Self-organization of Hippocampal Representations in Large Environments

Shuang Liu[1,2]
[1]University of Chinese Academy of Sciences
Beijing, China
Email: liushuang1@sia.cn

Bailu Si[2] and Yang Lin[2]
[2]Shenyang Institute of Automation
Chinese Academy of Sciences
Shenyang, China
Email: {sibailu,liny}@sia.cn

***Abstract*—The entorhino-hippocampal neural circuits of the mammalian brains are able to internally generate efficient spatial representations of large-scale environments. Hippocampal principal cells in mammalian brains form sparse codes of the positions in large environments. The underlying computational mechanisms of the formation of spatial memory in large-scale environments are still not well understood. We present a competitive learning network that integrates both spatial inputs from grid cells as well as contextual inputs from non-spatial cells. Through self-organisation, place units in the network form place fields at "randomly" interleaved locations, in the sense that both the recruitment process and the inter-field interval are memoryless, a property superior for spatial memory formation in large environments.**

I. Introduction

The ability to build map representations of and self-localize in large and uncertain environments is critical for a robot to accomplish spatial tasks [1]. Many robot navigation systems rely on high-precision sensors to build maps, yet still are not able to function in changing environments or in rainy or snowy weather conditions [2]. The map representations used in these systems are easy to be interpreted and used by humans, but scale linearly with the area of the explored environment. In large environments, both the memory to maintain and the time to update these maps become a great burden, and would require a huge database as well as computing and networking resources.

As a contrast, animals, such as birds or bats [3, 4], are able to efficiently navigate in large environments under natural conditions in the scale of hundreds or thousands of kilometers. For mammals, neuroscience studies pointed to the hippocampus and the entorhinal cortex as an internal GPS system in the brain. Grid cells in the medial entorhinal cortex (MEC) provide multi-scale periodic spatial codes [5], that efficiently represent distances in large-scale environments [6]. Many principal cells in the hippocampi of the mammalian brains fire whenever the animal passes specific locations of the environment [7], and therefore are called place cells, since each of such cells stably remembers certain locations, i.e. the place fields of the cells. As a whole, the place cell assemblies in the hippocampus constitute a cognitive map of the environment. One place cell is recruited into the cognitive map representation independent of its previous firing history either in the same environment [8] or in previously encountered environments [9]. Each place cell can be reused to encode multiple locations in the environment. Statistics reveal that the process of place field formation is memoryless [8]. Hippocampus place cells therefore provide sparse spatial map representations that scale sub-linearly in the area of the environment.

The formation of place fields has been studied theoretically in terms of the spatial inputs from the MEC [10, 11, 12, 13, 14], which sends strong projections to the hippocampus. The synaptic plasticity plays an important role in the self-organisation of place codes in the hippocampus [11, 15], and results in robust place representations [16]. The learning of the connections between the MEC and the hippocampus results in general network structures that are independent of the details of specific learning rules [17]. However, these modelling works are all done in small environments of size no more than several m^2. It is not clear whether these models are able to form spatial representations of large environments with statistical structures similar as recently observed experiments [8].

In our previous study, we suggested that place cells integrate both spatial inputs and non-spatial context inputs from the entorhinal cortex [15], a hypothesis consistent with experimental results showing that place cells indeed receive strong projections from non-spatial cells as well as from grid cells in the entorhinal cortex [18]. In this work, we investigate the nature of the synaptic plasticity of place cells, and reveal the network structures resulting from such a learning process. We show that our model is able to form spatial map representations of large environments with a memoryless recruitment process. The model is introduced in Section II and analysed theoretically in Section III, followed by the simulation results in Section IV. The paper is concluded in Section V with discussions.

II. Network Model of Place Cells

In the network, there are $M = 1000$ place units, the input of which is from two kinds of afferent activity (Fig. 1)

$$h_i(t) = \sum_{j=1}^{N} w_{ij}(t-1)\psi_j(t) + c_i. \quad (1)$$

c_i is the nonspatial input, which is simply modelled by random activities from a Normal distribution $\mathcal{N}(0, \sigma^2)$ [15], with the details left for future study. Note that c_is are constant once the network is given. N is the total number of grid cells connected to each place cell. $\psi_j(t)$ is the firing rate of grid cell j, which projects to place unit i with adaptable synapse w_{ij}.

Grid cells are organised into L modules [5], with 1000 cells in each. Each place cell receives inputs from random 5%

978-1-4799-1959-8/15/$31.00

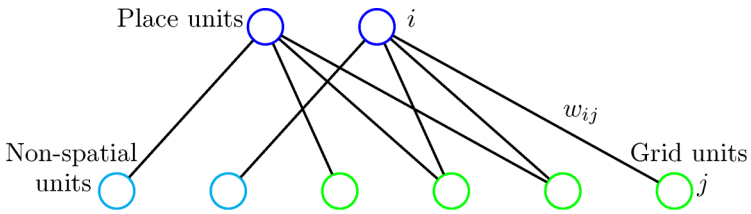

Fig. 1: Network model of place cells.

grid cells of each module, constituting the total number of connected grid cells $N = 50L$. Adding more number of grid cells into modules or add more connections to each place cell leads to similar results [15].

For simplicity, we only consider the firing rate of a grid cell in 1D environments

$$\psi_j(t) = \frac{1}{2} + \frac{1}{2}\cos(\frac{2\pi}{S_l}(x(t) - x_j)), \tag{2}$$

where $S_l = \lambda\rho^{l-1}$ $(l = 1, 2, \cdots, L)$ is the spacing between two firing fields of the cells in module l. $\lambda > 0$ is the minimal spacing. $\rho > 1$ is the ratio between two neighbouring modules. $x_j \in [0, S_l]$ is the spatial phase of grid cell j. $x(t)$ is the position of the animal at time t.

The firing rate of a place cell is determined by

$$r_i(t) = g(t)[h_i(t) - \mu(t)]_+, \tag{3}$$

where the threshold-linear transfer function $[x]_+ = 0$ for $x < 0$, and $[x]_+ = x$ otherwise. $\mu(t)$ is a threshold parameter and $g(t)$ is a gain factor, determined at each time step to make sure that both the sparsity $\frac{(\sum_i r_i(t)/N)^2}{\sum_i r_i^2(t)/N}$ and the mean firing rate of place cell assemblies is equal to the desired value $m = 0.1$. The parameters $\mu(t)$ and $g(t)$ mimic the overall inhibition on the principal cells from the interneurons in the hippocampus.

The connections from grid cells to place cells are learned according to a Hebbian rule

$$\tilde{w}_{ij}(t) = [w_{ij}(t-1) + \eta r_i(t)(\psi_j(t) - \kappa)]_+, \tag{4}$$

$$w_{ij}(t) = \frac{\tilde{w}_{ij}(t)}{\sqrt{\sum_j \tilde{w}_{ij}^2(t)}}. \tag{5}$$

Here η = 0.001 is a learning rate. $\kappa = 0.8$ is an inhibition parameter. The initial weight $w_{ij}(0)$ is set to $(1 - \gamma) + \gamma\epsilon$, where $\gamma = 0.9$, and ϵ is a random number uniformly sampled from $[0, 1]$.

III. Computational mechanisms underlying place field formation

The essence of Eq. 1 and Eq. 3 is that grid cells that are active together would provide input strong enough to drive a place cell to fire at the same location. The spatial information of the place cell mainly comes from the coactive grid cells, which are considered to maintain states of allocentric locations while the animal performs path-integration based navigation[19, 20, 21, 22]. The learning mechanism of Eq.4 would in turn increase the strength of the connections from these coactive grid cells to the downstream place cell. As a result, the place cell selects and wires groups of coactive grid cells to provide spatial inputs. Consider two coactive grid cells, their spatial phases satisfy the following condition

$$x_j + AS_{l_j} = x_k + BS_{l_k} = p_i, \tag{6}$$

where A and B are some non-negative integers. S_{l_j} is the spacing of cell j from module l_j. p_i is a firing location of the place unit i to which the cell j and k project. Eq.6 can be simplified into

$$x_j - x_k = BS_{l_k} - AS_{l_j} \equiv D_i, \tag{7}$$

meaning that the firing location p_i of place unit i is determined by the phase difference of the strong projecting grid cells. In principal, each place cell can have many independent groups of such grid cells, because grid cells with different spatial phases but the same phase difference are valid solutions of Eq. 6 as well.

In summary, the place fields of a place cell are encoded by (and is a readout of) the firing position of coactive grid cells that follow certain phase differences. These coactive grid cells are selected to provide spatial information to a place cell by Hebbian learning mechanisms. Since the spatial phases of grid cells are randomly distribution in the environments, the resulting place fields would show random and memoryless distribution. The non-spatial inputs to the place cells provide additional modulation on place field formation, resulting in diversified propensity across cells to form different number of place fields. The hippocampal place cell activities thus self-organise into sparse and 'random' codes, modulated by multi-modal non-spatial sensory inputs, and function as neural substrates for efficient navigation in large-scale environments.

IV. Simulation results

To verify the computational mechanisms of the model, we simulate the network and compare the results to experimental data [8]. In simulations, a virtual rat runs back and forth in a linear track of length 360 cm. Networks with the same parameter $\lambda = 32$, but different structural parameters L, ρ and σ are simulated. After the place units form stable firing fields due to the self-organisation process described in Eqs. 4-5, we analyse the statistics of the firing fields from all active cells.

A. The phase relation between grids determines the location of a place field

In a network without non-spatial inputs and with only two grid modules, the firing fields of each place unit repeat (Fig. 2A top row). This is because summing periodic firing patterns of grid cells leads to periodic firing fields of the output place units in the network. The period of the resulting place cell firing is equal to the least common multiple of the spacings of the two grid scales, i.e. 96 cm. The size of the environment is 360 cm, about 4 times as large as the least common multiple. As a result, four repeating firing fields are expressed in the environments for the active cells.

As a result of Hebbian learning, the grid cells that project strongly to a place cell after learning show clustered spatial

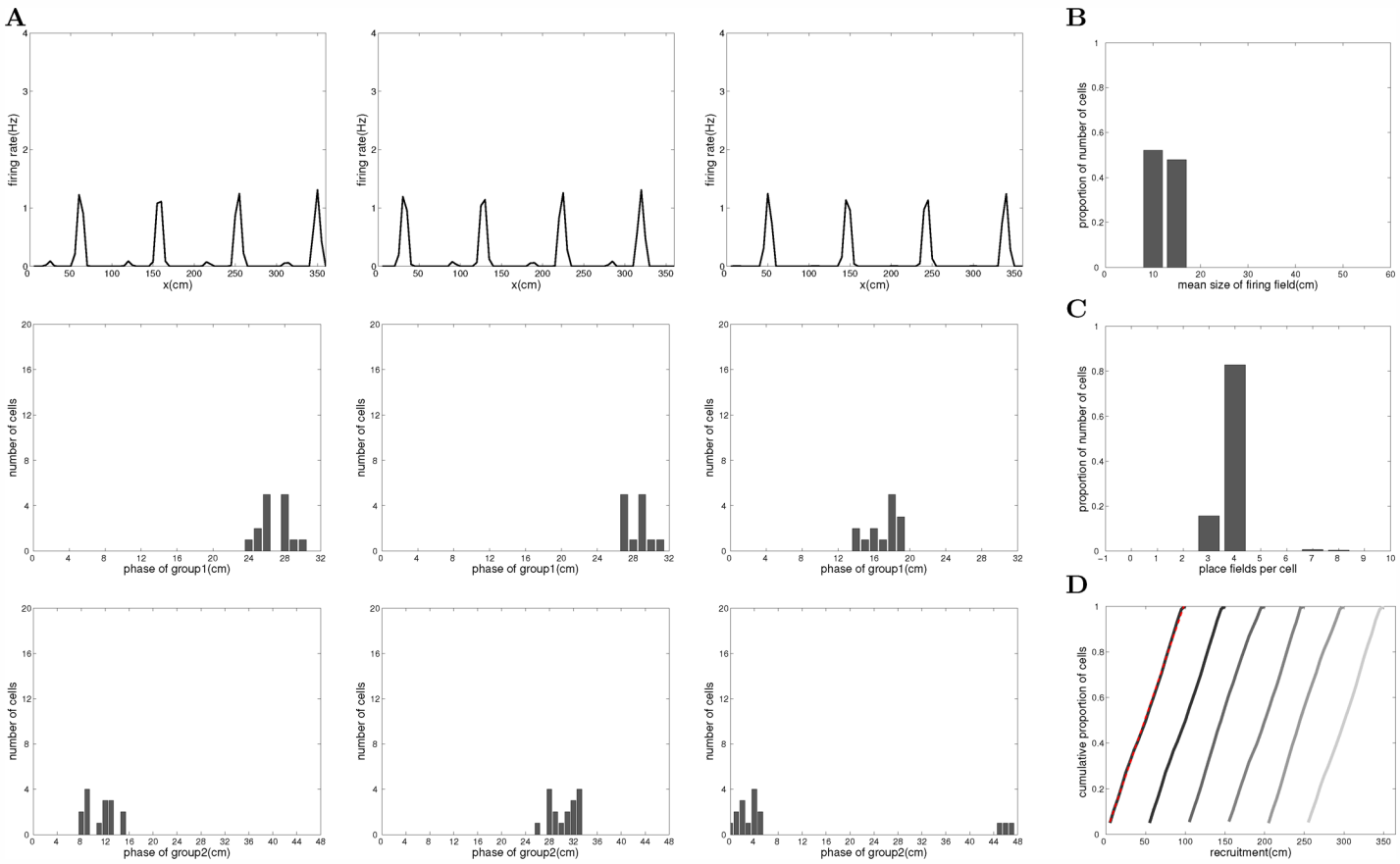

Fig. 2: Hebbian learning lead to clustered spatial phases of the downstream grid units that provide strong inputs. Parameters: $L = 2$, $\rho = 1.5$, $\sigma = 0$. (A) The firing fields (top row) of three example place units and the histogram (middle and bottom rows for the two grid modules) of the spatial phases of the grid cells that have the strongest weights (top 30%) to the downstream place unit; (B) The histogram of the mean field size of all active place units; (C) The histogram of the number of fields per place unit; (D) The cumulative proportion of the active cells in the total place cell population, counting from different starting positions (lines of different grayscales). The red broken line shows a linear fit.

phases (Fig. 2A, middle and bottom rows). The spatial phases of these grid cells roughly follow the relation in Eq. 7. For the three example shown in Fig. 2A, the coefficients are $(B = 1, A = 1)$, $(B = 0, A = 0)$ and $(B = 1, A = 1)$ respectively. The first firing location of the place cell is determined once A, B and the spatial phase x_j and x_k are fixed. Note that the fields of the place cell shown in the left column of Fig. 2 is a translated version of those of the cell shown in the right column. This is because their input grid cells share the same relative phase relation.

The field size is similar across place units (Fig. 2B), demonstrated by very peaked distribution around 12 cm. Due to the repeating firing patterns, units show multiple fields (Fig. 2C). We calculated the cumulative proportion of the active cells counting from different starting location, and plotted the recruitment curve as in [8]. The curves from different starting position share similar shape and can be fitted by straight lines (Fig. 2D).

When the ratio of grid scale $\rho = 1.421$, the realistic value observed from experiments, place units tend to have more place fields (Fig. 3A), resulting in broader distribution (Fig. 3B). This is because in this case the grids have more chances to produce overlapped inputs. The distribution of inter-field intervals is quite variable (Fig. 3C). The recruitment curves are linear (Fig. 3C), similar as in Fig. 2D.

B. Place cells encode larger environments with more grid modules

When the number of grid modules is increased to 5, place units express single place fields, due to larger least common multiple of the spacings (Fig. 4A-B). If the size of the environment is increased and becomes larger than the least common multiple of the spacings, place cells start to show repeating firing fields (data not shown). This is endowed by the nature of periodic inputs from grid cells.

With more modules, the size of fields is larger (Fig. 4C), due to the fact that place cells receive inputs from grids with larger scales. The network shows linear recruitment of place cells with respect to the size of the environment (Fig. 4D), similar as in Fig. 2D.

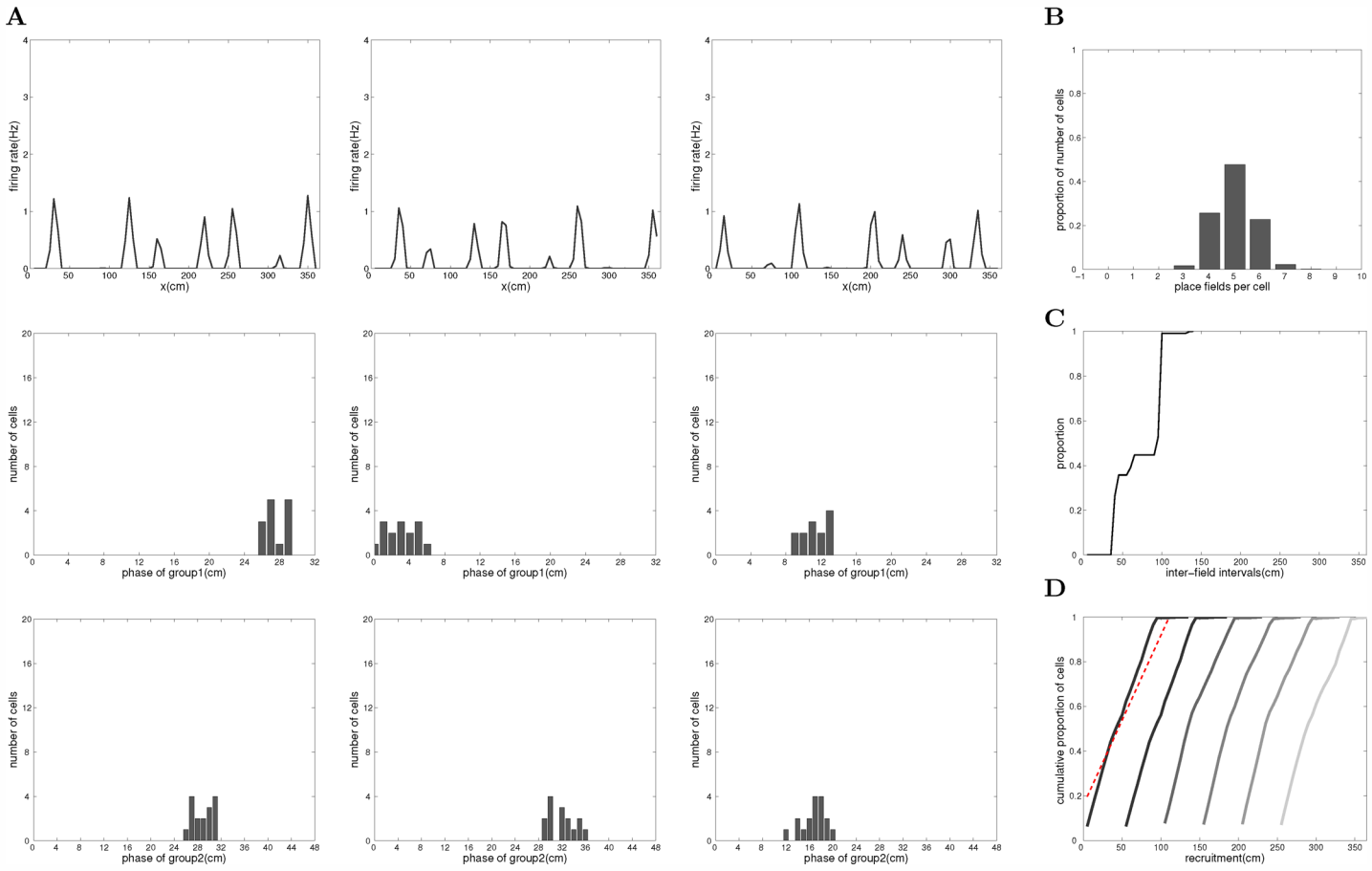

Fig. 3: The ratio of grid scale influences the number of place fields per place cell. Parameters: $L = 2$, $\rho = 1.421$, $\sigma = 0$. (A) The firing fields (top row) of three example place units and the histogram (middle and bottom rows for the two grid modules) of the spatial phases of the grid cells that have the strongest weights (top 30%) to the downstream place cell; (B) The histogram of the number of fields per place unit; (C) The cumulative distribution of the inter-field interval of all place units; (D) The cumulative proportion of the active cells in the total place cell population, counting from different starting positions (lines of different grayscales). The red broken line shows a linear fit.

C. Memoryless recruitment of place cells arises from non-spatial inputs

When the network receives non-spatial inputs ($\sigma > 0$), the firing fields of place units become random (Fig. 5A). The mean size of the fields does not change much (Fig. 5B vs. Fig. 2B), but the number of fields per cell become more diverse (Fig. 5C), and can be fitted by a poisson-gamma distribution [8]. The cumulative distribution of the inter-field intervals is close to that of an exponential distribution (Fig. 5D), indicating random appearing of place fields in the environment. The recruitment curve shown in Fig. 5E demonstrates that the network provides a memoryless process of generating new place fields. The distribution of the number of fields per cell, the cumulative distribution of the inter-field intervals and the recruitment curve are very similar to those observed in real rats [8]. The model proposed in this paper therefore might account for the computational mechanisms underlying the formation of place fields in large-scale environments.

V. Conclusion and Discussions

In this paper, we model the mechanisms of the self-organisation of spatial codes in hippocampal principal cells in large environments. The proposed model reproduces with high fidelity the statistics of hippocampal place fields, such as the distribution of the number of fields per cell, the cumulative distribution of the inter-field intervals and the memoryless recruitment curve. Thus the model suggests that place cells are able to form place fields in any random location due to the random phases of grid cells. The position of a place field however is determined by the phase difference between the projecting coactive grid cells, as a result of synaptic plasticity. The non-spatial inputs to place cells add rich morphology to the number of place fields and inter-field interval across cells. Overall, hippocampal place cell assembly constitutes sparse codes that are able to randomly emerge in a memoryless fashion when the animal explores large new environments.

In this study we only consider one-dimensional environment. Previous analysis showed that grid cells represent one-dimensional environments as linear subspaces from two

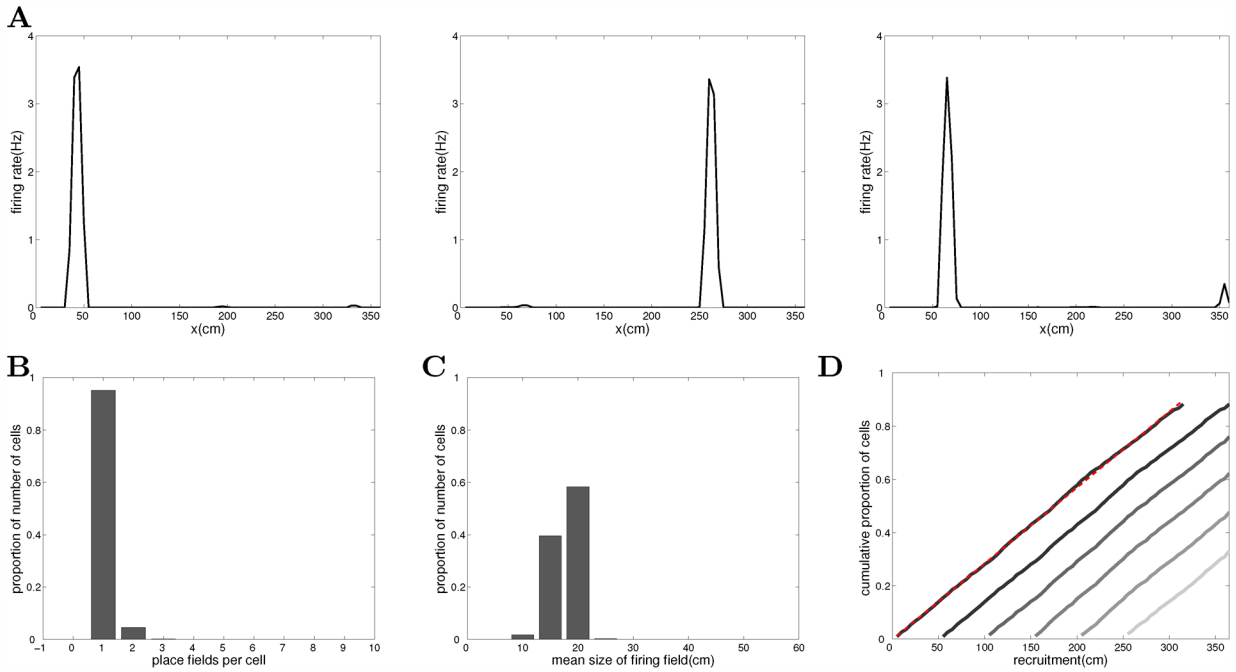

Fig. 4: The size of the environment that can be encoded by place cell ensemble increases with the number of grid modules. Parameters: $L = 5$, $\rho = 1.5$, $\sigma = 0$. (A) The firing fields of three example place units; (B) The histogram of the number of fields per place unit; (C) The histogram of the mean field size of all active place units; (D) The cumulative proportion of the active cells in the total place cell population, counting from different starting positions (lines of different grayscales). The red broken line shows a linear fit.

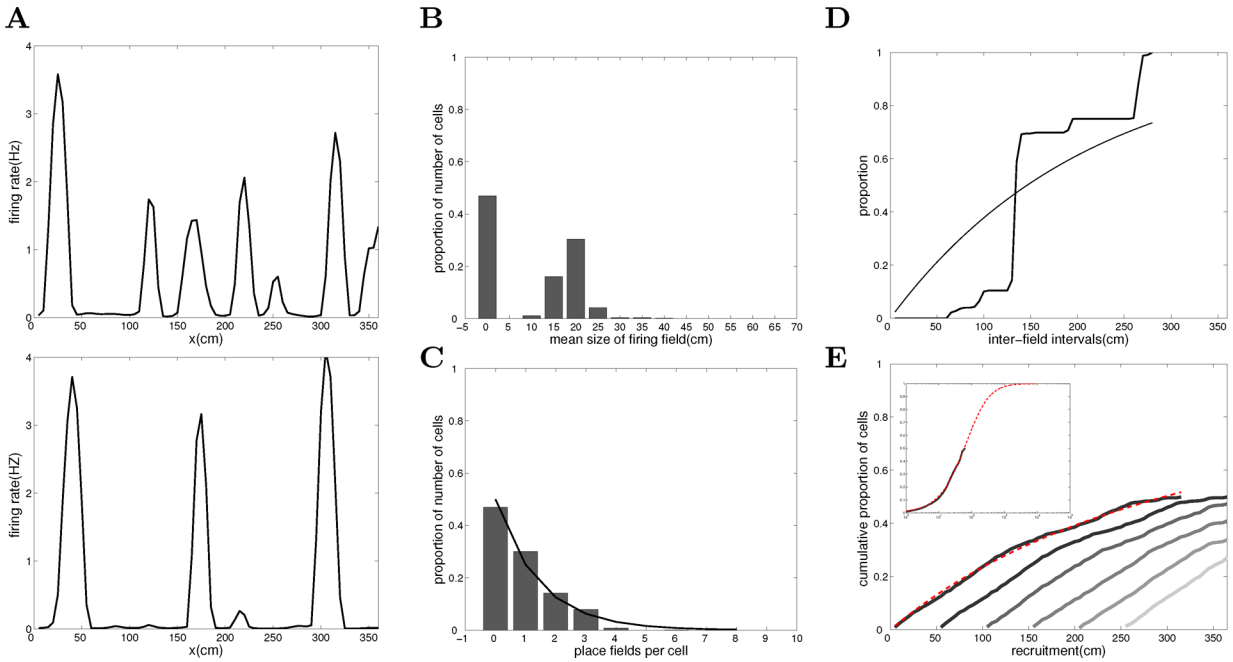

Fig. 5: Non-spatial inputs produce more place fields per units and memoryless recruitment of place units. Parameters: $L = 5$, $\rho = 1.5$, $\sigma = 3.5$. (A) The firing fields of two example place units; (B) The histogram of the mean field size of all place units; (C) The histogram of the number of fields per unit. The solid line shows the fit by a poisson-gamma distribution; (D) Cumulative distribution of the inter-field intervals. The solid line show the fit by an exponential distribution; (E) Cumulative distribution of the location where a cell first starts to fire counting only the fields from different positions on in the environment. The red broken line shows a fit by a gamma-Poission model [8]. The inset shows an extrapolation.

dimensional grids [23]. The results in this work are in principal valid for high dimensional environments. Future work includes the extension to 2D environments.

The segregation of spatial and non-spatial inputs is a key component of the model. It leads to further questions on whether there are distinctive properties between the two pathways, e.g. in terms of synaptic plasticity, the reliability and stability of the inputs, especially when global remapping happens as the animal travels between multiple environments. The modal assumes that the non-spatial inputs are constant. Experimental data indicate that the emergence of a place field can be induced by artificial stimulation or by attention during exploratory behaviour [24, 25]. It is well possible that the non-spatial inputs have some time dependent dynamics, which require further modelling work.

The results of this work can provide key insights on building neurorobots that are able to form efficient map representations of large environments to fulfil efficient navigation.

ACKNOWLEDGMENT

This work is supported by the "Hundred Talents Program" of CAS (No.Y3F9060901) to B.S..

REFERENCES

[1] Bailu Si, J. Michael Herrmann, and Klaus Pawelzik. Gain-based exploration: From multi-armed bandits to partially observable environments. In *Proceedings of the International Conference on Natural Computation*, pages 177–182. 2007.

[2] Lee Gomes. Hidden obstacles for google's self-driving cars. *MIT Technology Review*, August 2014. http://www.technologyreview.com/news/530276/hidden-obstacles-for-googles-self-driving-cars/.

[3] Raymond HG Klaassen, Thomas Alerstam, Peter Carlsson, James W Fox, and Åke Lindström. Great flights by great snipes: long and fast non-stop migration over benign habitats. *Biology letters*, 7(6):833–835, 2011.

[4] Asaf Tsoar, Ran Nathan, Yoav Bartan, Alexei Vyssotski, Giacomo DellOmo, and Nachum Ulanovsky. Large-scale navigational map in a mammal. *Proceedings of the National Academy of Sciences*, 108(37):E718–E724, 2011.

[5] Hanne Stensola, Tor Stensola, Trygve Solstad, Kristian Froland, May-Britt Moser, and Edvard I. Moser. The entorhinal grid map is discretized. *Nature*, 492(7427):72–78, 2012.

[6] Ila R Fiete, Yoram Burak, and Ted Brookings. What grid cells convey about rat location. *The Journal of Neuroscience*, 28(27):6858–6871, 2008.

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

[8] P Dylan Rich, Hua-Peng Liaw, and Albert K Lee. Large environments reveal the statistical structure governing hippocampal representations. *Science*, 345(6198):814–817, 2014.

[9] Charlotte B Alme, Chenglin Miao, Karel Jezek, Alessandro Treves, Edvard I Moser, and May-Britt Moser. Place cells in the hippocampus: Eleven maps for eleven rooms. *Proceedings of the National Academy of Sciences*, 111(52):18428–18435, 2014.

[10] T Solstad, EI Moser, and GT Einevoll. From grid cells to place cells: A mathematical model. *Hippocampus*, 16:1026–1031, 2006.

[11] ET Rolls, SM Stringer, and T Elliot. Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning. *Network: Comput Neural Sys*, 15:447–465, 2006.

[12] M. Franzius, R. Vollgraf, and L Wiskott. From grids to places. *J. Comput. Neurosci.*, 22:297–299, 2007.

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

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

[15] Bailu Si and Alessandro Treves. The role of competitive learning in the generation of DG fields from EC inputs. *Cognitive Neurodynamics*, 3(2):177–187, 2009.

[16] Amir H. Azizi, Natalie Schieferstein, and Sen Cheng. The transformation from grid cells to place cells is robust to noise in the grid pattern. *Hippocampus*, 24(8):912–919, 2014.

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

[18] Sheng-Jia Zhang, Jing Ye, Chenglin Miao, Albert Tsao, Ignas Cerniauskas, Debora Ledergerber, May-Britt B. Moser, and Edvard I. Moser. Optogenetic dissection of entorhinal-hippocampal functional connectivity. *Science*, 340(6128), 2013.

[19] BL McNaughton, FP Battaglia, O Jensen, EI Moser, and M-B Moser. Path integration and the neural basis of the "cognitive map". *Nature Reviews Neurosci.*, 7:663–78, 2006.

[20] MC Fuhs and DS Touretzky. A spin glass model of path integration in rat medial entorhinal cortex. *J. Neurosci.*, 26:4266–76, 2006.

[21] Yoram Burak and Ila R. Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput Biol*, 5(2):e1000291, 02 2009.

[22] Bailu Si, Sandro Romani, and Misha Tsodyks. Continuous attractor network model for conjunctive position-by-velocity tuning of grid cells. *PLoS Computational Biology*, 10(4):e1003558, 2014.

[23] Cristina Domnisoru, Amina A Kinkhabwala, and David W Tank. Membrane potential dynamics of grid cells. *Nature*, 495(7440):199–204, 2013.

[24] Doyun Lee, Bei-Jung Lin, and Albert K Lee. Hippocampal place fields emerge upon single-cell manipulation of excitability during behavior. *Science*, 337(6096):849–853, 2012.

[25] Joseph D Monaco, Geeta Rao, Eric D Roth, and James J Knierim. Attentive scanning behavior drives one-trial potentiation of hippocampal place fields. *Nature neuroscience*, 17(5):725–731, 2014.