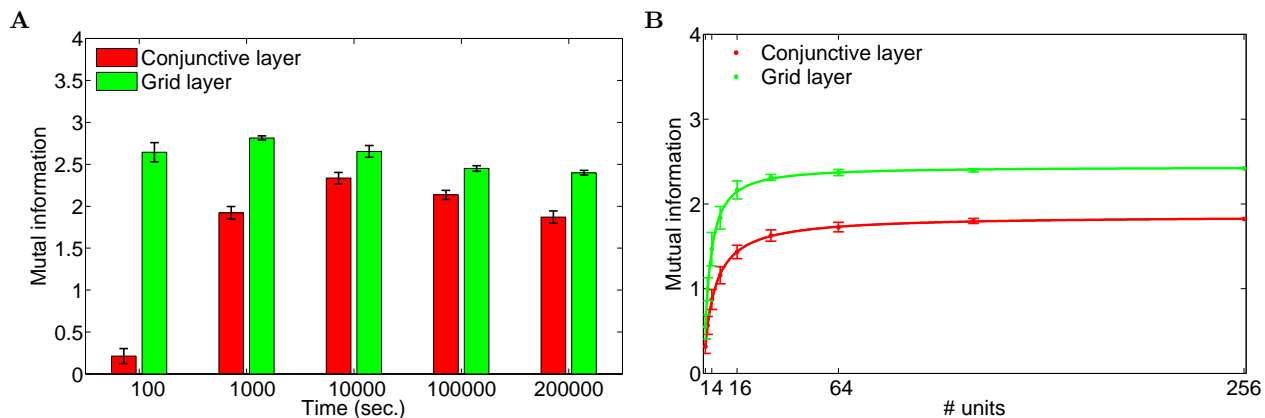

Figure 9: The grid layer carries more spatial information than the conjunctive layer does. (A) The average information, across 5 independent simulations, carried by all the units in the grid layer and in the conjunctive layer are compared for each stage during development. Error bars indicate $\pm$ standard deviations; (B) In one example simulation, the average information across 30 random population samples grows with the size of populations (dots in green for the grid layer and red for the conjunctive layer). Error bars indicate $\pm$ standard deviations. The dependence of the information on population size can be well characterized by sigmoid functions (solid lines, Eq.17) shown here solely at the end of development. At this final stage, the parameters of the fitted sigmoid function are $I_{sat} = 2.44$, $S_0 = 2.83$, $\lambda = 1.15$ for grid units, $I_{sat} = 1.86$, $S_0 = 4.66$, $\lambda = 0.99$ for conjunctive units.

by the development of consistent spatial responses across directions. In early stages of learning, an active conjunctive unit can become silent at directions slightly different from its preferred direction, simply because the head direction modulation brings the input below threshold. During competitive learning, if the connection from a place unit to a conjunctive unit is strengthened, the conjunctive unit receives stronger spatial input. The head-direction tuning then plays a progressively less dominant role in the response of conjunctive cells, and the spatial information conveyed by their activity becomes higher. Note that the information at the end of learning does not reach the maximal possible value of 4 bits, since the number of decoded locations is larger than the number of templates.

The most important difference between the two layers is that the information contained in the grid layer is higher than that in the conjunctive layer during all stages of development (Fig. 9A). This is also the case when fewer units are considered in the decoding (Fig. 9B). For all population sizes, from one unit to 256 units, the information in the grid layer at each time is higher than that in the conjunctive layer (only the information after 2×10^5 seconds of learning is shown). What is common in both layers is the increase in information with respect to population size in decoding. The information carried by grid units grows faster (with larger λ and smaller S_0, see *Methods*) and to higher asymptotic value I_{sat}.

The lower information carried by conjunctive units as compared to grid units is not due to the smaller number of active units in the population, but mainly due to their encoding position only within a narrow range of head directions. As the population size increases, in fact, the information carried by a population of conjunctive units quickly saturates and it does not reach the level obtained with 10 grid units even when all conjunctive units in the network are included in the population (Fig. 9B). In a more general analysis of encoding the conjunction of position and head-direction, conjunctive units may carry higher information than grid units on the condition, roughly speaking, that the base firing-rate (parameter c in Eq.2) be high enough to allow positional coding in non-preferred directions.

Discussion

In this paper, we model the emergence of grid and grid-by-head direction conjunctive responses of the cells in a local module of mEC. In the self-organization network model, mEC units form grid fields due to firing rate adaptation and competitive learning on the feedforward connections from place or anyway spatially modulated units. Both the grid layer and the conjunctive layer develop grids along similar time scales. Relative to the appearance of a periodic grid pattern, the alignment to a common orientation occurs later. The common grid orientation of conjunctive units results from the head-direction modulated collateral interactions, while the grids of grid units inherit that same orientation through the connections from the conjunctive to the grid layer. The spatial phases of the mEC units are randomly distributed, and the spacing is similar, due to the same adaptation time constant. The responses of grid units convey more information about position than those of the conjunctive layer, thus suggesting a computational advantage of differentiating the two populations.

The role of inhibition

In the network model, the homeostatic control of sparsity and mean activity, across units, helps it develop better grids. The biologically plausible implementation of homeostasis will be a direction for future research. Inhibition from interneurons is one possible mechanism for homeostasis in the activity of principal cells in mEC. Recent evidence has shown that inhibitory interneurons play key roles in regulating the response of stellate cells in layer II of mEC[16].

Inhibition between layer II stellate cells, however, has also been hypothesized to concur, through recurrent inhibitory circuitry, to determine the triangular firing maps themselves, a more substantial role than mere homeostasis[16]. Inhibition and adaptation might therefore cooperate in grid formation. In our model, units in the grid layer could well have additional inhibitory recurrent connections, which could be possibly learned through anti-Hebbian learning rules. Our model nevertheless shows that such inhibitory connections are not in principle needed for grid formation, even though the model would obviously still work with them, given an appropriate threshold, since the firing rate of each unit is determined based on the relative strength of its inputs.

The delay in synaptic transmission

The delay τ in synaptic transmission is an important component of the model, in order to develop distributed spatial phases for conjunctive units with similar head direction preference. This delay might be realized by excitatory postsynaptic current mediated by NMDA receptors, which are shown to have slow channel kinetics, lasting several hundred milliseconds[34]. NMDA receptors would then be expected, at least between conjunctive cells, in the deeper layers of mEC. If it turns out that there are no NMDA receptors in synapses between conjunctive cells of mEC, this may point towards alternative models, such as inhibition-based attractor network models.

Laminar organization in mEC

In this study, we only model the differentiation between grid units and conjunctive units. There are many layer-specific properties in mEC[11]. More detailed models are needed to analyze this lamination and the cellular properties it includes, and to uncover their functional relevance.

The model is based on excitatory collateral connections and on head-direction modulation, two elements thought to coexist in the deep layers of mEC[17,45]. We posit the role of the conjunctive cells in determining the observed common orientation of grid cells, given the lack of overt excitatory recurrent connections between layer II grid cells. In[37], instead, conjunctive units inherit the triangular grid pattern produced by the attractor dynamics of grid units, essentially the other way around, and then translate the grid pattern across the neural sheet. To distinguish between these model alternatives, mEC layer II could for example be inactivated, to assess the interaction between conjunctive cells and grid cells.

Entorhino-hippocampal interaction

A substantial amount of data show that entorhinal cortex and hippocampus interact with each other in a loop[6,7,22,59]. Most existing theoretical analyses only study one direction in the loop[2,4,14,20,28,40,44,46,47,49]. New insights on the dynamics of spatial memory may be gained by studying the whole loop of the entorhino-hippocampal system.

Our model addresses some salient aspects of the responses of mEC grid cells in a single environment. Previously we have shown that conjunctive units can self-organize on their own in multiple environments, by assuming a priori recurrent connections[48] and learning feedforward connections from place units, which show global remapping[15]. Learning recurrent and inter-layer connections in multiple environments would work with smaller learning rates than feedforward connections, as is the case in this study, so that the learned recurrent and inter-layer connections would reflect the correlation of the firing patterns across multiple environments. Such connections would facilitate the formation of fixed metric relations between the units, up to a common rotation[22]. Once recurrent and inter-layer connections are set up during postnatal development, they may serve path integration using movement inputs, as suggested by other models[8,21,37,41]. Therefore in adult animals, the responses of grid cells may be a complex result of network dynamics and experience-dependent plasticity[3].

One limitation of the model is that the head-direction tuning of the conjunctive units is imposed multiplicatively, in Eq. 1. The question of how head-direction selectivity can emerge in a learning process will be investigated in the future.

Predictions of the model

Several predictions can be drawn from this study, stimulating new experiments and analyses.

Weaker synaptic plasticity in collateral connections than in afferent connections

The feedforward weights mediate the correlation between environment-specific inputs and the responses of mEC units. The collateral weights mediate the correlation between the responses of conjunctive units. To capture the stable relationship between fields, the time constant for collateral learning should be larger (smaller learning rate) than the time constant for the reorganization of the fields. During postnatal development, synaptic plasticity between conjunctive cells is expected to be weaker than the plasticity of the synapses from place cells (or, equivalently, from spatially modulated cells in postrhinal cortex) to conjunctive cells.

Dependence of grid cells on conjunctive cells

The model suggests that the deep-to-superficial-layer connections play an important role in aligning the grids of the cells in superficial layers. The synaptic plasticity in the inter-layer connections can pick up the most common grid orientation of the units in the conjunctive layer. During postnatal development, inactivating the cells in the deeper layers of mEC, or blocking the inputs from those layers, if feasible, should increase the variability of the grid orientation of the cells in superficial layers, while the periodicity of the grids should not be disrupted.

Difference in the time courses of grid formation and grid alignment

Simulations show that grid alignment at the population level takes a longer time than the emergence of single grids, albeit unstable, in early stages of development. This is because the learning of collateral weights occurs over a longer time scale. The separation of grid formation from grid alignment needs to be assessed by statistical analyses of experimental data.

Dependence of spatial memory on grid cells

Our analysis shows that grid cells convey more spatial information. This could benefit the formation of spatial memories in the hippocampus, by the association of relevant stimuli such as visual, or olfactory ones, with specific places, irrespective of head directions[35]: associative learning may be easier if direction-invariant encoding of space is available. This suggestion is consistent

with the anatomical organization that includes superficial layers of mEC as key inputs to the hippocampus. While many alternative models would also lead to the expectation that hippocampal-dependent spatial memory tasks be impaired when grid cells lose their spatial specificity, our model would predict quantitatively more severe impairment following loss of specificity in the grid cells alone, than in the conjunctive cells alone, should these manipulations be feasible.

Methods

Simulation

A virtual rat is simulated to randomly explore a 125 by 125 cm^2 square environment with variable speed. At each step, the change in the running direction is sampled from a Gaussian distribution with zero mean and angular standard deviation $\sigma_{RD} = 0.2$ radians. The new direction cannot lead the rat outside the limits of the environment, otherwise a new running direction is sampled until a valid direction is chosen. The trajectory of the simulated rat is comprised of epochs with positive or negative acceleration. The lengths of epochs are from a Poisson distribution with 3 steps as the mean, roughly matching available behavioral data[45]. The speed at the end of each epoch is drawn from a truncated Gaussian distribution with mean speed equal to 40 cm/s and standard deviation equal to 16 cm/s. A two-sided truncation is applied to keep the speed non-negative and symmetrically distributed around the mean. The speed within each epoch is linearly interpolated between the speed at the start and at the end of the epoch. The overall appearance of the trajectory is comparable with those of actual rats.

Network model

Each conjunctive unit i is arbitrarily assigned with a preferred head direction θ_i to modulate its inputs. The overall input to a conjunctive unit i at time t is given by h_i^t

$$h_i^t = f_{\theta_i}(\omega_t)(\sum_j W_{ij}^{t-1} r_j^t + \rho_1^t \sum_k W_{ik}^{t-1} \psi_k^{t-\tau}). \quad (1)$$

$\psi_k^{t-\tau}$ is the activity of conjunctive unit k reverberated from collateral connections W_{ik}^t with a delay $\tau = 25$ steps. Each step in the simulation corresponds to 10 ms in real time. $\rho_1^t = 0.1t/T$ is the relative strength of the collateral inputs, with $T = 2 \times 10^7$ being the total number of steps in the simulation. The strength ρ_1^t of collateral connections is gradually increasing from zero in order to reduce the influence of the initial random weights. $f_{\theta_i}(\omega_t)$ is a tuning function that has maximal

value when the current head direction ω^t of the simulated rat is along the preferred direction θ_i, as in[58]

$$f_\theta(\omega) = c + (1 - c) \exp[\nu(\cos(\theta - \omega) - 1)]. \quad (2)$$

$c = 0.1$ and $\nu = 0.8$ are parameters determining the baseline activity and the width of head direction tuning. r_j^t is the firing rate of a "place unit" j relayed by the feedforward connection W_{ij}^t. The activity of a place unit is approximated by an exponential function centered in its preferred firing location $\vec{x}_{j0}$

$$r_j^t = \exp(-\frac{|\vec{x}^t - \vec{x}_{j0}|^2}{2\sigma_p^2}), \quad (3)$$

where $\vec{x}^t$ is the current location of the simulated rat. $|\cdot|$ is the Euclidean norm. $\sigma_p = 5cm$ is the width of the firing field. Place fields are evenly distributed in the environment.

Each grid unit m receives inputs from place units through feedforward connection W_{mj}^t and from conjunctive units through inter-layer connections W_{mi}^t (Fig. 1)

$$h_m^t = \sum_j W_{mj}^{t-1} r_j^t + \rho_2 \sum_i W_{mi}^{t-1} \psi_i^{t-1}. \quad (4)$$

$\rho_2 = 0.1$ is the strength of inter-layer connections.

The firing rate of a conjunctive unit or a grid unit n is determined through a threshold-nonlinear transfer function

$$\psi_n^t = \psi_{sat} \arctan[g^t(\alpha_n^t - \mu^t)]\Theta(\alpha_n^t - \mu^t), \quad (5)$$

where $\psi_{sat} = 2/\pi$ normalizes the firing rate into arbitrary units. $\Theta(\cdot)$ is the Heaviside function. The variable α_n^t represents a time-integration of the input h_n, adapted by the dynamical threshold β_n

$$\alpha_n^t = \alpha_n^{t-1} + b_1(h_n^{t-1} - \beta_n^{t-1} - \alpha_n^{t-1}),$$
$$\beta_n^t = \beta_n^{t-1} + b_2(h_n^{t-1} - \beta_n^{t-1}), \quad (6)$$

where β_n has slower dynamics than α_n (b_2 is set to $b_2 = b_1/3$).

The mean activity $a = \sum_n \psi_n^t/N$ and the sparsity $s = (\sum_n \psi_n^t)^2/(N \sum_n \psi_n^{t\,2})$ in each layer are kept within 10% relative to pre-specified values, $a_0 = 0.1$ and $s_0 = 0.3$ respectively. The biological mechanism of homeostasis within a local network is not the main focus of the current paper. As a technical shortcut, we simply resort to bisection and Newton-Raphson methods to choose an appropriate gain g^t and a threshold μ^t at each step[43]. First, the bracketing and bisection method is applied to find μ^t such that the error function of sparsity $E_s(\mu^t) \equiv s - s_0$ is within 10% error bound of s_0. Then one Newton-Raphson step is taken to get a better root g^t for the error function of mean activity $E_a(g^t) \equiv a - a_0$

$$\Delta g^t = -\frac{E_a(g^t)}{E_a'(g^t)}, \quad (7)$$

where $E'_a(g^t)$ is the derivative of the error function with respect to the gain

$$E'_a(g^t) = \frac{\psi_{sat}}{N} \sum_n \Theta(\alpha_n^t - \mu^t) \frac{\alpha_n^t - \mu^t}{1 + g^{t^2}(\alpha_n^t - \mu^t)^2}. \quad (8)$$

Normally it takes only a few iterations of these two steps to keep both the mean activity and sparsity within 10% relative error bounds.

Self-organization of connection weights

All weights in the network self-organize when the virtual rat explores the environment. The weights from place units to grid units or conjunctive units are changed according to a Hebbian rule

$$\Delta W_{nj}^t = \epsilon(\psi_n^t r_j^t - \bar{\psi}_n^{t-1} \bar{r}_j^{t-1}). \quad (9)$$

Here $\epsilon = 5 \times 10^{-3}$ is a positive learning rate. $\bar{\psi}_n^t$ and $\bar{r}_j^t$ are estimated mean firing rates

$$\begin{aligned} \bar{\psi}_n^t &= \bar{\psi}_n^{t-1} + \eta\left(\psi_n^t - \bar{\psi}_n^{t-1}\right), \\ \bar{r}_j^t &= \bar{r}_j^{t-1} + \eta\left(r_j^t - \bar{r}_j^{t-1}\right). \end{aligned} \quad (10)$$

$\eta = 0.05$ is a positive averaging factor.

The collateral weights between conjunctive units are adapted according to

$$\Delta W_{ik}^t = \zeta \psi_i^t (\psi_k^{t-\tau} - \kappa), \quad (11)$$

and the weights from conjunctive units to grid units are learned by a similar Hebbian rule

$$\Delta W_{mi}^t = \zeta \psi_m^t \psi_i^t. \quad (12)$$

Here $\zeta = 2 \times 10^{-5}$ is a positive learning rate much smaller than the learning rate for feed-forward weights. $\kappa = 0.1$ is an inhibition factor. All the weights in the network are initialized as random numbers $(1-\gamma) + \gamma u$. $\gamma = 0.1$, and u is a random variable uniformly distributed in $[0, 1]$.

After initialization or weight changes, all weights are normalized into unitary L_2 norm

$$\sum_j W_{ij}^{t\,2} = 1. \quad (13)$$

Spatial information in self-organized codes

The spatial information conveyed by the network activity is measured by how much the firing rates can tell about the actual position of the virtual rat. A maximal likelihood decoding procedure based on Euclidian-distance is adopted to predict the position of the rat from the population activity of the network[12]. First, the region of 40 by 40 cm^2 in the center of the environment is divided into 4 by 4 equal-sized spatial bins, denoted by

ξ. The size of the region is chosen to fit into the period of the grids. Note that the spacing of the grids formed in the network is about 53 cm (Fig. 2 and 5). During a random exploration of 2×10^6 steps, the population activities of the grid layer and the conjunctive layer are stored as templates $\vec{\Psi}^{temp}(\xi)$ for each spatial bin when the virtual rat runs closest to the center of each spatial bin. Then the virtual rat explores for another 2×10^6 steps. When the rat runs into one of the bins ξ^t at time t, a probability distribution over the bins is constructed

$$P(\xi|\vec{\Psi}(\xi^t)) = \frac{|\vec{\Psi}(\xi^t) - \vec{\Psi}^{temp}(\xi)|}{\sum_{\xi'} |\vec{\Psi}(\xi^t) - \vec{\Psi}^{temp}(\xi')|}, \quad (14)$$

where $|\cdot|$ is the Euclidean norm of a vector.

The maximum likelihood prediction of the actual spatial bin, which the rat occupies, is

$$\hat{\xi} = \arg\max_\xi P(\xi|\vec{\Psi}(\xi^t)). \quad (15)$$

The mutual information between the actual bin ξ and the decoded bin $\hat{\xi}$ can be calculated by

$$I = \sum_{\xi,\hat{\xi}} P(\xi, \hat{\xi}) \log_2 \frac{P(\xi, \hat{\xi})}{P(\xi)P(\hat{\xi})}, \quad (16)$$

where $P(\xi)$ and $P(\hat{\xi})$ are the marginal distributions considering all the decoding steps, and $P(\xi, \hat{\xi})$ is the joint distribution of the actual and decoded bins.

The extracted information I is a measure of the spatial information carried by the units in the network. The information content depends on the population size considered in the decoding step (Eq.14-15). We vary the population size from 1,2,. . . to 256 units in geometric progression. For population size smaller than 256, 30 different random populations are sampled for decoding to get the average information.

We fit a sigmoid function with three parameters to the information carried by a population of units[51]

$$I(S) = \frac{I_{sat}}{1 + (S_0/S)^\lambda}, \quad (17)$$

where S is the number of units in the population. Parameter $I_{sat} > 0$ is the asymptotic information when S goes to infinity. $S_0 \geq 1$ is the number of units needed to conveys $I_{sat}/2$ bits information. $\lambda > 0$ is a power factor controlling the increase rate of the information. The derivative of the information with respect to population size is maximal at $S = 1$, since

$$I'(S) = \frac{\lambda I_{sat}}{S[(\frac{S_0}{S})^{\lambda/2} + (\frac{S_0}{S})^{-\lambda/2}]^2} \leq \frac{\lambda I_{sat}}{4S} \leq \frac{\lambda I_{sat}}{4}. \quad (18)$$

Ad hoc collateral weights in the conjunctive layer

To see the difference between the self-organized and the ad hoc collateral weights in a previous study[48], we also perform simulations in which the learning of collateral weights (Eq.11) kicks in only after the units in the network forms grid fields with fixed ad hoc collaterals and adaptable feedforward connections.

The ad hoc collateral weights try to capture the correlation of the conjunctive responses of the units when the rat passes through their place fields. Temporarily assigning to conjunctive unit i a random auxiliary place field at position (x_i, y_i) in the environment, the ad hoc weights from conjunctive unit k to i is defined by

$$W_{ik} = [f_{\theta_k}(\omega_{ki})f_{\theta_i}(\omega_{ki})\exp(-\frac{d_{ki}^2}{2\sigma_f^2}) - \kappa]^+, \qquad (19)$$

where $[\cdot]^+$ is a threshold function, with $[x]^+ = 0$ for $x < 0$, and $[x]^+ = x$ otherwise. κ is the inhibition parameter defined in Eq.11. $f_\theta(\omega)$ is the head direction tuning function defined in Eq. 2. ω_{ki} is the direction of the trajectory from field k to i. $d_{ki} = \sqrt{[x_i - (x_k + \ell\cos\omega_{ki})]^2 + [y_i - (y_k + \ell\sin\omega_{ki})]^2}$ is the Euclidean distance between field k and i with offset $\ell = 10cm$. ℓ is chosen to equal to the distance the simulated rat covers with its average speed 40cm/s during τ steps. $\sigma_f = 10cm$ is the width of spatial tuning. Normalization as in Eq.13 is also applied to W_{ik}. Note that once the ad hoc weights are assigned, their auxiliary fields are removed from the simulation.

[†] Electronic address: `bailu.si@weizmann.ac.il`

[‡] Electronic address: `ale@sissa.it`

[1] Ainge J, Langston R (2012) Ontogeny of neural circuits underlying spatial memory in the rat. Front Neural Circuits 6

[2] de Almeida L, Idiart M, Lisman J (2009) The input-output transformation of the hippocampal granule cells: from grid cells to place fields. J Neurosci 29(23):7504–12

[3] Barry C, Ginzberg LL, O'Keeefe J, Burgess N (2012) Grid cell firing patterns signal environmental novelty by expansion. Proceedings of the National Academy of Sciences 109(43):17,687–17,692

[4] Blair H, Gupta K, Zhang K (2008) Conversion of a phase- to a rate-coded position signal by a three-stage model of theta cells, grid cells, and place cells. Hippocampus 18(12):1239–55

[5] Boccara CN, Sargolini F, Thoresen VHH, Solstad T, Witter MP, Moser EI, Moser MBB (2010) Grid cells in pre- and parasubiculum. Nature neuroscience 13(8):987–994

[6] Bonnevie T, Dunn B, Fyhn M, Hafting T, Derdikman D, Kubie JL, Roudi Y, Moser EI, Moser MB (2013) Grid cells require excitatory drive from the hippocampus. Nat Neurosci 16(3):309–317

[7] Brun VH, Otnass MK, Molden S, Steffenach HAA, Witter MP, Moser MBB, Moser EI (2002) Place cells and place recognition maintained by direct entorhinal-hippocampal circuitry. Science 296(5576):2243–2246

[8] Burak Y, Fiete IR (2009) Accurate path integration in continuous attractor network models of grid cells. PLoS Comput Biol 5(2):e1000,291

[9] Burgess N (2008) Grid cells and theta as oscillatory interference: theory and predictions. Hippocampus 18(12):1157–1174

[10] Burgess N, Barry C, O'Keefe J (2007) An oscillatory interference model of grid cell firing. Hippocampus 17(9):901–812

[11] Canto CB, Witter MP (2011) Cellular properties of principal neurons in the rat entorhinal cortex. II. the medial entorhinal cortex. Hippocampus 22:1277–1299

[12] Cerasti E, Treves A (2010) How informative are spatial ca3 representations established by the dentate gyrus? PLoS Comput Biol 6(4):e1000,759

[13] Chen L, Lin L, Green E, Barnes C, McNaughton B (1994) Head-direction cells in the rat posterior cortex. i. anatomical distribution and behavioral modulation. Exp Brain Res 101(1):8–23

[14] Cheng S, Frank LM (2011) The structure of networks that produce the transformation from grid cells to place cells. Neuroscience 197(1):293–306

[15] Colgin LL, Moser EI, Moser MB (2008) Understanding memory through hippocampal remapping. Trends in Neurosciences 31(9):469–477

[16] Couey JJ, Witoelar A, Zhang SJ, Zheng K, Ye J, Dunn B, Czajkowski R, Moser MB, Moser EI, Roudi Y, Witter MP (2013) Recurrent inhibitory circuitry as a mechanism for grid formation. Nat Neurosci 16(3):318–324

[17] Dhillon A, Jones RSG (2000) Laminar differences in recurrent excitatory transmission in the rat entorhinal cortex in vitro. Neuroscience 99(3):413–422

[18] Doeller C, Barry C, Burgess N (2010) Evidence for grid cells in a human memory network. Nature 463(7281):657–61

[19] Fiete IR, Burak Y, Brookings T (2008) What grid cells convey about rat location. The Journal of neuroscience 28(27):6858–6871

[20] Franzius M, Vollgraf R, Wiskott L (2007) From grids to places. J Comput Neurosci 22:297–299

[21] Fuhs M, Touretzky D (2006) A spin glass model of path integration in rat medial entorhinal cortex. J Neurosci 26:4266–76

[22] Fyhn M, Hafting T, Treves A, Moser MB, Moser E (2007) Hippocampal remapping and grid realignment in entorhinal cortex. Nature 446:190–94

[23] Giocomo LM, Zilli EA, Fransén E, Hasselmo ME (2007) Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing. Science 315(5819):1719–

1722

24. Giocomo LM, Moser MB, Moser EI (2011) Computational Models of Grid Cells. Neuron 71(4):589–603

25. van Haeften T, Baks-Te-Bulte L, Goede PH, Wouterlood FG, Witter MP (2003) Morphological and numerical analysis of synaptic interactions between neurons in deep and superficial layers of the entorhinal cortex of the rat. Hippocampus 13(8):943–952

26. Hafting T, Fyhn M, Molden S, Moser MBB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. Nature 436(7052):801–806

27. Hasselmo ME (2008) Grid cell mechanisms and function: contributions of entorhinal persistent spiking and phase resetting. Hippocampus 18(12):1213–1229

28. Hayman R, Jeffery K (2008) How heterogeneous place cell responding arises from homogeneous grids – a contextual gating hypothesis. Hippocampus 18:1301–1313

29. Köhler C (1986) Intrinsic projections of the retrohippocampal region in the rat brain. III. The medial entorhinal area. J Comp Neurol 246(4):149–169

30. Kropff E, Treves A (2008) The emergence of grid cells: Intelligent design or just adaptation? Hippocampus 18:1256–1269

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

32. Kubie JL, Fenton A (2012) Linear look-ahead in conjunctive cells: An entorhinal mechanism for vector-based navigation. Frontiers in Neural Circuits 6(20)

33. Langston RF, Ainge JA, Couey JJ, Canto CB, Bjerknes TL, Witter MP, Moser EI, Moser MBB (2010) Development of the spatial representation system in the rat. Science 328(5985):1576–1580

34. Lester R, Clements J, Westbrook G, Jahr C (1990) Channel kinetics determine the time course of nmda receptor-mediated synaptic currents. Nature 346(6284):565–567

35. Maaswinkel H, Whishaw IQ (1999) Homing with locale, taxon, and dead reckoning strategies by foraging rats: sensory hierarchy in spatial navigation. Behavioural Brain Research 99(2):143 – 152

36. Mathis A, Herz A, Stemmler M (2012) Optimal population codes for space: grid cells outperform place cells. Neural Comput 24(9):2280–317

37. McNaughton B, Battaglia F, Jensen O, Moser E, Moser MB (2006) Path integration and the neural basis of the "cognitive map". Nature Reviews Neurosci 7:663–78

38. Melzer S, Michael M, Caputi A, Eliava M, Fuchs E, Whittington M, Monyer H (2012) Long-range-projecting gabaergic neurons modulate inhibition in hippocampus and entorhinal cortex. Science 335(6075):1506–10

39. Mhatre H, Gorchetchnikov A, Grossberg S (2010) Grid cell hexagonal patterns formed by fast self-organized learning within entorhinal cortex. Hippocampus DOI 10.1002/hipo.20939

40. Molter C, Yamaguchi Y (2008) Entorhinal theta phase precession sculpts dentate gyrus place fields. Hippocampus 18:919–930

41. Navratilova Z, Giocomo LM, Fellous JMM, Hasselmo ME, McNaughton BL (2011) Phase precession and variable spatial scaling in a periodic attractor map model of medial entorhinal grid cells with realistic after-spike dynamics. Hippocampus DOI 10.1002/hipo.20939

42. O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain research 34(1):171–175

43. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical Recipes 3rd Edition: The Art of Scientific Computing, 3rd edn. Cambridge University Press, New York, NY, USA

44. Rolls E, Stringer S, Elliot T (2006) Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning. Network: Comput Neural Sys 15:447–465

45. Sargolini F, Fyhn M, Hafting T, Mcnaughton BL, Witter MP, Moser MB, Moser EI (2006) Conjunctive Representation of Position, Direction, and Velocity in Entorhinal Cortex. Science 312(5774):758–762

46. Savelli F, Knierim JJ (2010) Hebbian analysis of the transformation of medial entorhinal grid-cell inputs to hippocampal place fields. J Neurophysiol 103(6):3167–83

47. Si B, Treves A (2009) The role of competitive learning in the generation of DG fields from EC inputs. Cognitive Neurodynamics 3(2):177–187

48. Si B, Kropff E, Treves A (2012) Grid alignment in entorhinal cortex. Biological Cybernetics 106(8-9):483–506

49. Solstad T, Moser E, Einevoll G (2006) From grid cells to place cells: A mathematical model. Hippocampus 16:1026–1031

50. Stackman R, Taube J (1998) Firing properties of rat lateral mammillary single units: head direction, head pitch, and angular head velocity. J Neurosci 18(21):9020–37

51. Stella F, Cerasti E, Treves A (2013) Unveiling the metric structure of internal representations of space. Frontiers in Neural Circuits 7(81)

52. Taube J, Burton H (1995) Head direction cell activity monitored in a novel environment and during a cue conflict situation. J Neurophysiol 74(5):1953–71

53. Taube JS, Muller RU, Ranck JB (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. The Journal of neuroscience 10(2):420–435

54. Treves A (2003) Computational constraints that may have favoured the lamination of sensory cortex. Journal of Computational Neuroscience 14:271–282

55. Treves A, Kropff E, Biswas A (2005) On the triangular grid of entorhinal place fields. Society for Neuroscience abstract 198.11

56. Wills TJ, Cacucci F, Burgess N, O'Keefe J (2010) Development of the hippocampal cognitive map in preweanling rats. Science 328(5985):1573–1576

57. Yartsev MM, Witter MP, Ulanovsky N (2011) Grid cells without theta oscillations in the entorhinal cortex of bats. Nature 479(7371):103–107

58. Zhang K (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. The Journal of neuroscience 16(6):2112–2126

59. Zhang SJ, Ye J, Miao C, Tsao A, Cerniauskas I, Ledergerber D, Moser MB, Moser EI (2013) Optogenetic dissection of Entorhinal-Hippocampal functional connectivity. Science 340(6128)

60. Zilli E (2012) Models of grid cell spatial firing published 2005-2011. Frontiers in neural circuits 6

61. Zilli EA, Hasselmo ME (2010) Coupled noisy spiking neurons as velocity-controlled oscillators in a model of grid cell spatial firing. The Journal of neuroscience 30(41):13,850–13,860