Published in final edited form as:

*Neuron*. 2020 August 19; 107(4): 703–716.e4. doi:10.1016/j.neuron.2020.05.022.

# A distributed neural code in the dentate gyrus and in CA1

**Fabio Stefanini**[1,2], **Lyudmila Kushnir**[3], **Jessica C. Jimenez**[4,5], **Joshua H. Jennings**[6], **Nick I. Woods**[7,8], **Garret D. Stuber**[9], **Mazen A. Kheirbek**[7,10,11,12,+], **René Hen**[4,5,13,+], **Stefano Fusi**[1,2,13,6,+,v]

[1]Center for Theoretical Neuroscience, College of Physicians and Surgeons, Columbia University

[2]Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University

[3]GNT - LNC, Départment d'études cognitives, École Normale Supérieure, INSERM, PSL Research University, 75005, Paris, France

[4]Departments of Neuroscience, Psychiatry & Pharmacology, Columbia University, New York, NY, USA

[5]Division of Integrative Neuroscience, Department of Psychiatry, New York State Psychiatric Institute, New York, NY, USA

[6]Department of Bioengineering, Stanford University, Stanford, CA, 94305, USA

[7]Neuroscience Graduate Program, University of California, San Francisco, San Francisco, CA, USA

[8]Medical Scientist Training Program, University of California, San Francisco, San Francisco, CA

[9]4158, USA 9Center for the Neurobiology of Addiction, Pain, and Emotion, Department of Anesthesiology and Pain Medicine, Department of Pharmacology, University of Washington, Seattle, WA 98195, USA

[10]Department of Psychiatry, University of California, San Francisco, San Francisco, CA, USA

[11]Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, USA

[12]Kavli Institute for Fundamental Neuroscience, University of California, San Francisco, San Francisco, CA, USA

[13]Kavli Institute for Brain Sciences, Columbia University

[+]Senior and corresponding authors: mazen.kheirbek@ucsf.edu, rh95@cumc.columbia.edu, sf2237@columbia.edu.

[v]Lead contact: Stefano Fusi, sf2237@columbia.edu

AUTHOR CONTRIBUTIONS

Conceptualization, S.F., R.H., M.A.K., F.S., and L.K.; Methodology, M.A.K., R.H., and G.D.S.; Software, F.S., and L.K.; Formal Analysis, F.S., L.K., S.F.; Investigation, J.H.J., M.A.K., J.C.J., N.I.W.; Resources, J.H.J, G.D.S, M.A.K., R.H.; Data Curation, J.C.J., F.S., M.A.K.; Writing - Original Draft, F.S., S.F.; Writing - Review & Editing: F.S., S.F., M.A.K., R.H.; Visualization, F.S.; Supervision, M.A.K., S.F., R.H.; Funding Acquisition, F.S., M.A.K., S.F., R.H..



DECLARATION OF INTERESTS

The authors declare no competing interests.

## Abstract

Neurons are often considered as specialized functional units that encode a single variable. However, many neurons are observed to respond to a mix of disparate sensory, cognitive and behavioral variables. For such representations information is distributed across multiple neurons. Here we find this distributed code in the dentate gyrus and CA1 subregions of the hippocampus. Using calcium imaging in freely moving mice, we decoded the animal's position, direction of motion and speed from the activity of hundreds of cells. The response properties of individual neurons were only partially predictive of their importance for encoding position. Non-place cells encoded position and contributed to position encoding when combined together with other cells. Indeed, disrupting the correlations between neural activities decreased decoding performance, mostly in CA1. Our analysis indicates that population methods, rather than classical analyses based on single cell response properties, may more accurately characterize the neural code in the hippocampus.

## eToc:

Using high-resolution calcium imaging and machine learning techniques, Stefanini et al. studied how information about position, direction of motion and speed is represented in the hippocampus. They show that neurons cooperate to encode this information and cells with easily interpretable responses are not necessarily the most important ones.

### Keywords

hippocampus; dentate gyrus; calcium imaging; decoding; population coding; distributed representations; correlated activity

## INTRODUCTION

The hippocampus has been extensively studied in experiments on navigation and spatial memory. The responses of some of its cells are easily interpretable as these tend to fire only when the animal is at one location in an environment (place cells). However, it is becoming clear that in many brain areas, which include the hippocampus and entorhinal cortex, the neural responses are very diverse (Rigotti et al., 2013; Eichenbaum, 2018; Fusi et al., 2016; Hardcastle et al., 2017) and highly variable in time (Fenton and Muller, 1998; Ziv et al., 2013; van Dijk and Fenton, 2018). Place cells might respond at single or multiple locations, in an orderly (grid cells) or disorderly way and multiple passes through the same location typically elicit different responses. Part of the diversity can be explained by assuming that each neuron responds non-linearly to multiple variables (mixed selectivity) (Rigotti et al., 2013; Kriegeskorte and Douglas, 2019; Saxena and Cunningham, 2019). Some of these variables may not be monitored in the experiment and hence contribute to what might appear as noise. A neural code based on mixed selectivity is highly distributed because some variables can be reliably decoded only by reading out the activity of a population of neurons. It was recently shown that the mixed selectivity component of the neuronal responses is important in complex cognitive tasks (Rigotti et al., 2013; Fusi et al., 2016) because it is a signature of the high dimensionality of the neural representations. Place cell discharges

are also highly variable (Fenton and Muller, 1998) to the extent that the variability, not the spatial tuning alone, can capture changes due to learning in a spatial memory task (Olypher et al., 2003; Kelemen and Fenton, 2010; van Dijk and Fenton, 2018). These recent studies naturally pose the question of how position is encoded within the population activity in the hippocampus. To answer this question, we used calcium imaging to record the activity of large populations of neurons in the dentate gyrus (DG), a region of the hippocampus in which the neural responses are highly sparse and diverse (Leutgeb et al., 2007; Danielson et al., 2016; van Dijk and Fenton, 2018), and in CA1, a region that has been extensively studied in relation to spatial navigation using both electrophysiology (Moser et al., 2008; Harvey et al., 2009; Keinath et al., 2014; Agarwal et al., 2014) and imaging (Dombeck et al., 2010; Ziv et al., 2013).

We show that the position of a mouse freely exploring an environment can be decoded from the activity of a few tens of granule cells (GCs) of DG with an accuracy comparable to that of CA1. Using machine learning techniques, we ranked neurons by their contribution to position encoding. We found that trial averaged, single neuron tuning properties are insufficient to predict a neuron's contribution to position encoding at the population level. Cells that were not spatially tuned according to a statistical test based on spatial information (non-place cells) also contributed to the population code, to the extent that position could be decoded from the ensemble of these untuned cells alone in both areas. We further found that neurons in both DG and CA1 reliably encoded other variables such as the direction and speed of movement. These neurons were not distinct from the neurons that encoded position, i.e., the majority of neurons encoded multiple variables and contributed to all of them. We then found that destroying correlated activities among neurons while maintaining their spatial tuning had an impact on decoding performance in CA1 but not in DG. Taken together, these results show that the information encoded at the population level is far richer than at the single cell level and allowed us to uncover the strong robustness of DG and CA1 spatial coding through the distributed nature of their neural representation.

## RESULTS

We studied the neural code in the DG and in the CA1 area of the hippocampus of freely moving mice. We used miniaturized head-mounted microscopes to perform calcium imaging of granule cells (GCs) in the DG and of pyramidal cells in CA1. To image cell activity patterns we injected a virus encoding the calcium indicator GCaMP6 and implanted a gradient index (GRIN) lens for chronic imaging (Fig. 1). Four weeks after surgery, we imaged cellular activity while mice foraged for sucrose pellets in an open field arena. We then used a recently developed algorithm for reliably extracting the GCaMP signals from the raw videos, CNMF-e (Zhou et al., 2018) (Fig. 1d, e). This algorithm separates local background signals due to changes in fluorescence in the neuropil from the signals due to calcium concentration changes in individual cells. This was necessary to identify signal sources in our granule cells imaging data without introducing spurious distortions or correlations among cells due to artifacts. We identified a total of 1109 DG cells across 3 animals, among which 352 (32%) were significantly tuned to position and a total of 863 CA1 cells across 3 animals, among which 38 (4%) were significantly tuned to position (see Methods and Fig. S2).

The low fraction of place cells in CA1 seems to be in contrast with reports from previous studies in CA1 (Ziv et al., 2013; Pfeiffer and Foster, 2013; Meshulam et al., 2017; Talbot et al., 2018). However modern tools for source extraction from calcium imaging can detect cells with very low activity, which are largely underestimated in electrophysiological recordings as other recent studies have also suggested (Dipoppa et al., 2018; Tang et al., 2018). If one excludes from the analysis these low firing rate cells, the fraction of place cells becomes significantly higher (see Fig. S3 and Table S1). Moreover, while calcium signal extraction may miss isolated spikes, this is likely less of an issue for dentate gyrus granule cells than pyramidal cells since the former are often active in burst, as electrophysiology studies have shown (Pernía-Andrade and Jonas, 2014) and as our data also shows (Fig. 1e). See Fig. S3 and Table S1 for a brief review of the literature comparing firing rates in DG and CA1 across studies.

The first step of our analysis was to assess whether the position of the animal is encoded in the recorded neural activity during mobility. We therefore removed all the time bins in which the animal was slower than 2 cm/s for a period longer than 1 s, after confirming by visual inspection that this procedure would exclude moments of immobility. To decode position, we discretized the $x$ and $y$ coordinates of the animal by dividing the arena into 64 regions (8 by 8 grid) (50 cm square arena for DG mice, 50 × 28 cm for CA1 mice). We then trained a battery of linear classifiers for each pair of discrete locations. Each session was divided into 10 one-minute long intervals, 9 of which were used to train the classifiers and the remaining ones to test them (10-fold cross validation). We used a majority rule (Bishop, 2006) to combine the outputs of the linear classifiers as an instantaneous estimate of the animal's location, using the center of the selected location as the decoded position.

In both areas, the median decoding error was comparable to the animal size, revealing for the first time that instantaneous position can be decoded from DG GCs population activity (Fig. 2). Our analysis of the CA1 data shows a comparable decoding accuracy in DG and CA1 after correcting for number of cells (Fig. 2c and Fig. S4). The accuracy was slightly higher than the one observed in previous studies in CA1 (Ziv et al., 2013). Different decoding strategies, such as decoding from raw Calcium traces or events, produced similar results (see Fig. S6). The decoding error was found to weakly correlate to the speed of movement (see Fig. S7). To our knowledge, this is the first time that decoding of position from populations of DG cells has been reported.

We could also decode the direction of motion of the animal in both regions and its speed only in DG. Speed was weakly correlated with the overall level of activity in DG and we could decode it in two animals out of three using linear regression (Fig. 2b, c). To decode the direction of motion we divided the full range of possible directions into 8 angular bins and labeled time bins according to the instantaneous discrete direction of motion of the mouse (see Methods). To our knowledge, this is also the first time that decoding of direction and speed of motion from populations of DG and CA1 cells has been reported although direction tuning has been previously observed in CA1 pyramidal cells in rats (Acharya et al., 2016). We did not find differences in decoding performance for direction of motion between the DG and CA1 areas (Fig. 2c).

To better characterize the neural code, we tried to determine what features of the response properties of individual neurons are important for encoding the variables we could decode. It is important to realize that these response properties could be dissociated from the contribution of a cell to the accuracy of a decoder that reads out a population of neurons. For example, there could be neurons that are only weakly selective to position and so individually would not pass a statistical test for spatial tuning. However, when combined with other neurons, they can still contribute to position encoding. Alternatively, there are situations in which the decoder might assign a large weight to neurons that are weakly selective or even not selective at all but they are correlated to selective neurons. This situation can be illustrated with the intentionally extreme case of Fig. 3, in which we show how the responses of individual neurons can be dissociated from their importance for the decoder. A simulated animal visits two locations of the arena multiple times. The activity of two hypothetical neurons is represented in the activity space (Fig. 3b), with the horizontal and vertical axes representing the activity of the first and the second neuron respectively. At each pass through each location the two neurons have different activity due to other variables that might also be encoded, e.g., the direction of movement, the speed of the animal, or other variables that are not under control in the experiment. Each point in the activity plot represents the activity of the neurons in a single pass. The responses of neuron 2 to the two different locations have the same distribution (Fig. 3b,c). A cell with such response properties is untuned to space (a non-place cell) and therefore it is typically considered unimportant for encoding position. However, a linear decoder trained to decode the position of the animal can make use of the untuned neuron because of the correlations between the activities of the two neurons. While the activity of neuron 1 is only partially predictive of the animal's location (the distributions partially overlap), by reading out neuron 2 together with neuron 1 it is possible to decode position with no errors using a linear decoder. In such situation, the linear decoder would assign equal weights to the two neurons, as shown in Fig. 3c.

In the real data, there might be a spectrum of different situations that are less extreme than the one illustrated in Fig. 3 in which a decoder can take advantage of weakly tuned cells. Cells like the untuned one in Fig. 3 or weakly tuned cells can "cooperate" with more tuned cells to encode more precisely a variable like position. This is a situation similar to the one of Fig. 3, in which the correlations between the activities of different neurons would be important. However, there might also be weakly tuned cells that are uncorrelated, but when combined together would contribute to the accuracy of a decoder. In both cases, the decoder can use the weakly tuned cells to improve its accuracy. Analogously a downstream neuron can in principle harness the activity of weakly tuned neurons to readout the animal's position.

In our analysis we took the perspective of such a readout neuron and analyzed the weights assigned to cells by our decoder to determine the importance of input neurons in a population for encoding position. The procedure we adopted was to first train the position decoder on each pair of locations and then to combine the resulting weights to obtain a single importance index ($\omega$) for each cell (see Methods). Similar methods are used to assess the importance of individual features in a feature space (Haufe et al., 2014; Mladenić et al., 2004) and have been recently used to identify important synapses in learning models (Zenke

et al., 2017). We then ranked the neurons according to this importance index and estimated the decoding accuracy for populations of 50 neurons (Fig. 4a) to assess the validity of our approach. The 50 neurons with the largest importance index indeed performed significantly better than the worst 50 neurons, though position could be decoded above chance level even from the worst neurons. The accuracy decreases progressively between the performance for the best and for the worst neurons, validating the method for ranking the neurons on the basis of the importance index. We also controlled that the ranking was stable within the session (see Fig. S8) and that it was not due to poor cell segmentation (Fig. S24).

The observation that most neurons could contribute to the decoding of position indicates that the neural code is highly distributed. Indeed, the importance index is rather similar for most of the cells. To quantify the distribution of importance across cells, we used the Gini coefficient, a quantity that is often used to represent wealth inequality in a country. A high Gini coefficient indicates high inequality, as in a dedicated code whereby few neurons strongly encode a given variable, whereas low values correspond to an equal distribution of resources as in a distributed code. We observed low values both in DG and CA1, indicating that different neurons tend to contribute equally to the encoding of position, a signature of a distributed code rather than a situation in which only a few cells are important (see Fig. S13).

Not too surprisingly, one important feature of an individual neuron is its average activity, which is strongly correlated with the importance index and hence to the overall ability to encode position (see also Fig. 5a, b and Fig. S19). However, inspection of the firing fields of Fig. 4b indicated that there were no other obvious properties that predicted whether a neuron is important or not in both DG and CA1 neuron populations. We then identified which neurons were spatially tuned and called them place cells if the spatial information contained in their activity was statistically significant (see Methods for details). The difference between the spatial information for the recorded activity and the spatial information obtained for shuffled data, properly normalized, is what we defined as significance of spatial information (SSI). It is indeed a measure used to assess whether a cell is a place cell or not relative to a null distribution (Skaggs et al., 1993; Panzeri et al., 2007; Meshulam et al., 2017; Danielson et al., 2017; Allegra et al., 2019).

From Fig. 4b it is clear that in our data there are non-place cells that have a large importance index (see also Fig. S17). The animal position could be decoded from these cells alone in both DG and CA1 (Fig. 4c and Fig. S17). This indicates that the activity of the non-place cells contains some spatial information. However, because of noise and limited data, the activity of these cells did not pass the statistical test that we adopted to characterize place cells.

While the SSI is a property of the single cell, the importance index depends on the contribution of a cell to the population code. We thus analyzed the relation between each cell’s SSI and its importance index. Although we did not find a one to one correspondence between SSI and importance index, the two quantities were correlated (Fig. 5a, b, c and Fig. S19) indicating that some individual response properties are at least partially informative about the importance of a cell in encoding position. To compute the SSI one has to compute

the spatial information and subtract a baseline obtained by shuffling the activity. The spatial information without the baseline subtraction, which is sometimes used as a measure of the tuning of the cells, was actually negatively correlated with the importance index (see Fig. S9a and (Kelemen and Fenton, 2010)). However, the baseline was also negatively correlated with the importance index (Fig. S9b). The net effect is that the correlations between SSI and importance index were positive. The negative correlations are a reflection of the sampling bias problem that affects cells with low activity (Panzeri et al., 2007). Low activity cells tend to be more selective because the fluctuations of the activity are relatively large. However, these cells are typically unrealiable (e.g., they fire at a particular location only for one passage) and hence their importance index is low.

We performed a similar analysis of importance for the direction of movement. In Fig. 6 (see also Fig. S18) we show that we could rank the cells according to their contribution to decoding (Fig. 6a and Fig. S18) and that the important cells were highly heterogeneous in their direction tuning (Fig. 6b). Considering all recorded cells, we also found that a cell's activity correlated with the importance index for direction of movement in DG and in CA1 (Fig. 6b). We defined the significance of directional information (SDI) in a similar way to the SSI by comparing the mutual information between direction of motion and a cell's activity to a distribution obtained by shuffling the cell's calcium events in time. The importance index and this directional information were correlated both in DG and in CA1 (Fig. 6d and Fig. S20).

All these analyses indicate that single neuron properties are only partially predictive of the importance of a cell for decoding. Moreover, the importance is not an intrinsic property of an individual cell because it clearly changes depending on which other cells are part of the population of neurons that are used by the decoder. This is illustrated in Fig. 3, in which the untuned cell is important when combined together with the cell represented on the horizontal axis, but it would be useless if combined together with another untuned cell, or with an uncorrelated tuned cell.

Since we could decode at least two variables from the neural activities, we were wondering whether we could identify some form of specialization in which segregated groups of neurons encoded different variables. In Fig. 7 we report the importance index for the direction of movement versus the importance index for position (Fig. 7a and Fig. S21). The situation in which different variables would be encoded by segregated populations of neurons would predict a negative correlation between these two importance indices: cells with large importance index for position should have a small importance for the direction of movement, and vice versa. Instead, for both the regions we analyzed, we found a positive correlation between the two quantities, with a higher correlation in CA1, suggesting that neurons that are important for encoding one variable are also important for encoding the other. This is partially explained by the fact that for both position and direction of movement the most active cells tend to be the most important ones. However, when we regressed out the components explained by the activity, we still found a positive correlation between the importance indices of the two variables (Fig. 7a). In addition, this could not be explained by a correlation between direction of movement and position (Fig. S11).

We then focused on those cells that had a high importance for one variable and not for the other as candidate specialized cells. However, we could decode position from the most important cells for encoding direction of motion and vice versa, showing that even the most important cells for one variable carry information about the other variable in both regions (Fig. 7). We conclude that in both CA1 and DG, neurons have mixed selectivity to the variables we decoded, in line with recent studies both in CA1 (Meshulam et al., 2017) and in the cortex (Rigotti et al., 2013; Fusi et al., 2016; Hardcastle et al., 2017; Lindsay et al., 2017). See also the Discussion.

So far, we have shown that the code that is used to represent position is distributed, i.e., all active cells contribute to some extent to the population code. We therefore sought to see if correlations between the activities of different neurons contribute to the decoding performance, in a way similar to what we described in Fig. 3. To understand the contribution of correlations to the encoding of position it is important to distinguish different components, and in particular the correlations generated by the signal (i.e., the position of the animal) and those that can be considered as noise (i.e., not related to the encoding of position). The signal component is induced by the tuning properties of individual neurons. For example, two place cells that have highly overlapping fields are going to be correlated since they tend to be co-active when the animal is at a particular location. Noise correlations represent the component that cannot be explained by the signal and they are essentially due to the fact that every time the animal is at particular location the neural response can be different. Noise correlations can be beneficial, detrimental or irrelevant for the neural code (Abbott and Dayan, 1999; Schneidman et al., 2003; Brody, 1999). However, our initial hypothesis was that a large portion of the noise variance can be explained by the fact that neurons encode multiple variables besides position (see Discussion). For example, the different points that encode the same position in Fig. 3 might correspond to visitations in which the head direction and/or the speed were different. In this case, destroying the correlations would result in a decrease or no change in decoding performance (Fig. 8a).

We devised a procedure to shuffle the data in a way that destroys the noise correlations across neurons maintaining the spatial tuning of each cell (see Methods and Fig. S14). We then studied the effect of this procedure on the decoding accuracy for position. At each pass through a location, we randomly picked the activity of a cell from the pool of recordings corresponding to that location and that cell (Fig. 8b). We then corrected for the different time spent in each pass at the same location and repeated the procedure for all cells independently. By using this procedure we effectively destroyed the noise correlations between neurons because, after this manipulation, each cellâ s activity was independent from the others. However, by restricting the manipulation to each discrete location, we did not alter the spatial tuning of the cells (Fig. 8c) nor the signal correlations among neurons induced by their tuning profiles. By comparing the performance of the decoder on the modified data to the one on the original data, we could then assess the contribution of the noise correlations to decoding. This is a direct test of the presence of a structure in the neural representations that is beneficial for representing information (Abbott and Dayan, 1999; Averbeck and Lee, 2006; Pillow et al., 2008; Eyherabide and Samengo, 2013). In 4 of the 6 analyzed animals we found that the decoding error increased when correlations were destroyed through the shuffling procedure, revealing the importance of correlations

(Fig. 8d). The effect was very consistent in CA1 neurons where performance was reduced by about 20%, whereas almost no effect was observed in DG (see Fig. S14). Pairwise correlations were found to be lower in DG than in CA1 (Fig. S16) and this may partially explain the main effect that disrupting correlations can lead to different changes in decoding accuracy in the two areas. However, it might only partially explain the effect because the correlations are not completely absent in DG and those that we observed certainly change after disrupting the noise correlations (for an analysis on how destroying correlations impacts pairwise correlations and importance index see Fig. S16 and Fig. S23).

## DISCUSSION

Neurons in the DG and in CA1 have rather diverse response properties and often the responses are not easily interpretable (Danielson et al., 2016; Leutgeb et al., 2007). Despite this seemingly disorganized neural code, it is possible to decode from the activity of a population of neurons the position, the speed and the direction of motion of the animal. Neurons respond to mixtures of the decoded variables as observed in other high cognitive brain areas (Rigotti et al., 2013; Fusi et al., 2016). The information about these variables is highly distributed across neurons to the point that the responses of individual neurons are only weakly predictive of their contribution to the neural code. It is therefore crucial to consider neurons in one region of the brain as part of an ensemble to assess their importance for processing and transferring information about a particular variable.

One implication of such distributed neural code is that it can be misleading to characterize the function of a brain region based only on the statistics of individual neuron properties. In the specific case of position encoding, for instance, it is not possible to conclude to what extent the position of the animal is encoded only by analyzing the tuning of individual cells to space. Indeed, populations of cells whose activities do not pass a selectivity criterion for space encoding, for example through an information theoretical approach, may still encode position via the ensemble activity patterns, as we showed by decoding position and direction of motion from untuned cells in both DG and CA1 regions of the hippocampus.

The population coding rescues the ability of these areas to encode position despite the sparsity of its activity and the variability of its representations. Here we show that indeed even few tens of cells encode position with high precision in both analyzed areas. Furthermore, the decoding was accurate even when model training and model test periods were separated by up to 18 minutes, indicating that at the population level the representations were stable, despite the elevated variability of individual cells (see Fig. S12).

Our findings are in line with studies suggesting that session averaged, single cells statistics fall short in describing the activities of hippocampal cells. For example, while place fields are widely used to analyze DG activities in remapping studies (Leutgeb et al., 2007), it is only when sub-second network discharge correlations are taken into account in the analysis that memory discrimination signals can be revealed (van Dijk and Fenton, 2018). More importantly, we address one important question about the role of non-place cells in CA1 and DG areas of the hippocampus. A recent work by Meshulam and colleagues (Meshulam et al., 2017) used a maximum entropy model to describe the neural activity recorded in

CA1. The model is constructed from the second order statistics (the correlations between neurons) and it accurately predicts the activity of each neuron from the state of all the other neurons in the network, regardless of how well that neuron codes for position. They conclude that correlation patterns in CA1 hippocampus only partially arise from place encoding. Moreover, their results suggest that understanding the neural activity may require not only knowledge of the external variables modulating it (i.e., the position of the animal) but also of the internal network state. Our results indicate that the correlation patterns not due to position encoding can be partially explained by the encoding of other external variables (e.g., the direction of movement). However, it is likely that some components of the correlation patterns encode the internal state of the animal, as suggested by Meshulam and colleagues (Meshulam et al., 2017). Our analysis also shows directly that the non-place cells contribute to the encoding of the position of the animal. This is partially due to the fact that some of the cells, if taken individually, encode position only weakly and hence do not pass the statistical test for being categorized as place cells (the criterion used by Meshulam and colleagues to select place cells is similar to ours (Meshulam et al., 2017)). However, non-place cells contribute to position encoding also thanks to their correlations with place cells. Indeed, when the noise correlations are destroyed, the decoding accuracy decreases in CA1. This is compatible with their observation that place and non-place cells belong to the same network when the patterns of correlations are considered, but it also goes beyond their analysis because it shows directly that correlations are important for encoding position. In conclusion, to study the neural code in one particular region of the brain one has to consider all cells in a population as the tuning properties which are based on single cell statistics might not be sufficient to understand how task relevant variables are encoded (see also (Fusi et al., 2016; Kriegeskorte and Douglas, 2019; Saxena and Cunningham, 2019)).

### Poor spatial tuning and the advantages of mixed selectivity

One of the important observations that we discussed in the article is that there is a large proportion of cells that exhibit poor spatial tuning. The computational advantage of poor spatial tuning can be understood only when one considers a situation in which the neurons in a population encode not only the position of the animal, but also several other variables (e.g., head direction, the velocity of the animal and other unknown variables that are not under control in our experiment). This can be implemented in different ways. For instance, each variable could be encoded by a different group of highly specialized neurons. However, these representations are low dimensional and hence they greatly limit the number of combinations of input variables that a linear readout, or a downstream neuron can respond to (see for instance (Fusi et al., 2016)). One simple example is a downstream neuron that must respond when the animal is looking at the center of the arena from two opposite corners. Such a simple situation is equivalent to the exclusive-or (XOR) problem in which the combinations of variables (position and head direction) that should activate the neuron (animal looking at the center of the arena) and those that should not (same positions, animal looking in the opposite directions) are not linearly separable. Instead, when head direction and position are mixed non-linearly the neural representations can be high dimensional and a linear readout can separate any set of combinations of inputs from the others. This is why in most problems that involve at least two variables, mixing non-linearly all the relevant variables is beneficial. Mixing position with other variables, like head direction, leads to

relatively poor spatial tuning but it also confers more flexibility to linear readouts, greatly increasing the computational ability of the network.

In our paper we showed that also our data supports the hypothesis for mixed selectivity, already entertained by previous research, in the specific case of spatial representations in the hippocampus. We verified this hypothesis by showing that neurons in DG and CA1 encode multiple spatial variables, not only position, through decoding. In such a situation it is possible, and very likely, to find that neurons that do not explicitly encode position, i.e., they do not have a significant spatial tuning, but are still important to discriminate between pairs of locations through correlations imposed by the geometry of the neural representations. In the cartoon of Fig. 3, for instance, we show how the modulation of activity imposed by correlations can be used by a decoder to perfectly discriminate two locations (see also recent reviews in (Kriegeskorte and Douglas, 2019; Saxena and Cunningham, 2019)). That modulation must therefore be intended as the result of the population response to combinations of variables which include position, speed, movement direction, and possibly other variables that were not under our control (Allegra et al., 2019). We modeled such situation in Fig. S15 and found regimes in which correlations imposed by other variables help and situations where they do not have a significant effect on decoding performance depending on the geometry of the neural representations. Taken together, our results show the advantages of a distributed code in that it can reliably represent multiple combinations of variables.

## The encoding role of correlations

Destroying correlations among neurons did not have a strong impact in decoding performance in DG neurons but it consistently reduced decoding performance in CA1 data. Whether neural correlations are used in the population code is a long-standing question. In the data, it has been shown in the past that the pair-wise correlations accounted only for about 10% of the information contained in neural activities (Averbeck and Lee, 2006; Latham and Nirenberg, 2005; Schneidman et al., 2006) while using models that exploit higher-order correlations can recover about 20% of information related to the stimulus in a population of retinal ganglion cells (Pillow et al., 2008). Here we showed that the disruption of correlations leads to relatively modest but statistically significant decrease in decoding accuracy in CA1 but not in DG.

These observations indicate that the correlations that we are destroying should be considered as signal correlations rather than noise correlations, at least in CA1. The variability across visitations can probably be explained by the fact that neurons encode multiple variables in a consistent way and may induce the observed neural correlations (Wood et al., 1999; Allegra et al., 2019). This situation would be similar to the one discussed in Fig. 3 (e.g., pass 1 and 2 would correspond to two visitations of location A with a different direction of motion), i.e., the disruption of the correlations decreases the performance of the decoder. In Fig. S15 we show in simulations that this is indeed the case. We considered a model in which the neural activity depends on multiple variables, for instance the position of the animal, the direction of motion, etc. Each variable can assume a discrete set of different values and every set of values of the encoded variables defines one specific condition. We

then constructed different neural representations by arranging the different conditions in the space of neural activities. In particular, we considered two scenarios, one with unstructured representations, whereby different conditions are represented by different random vectors in the activity space, and one with a kind of structured geometry which is beneficial for generalization across conditions (Bernardi et al., 2018). In both cases the encoded variables are linearly separable, i.e., they can be decoded with a linear classifier. We then compared the linear decoder's performance before and after destroying the correlations, as we did in the real data, for different numbers of conditions in each scenario.

In most of the scenarios that we simulated, the decoder performance is either disrupted or it remains the same when the correlations are destroyed. The beneficial effect of the correlations is maximal when the representations are fairly unstructured. In the case of random representations, the effect is maximal for a certain number of conditions. This number depends on the number of encoded variables and on the number of values that each variable can hold (i.e., the total number of conditions). Our experimental observations that show that the decoder's performance is disrupted more in CA1 than in DG are compatible with a scenario in which the representations in CA1 are unstructured, similarly to the simulated representations obtained with the random model. Our results also show that the representations in DG are compatible with at least two scenarios: 1) they could be structured, as we described them in Fig. S15d–f, or 2) they could also be unstructured as in CA1 but with a different number of encoded variables, either very small or very large. It is important to stress that the scenarios studied in Fig. S15 are all plausible in the sense that they are based on representations that have already been observed in other studies (Rigotti et al., 2013; Bernardi et al., 2018). However, the examples we report are certainly not exhaustive and so we cannot exclude that other codes that we did not consider may be more appropriate to describe DG and CA1 represention.

One alternative explanation for the difference between CA1 and DG comes from the fact that the performance reduction that follows the disruption of correlations depends on the level of activity of the cells, and that CA1 and DG exhibit different levels of activity. However, in our data this difference in activity levels could not fully account for the difference between CA1 and DG in the effect of destroying correlations as we did not observe any performance reduction in DG when the level of activity was matched to the one observed in CA1 (Fig. S14).

Our simulations in which multiple variables are encoded are compatible with recent models of the hippocampus that emphasize its role in memory compression (Gluck and Myers, 1993; Benna and Fusi, 2019), and memory prediction (Dayan, 1993; Stachenfeld et al., 2014, 2017; Gershman et al., 2012; Recanatesi et al., 2018; Whittington et al., 2019): for all these models the neural representations in the hippocampus are constructed by learning the statistics of the sensory experiences in order to generate a compressed representations of the memories to be stored, or, when focused on temporal sequences, to generate a prediction of the next memory (successor representation). Future theoretical work will establish more quantitatively whether this scenario is fully compatible with our observations and what the different roles of CA1 and DG could be in this compression process.

### Average activity in DG is larger than in CA1

One of the observations that require some discussion is that the average activity in DG is larger than in CA1 in our data. This might sound surprising, but a careful review of the literature shows that our observations are compatible with other studies (see Fig. S3 and Table S1). Indeed, the average firing rate that is reported varies from study to study depending on the recording technique, the type of experiment and whether rats or mice have been employed. Our conclusion is that our results fall within the range of values reported in the existing literature. The review reported in the Supplementary Material is not exhaustive by any means, but we believe it is highly representative of the existing literature.

### Conclusion

Our results strengthen the hypothesis that the neural code in the DG and in the CA1 area of the hippocampus is highly distributed and that it is important to analyse it using a population approach (Fenton et al., 2008; Meshulam et al., 2017; van Dijk and Fenton, 2018). The analysis of the averaged response properties of individual neurons is certainly informative but it is not sufficient to characterize the neural code of a brain area. Critically, the role of the DG and the CA1 area of the hippocampus should be revisited in light of our observations. The methods that we propose will shed new light on the general role of other brain areas implicated in high level cognitive functions such as spatial navigation and in which place cells are not observed.

## STAR ★ METHODS

### Resource Availability

**Lead contact—**Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Stefano Fusi (sf2237@columbia.edu).

**Materials Availability—**This study did not generate new unique reagents.

**Data and Code Availability—**The datasets and analysis code supporting the current study are available from the lead contact on request.

### Experimental Model and Subject Details

**Mice—**All procedures were conducted in accordance with the U.S. NIH Guide for the Care and Use of Laboratory Animals and the institutional Animal Care and Use Committees at New York State Psychiatric Institute and UCSF. Adult male C57BL/6J mice were supplied by Jackson Laboratory and were used beginning at 8–12 weeks of age. Mice were co-housed with litter mates (2–5 per cage). Mice were maintained with unrestricted access to food and water on a 12-hour light/dark cycle.

**Viral Constructs—**For calcium imaging, AAVdj-CaMKII-GCaMP6m was packaged and supplied by Stanford Vector Core at titers of ~ $4X10^{12}$ vg/ml, and AAV1-Syn-GCaMP6f.WPRE.SV40 was packaged and supplied by U Penn Vector Core at titers of ~ $2X10^{12}$ vg/ml.

## Method Details

**Calcium imaging.—**Mice were prepared for in vivo calcium imaging as previously described (Resendez et al., 2016). For dorsal DG imaging, mice were injected with a virus encoding GCaMP6m (AAVdj-CaMKII-GCaMP6m) at the following coordinates: −1.95AP, 1.4ML, 2.2, 2.1, 2.0, 1.9 DV, ~ 90 nl per site) and a ~ 1.0 mm diameter, ~ 4 mm long GRIN lens (Inscopix, Palo Alto, CA) was implanted at (−2.0AP, −1.4ML, −1.95 DV). For dorsal CA1 imaging, mice were injected with a virus encoding GCaMP6f (AAV1-Syn-GCaMP6f.WPRE.SV40) at the following coordinates: (-2.15AP, 1.85ML, −1.55, −1.65DV, 256nl per site) and a GRIN lens was implanted at (-2.15AP, 1.30ML, −1.30DV). Three weeks after surgery, mice were checked for GCaMP expression with a miniaturized microscope (Inscopix, Palo Alto, CA) with procedures previously described (Resendez et al., 2016). Anesthetized mice were checked for GCaMP+ neurons and a baseplate was attached to the skull at the optimal imaging plane. For all the mice presented in this report the histology confirmed the adequate placement of the lens (Fig. S22). For dorsal DG imaging, one week later, mice were imaged during foraging in an open field task and were habituated to the room and enclosure (30min), then 24 hours later they were imaged as they foraged for sucrose pellets in an open field enclosure (50cm $^2$). For dorsal CA1 imaging, mice were imaged during exploration of an open field enclosure. Mice were habituated to the room and enclosure (10 minutes) and then imaged 30 minutes later. Imaging frames were recorded with nVista acquisition software (Inscopix, Palo Alto, CA), and time-synced behaviour was acquired using EthoVision XT 10. Calcium imaging videos were acquired at 15 frames per second with 66.56 ms exposure.

## Quantification and Statistical Analysis

**Behaviour data pre-processing.—**The behaviour was recorded using a webcam (Logitech) mounted on the ceiling about 3 feet above the arena. The instantaneous position of the animal was then extrapolated from the video using custom code written in Python using the Scikit-image library (version 0.13.0). We first applied a 9 points piecewise affine transformation to correct for barrel camera distortions. We then applied a smoothing filter with a Gaussian profile to reduce the effect of pixel intensity noise due to low lighting and low image resolution and applied a threshold to the gray-scale converted image to get a few contiguous regions of pixels as candidate animal tracking. We then used a method based on the determinant of the Hessian to identify blobs in the pre-processed images and verified that the largest blob was consistently found to be corresponding to the animal silhouette. Hence, we used the centre of the largest blob as the tracked position of the mouse. We further temporally aligned the position data to the imaging data using linear interpolation and smoothed them with a 7 frames time window. Lastly, we identified the time bins in which the speed of the animal was lower than 2 cm/s for more than 1 s and discarded them from the analysis, unless specified.

**Signal extraction and spike deconvolution.—**All calcium movies were initially processed in Mosaic (Inscopix, Palo Alto, CA) for spatial binning and motion correction and subsequently analysed using a recently developed software algorithm written in Matlab (Mathworks) called CNMF-e (Zhou et al., 2018). Briefly, the algorithm separates the large, low-frequency fluctuating background components from the signal produced by of

multiple sources in the data, allowing the accurate source extraction of cellular signals. It involves a constrained non-negative matrix factorization problem optimized for endoscopic data whereby calcium temporal dynamics and the shape of spatial footprints are used as constraints. It includes 3 main steps which are iterated: obtain a first estimate of spatial and temporal components of single neurons without direct estimation of the background; estimate the background given the estimated neurons' spatio-temporal activity; update the spatial and temporal components of all neurons while fixing the estimated background fluctuations. In each of these steps, manual intervention guided by visual inspection based on temporal profile and spatial footprint shape allowed to further improve the quality of the signal extraction. The result of this process consists of a list of deconvolved calcium events for each cell with associated time-stamp and magnitude and the convolved trace with a calcium decay profile estimated for each cell independently on the basis of the raw trace.

For our decoding analysis, we did not use the original traces, rather we used the events extracted with CNMF-e convolved with an exponential kernel. The time constant of the kernel was optimized to maximize the cross-validated position decoding performance and was equal for all neurons. The results depend only weakly on the kernel time constant, and qualitatively are the same (see Fig. S1). All other quantities derived from the calcium traces were computed using the calcium events, unless specified otherwise, and therefore their values do not depend on the shape of the kernel.

**Place fields and heading direction tuning.—**Place fields for each extracted source were constructed in a manner similar to established method applied to electrophysiology data (Leutgeb et al., 2007). We used the calcium events of each cell as its putative spiking activity. We then summed the total number of events that occurred in a given location, divided by the amount of time the animal spent in the location and smoothed using a Gaussian kernel centered on each bin. The rate in each location × was estimated as

$$r(x) = \frac{\sum_{i=1}^{n} g\left(\frac{s_i - x}{h}\right)}{\int_0^T g\left(\frac{y(t) - x}{h}\right) dt}$$

where $g$ is a Gaussian smoothing kernel, $h = 5$ sets the spatial scale for smoothing, $n$ is the number of events, $s_i$ is the location of the $i$-th event, $y(t)$ the location of the animal at time $t$ and $[0, T)$ the period of the recording. In this and all subsequent analysis we removed the time bins in which the animal had a speed of less than 2 cm/s for more than 1 s, unless specified otherwise. Similarly, for heading direction tuning, we first discretized the directions of motion into 8 angular bins of 45 degrees each and then computed the mean event rate for each cell in each of the 8 bins.

**Spatial information statistics.—**To quantify the statistical significance of the rate maps we measured their specificity in terms of the information content of cell activity (Skaggs et al., 1993; Danielson et al., 2017; Allegra et al., 2019). We used a 16×16 square grid and computed the amount of Shannon information that a single event conveys about the animal's location. The spatial information content of cell discharge was calculated as a mutual

information score between event occurrence per cell and animal position or equivalently using the formula:

$$SI = \sum_{i=1}^{N} p_i \frac{r_i}{r} \log_2 \frac{r_i}{r}$$

where $i$ is the spatial bin number, $p$ is the probability for occupancy of bin $i$, $r_i$ is the mean event rate at bin $i$ and $r$ is the overall mean event rate. We applied the same formula to the direction of motion after discretizing the full angle to 8 bins of 45 degrees. For both measures, we corrected for the sampling bias problem in information measures (Panzeri et al., 2007) using shuffled distributions of event occurrences as follows. For each cell independently, we discretized time, generating a long vector of 0's (no event) and 1's (event). We then randomly permuted the elements of this vector and for each permutation we computed the resulting spatial information. We repeated this procedure 1000 times, therefore obtaining 1000 values of spatial information to which we compared the original information content (Ziv et al., 2013; Danielson et al., 2017; Meshulam et al., 2017; Allegra et al., 2019). We labeled a cell as place cells or a heading direction cell if the original value of spatial information exceeded 3 sigmas from the shuffled distribution (see also S2 and S3 and Table S1).

**Decoding position.—**For all the datasets, unless otherwise specified, we used 10-fold cross validation to validate the performance of the decoders. We divided the trial in 10 temporally contiguous periods of equal size in terms of number of datapoints after excluding datapoints corresponding to immobility. We then trained the decoders using the data from 9 of them and tested on the remaining data. To decode the position of the animal, we first divided the arena into 8×8 equally sized, squared locations. We then assigned at each time bin the label of the discrete location in which the animal was found. For each pair of locations, we trained a Support Vector Machine (SVM) classifier (Cortes and Vapnik, 1995) with a linear kernel to classify the cell activities into either one of the two assigned locations using all the identified cells, unless specified otherwise. We used only the data corresponding to the two assigned locations and to correct for unbalanced data due to inhomogeneous exploration of the arena we balanced the classes with weights inversely proportional to the class frequencies (Pedregosa et al., 2012). The output of the classifiers was then combined to identify the location with the largest number of votes as the most likely location (Bishop, 2006). For each choice of train and test set, we computed the median decoding error as the median of the physical distance between the centre of the decoded discrete location and the actual position of the mouse in each time bin of the test set, unless otherwise specified. The final decoding performance was then computed as the mean of all the median errors across the different choices of train and test sets.

**Chance level decoding performance—**To assess the statistical significance of our decoders, we computed chance distributions of decoding errors from shuffled data. This can be done in different ways and we chose a conservative procedure that maintained some structure of the data while destroying the relation between the behavior, e.g., the animal's position, and the calcium event time series. Briefly, we discretized time obtaining a vector

of positions (or other behavioral variables). We then flipped this vector in time (e.g., the last data point of position became the first datapoint and vice versa) an then shifted the whole vector in time by a random amount in a torus, i.e., points that went beyond the time limits of the data were reinserted from the other side. This procedure destroys the relation between behavior and neural activity, but preserves the time correlations of both the time series representing behavior and, of course, the time series of the neural activity (which remains untouched). For each random shift, we trained a new decoder on the data and pooled all the errors obtained. We finally assessed the statistical significance of the decoding error for the 10-fold cross-validation of the original data by comparing it to the distribution of errors obtained from the manipulated data using the non-parametric Mann-Whitney U test, from which we obtained a p-value of significance. We implicitly assumed that the 10 folds are statistically independent (the 10 testing time intervals considered for the 10 folds did not have any overlap). This is the procedure we used in all our figures unless specified otherwise.

Another less conservative shuffling strategy is to manipulate the calcium events. We assigned a random time bin to each calcium event for each cell independently while maintaining the overall density of calcium events across all cells, i.e., by choosing only time bins in which there were calcium events in the original data and keeping the same number and magnitude of the events in each time bin. This method destroys spatial information as well as temporal correlations but keeps the overall activity across cells. We verified that our results did not depend on the particular strategy adopted (see Fig. S3 and Table S1).

**Decoding the direction of motion.—**One behaviourally relevant quantity that was available to us was the direction of motion of the animal. Unfortunately, the visual tracking didn’t allow for a direct estimate of the direction of motion. The head direction was also not easily measurable so we resorted to using the positional information to extract the direction of motion. We computed it by using two subsequent datapoints in the animal x-y trajectory. We discretized the values into 8 angles and then applied similar decoding strategies as for position decoding, i.e., we used a battery of linear-kernel SVM decoders to distinguish between pairs of angles after balancing the dataset through class weighting. We report the median error in radiant on the left-out data of the 10-fold cross validation. We applied the methods described above for position decoding for assessing the statistical significance of the results.

**Decoding speed.—**To decode the speed of movement of the animal we first computed the speed of motion using two consecutive positions and assigned the computed speed to the later time bin among the two. To decode the instantaneous speed of motion we used Lasso (Tibshirani, 1996), a linear regression analysis method that minimizes the sum of squared errors while selecting a subset of the input cells to improve decoding accuracy and interpretability of the results. We applied the methods described above for position decoding for assessing the statistical significance of the results.

#### 0.0.0.15. Bayesian decoder.

The Bayesian decoder is a theoretical optimal probabilistic method to decode information for the activity of the neural population. It is based on the Bayes rule and has been extensively used to decode position from electrophysiological data from the hippocampus (Zhang et al., 1998; Wilson and McNaughton, 1993). Briefly, if $x$ is a discrete position in the arena, we estimate the position using:

$$P(x|r_i) = P(r_i|x)P(x)/P(r_i)$$

where $r_t$ is the activity of the population at time $t$ and assuming independent activity of different neurons. The algorithm computes $P(x \mid r_t)$ for all discrete positions and assigns the predicted position to the one that maximises it:

$$\hat{x}_t = \underset{x}{\operatorname{argmax}} P(x|r_t).$$

**Importance index.—**The importance index was introduced to quantify the contribution of each cell in a population to the decoding of a given quantity. We applied a modified version of a traditional method for feature selection in machine learning. In our analysis, a feature of the input space consists of one DG cell. Feature selection is performed using the weights of the decoder after fitting model to the data. In our case, since we employed multiple decoders, one for each pair of physical location in the arena, we introduced a method to combine the weights assigned to the cells by each decoder. We defined the importance index of cell $i$ as:

$$\omega_i = \sum_k \frac{|w_{ik}|}{\sum_j |w_{jk}|}$$

where $w_{ik}$ is the weight of the $k$-th decoder assigned to the $i$-th cell (and equivalently $w_{jk}$ is the weight of the $k$-th decoder assigned to the $j$-th cell). The indices $i$, $j$ run through all cells in the population and $k$ runs through all the binary decoders.

**Procedure to destroy correlations.—**To destroy correlations without impacting the spatial information of single neurons, we considered multiple passes through single discrete locations in the arena. We then shuffled the calcium event occurrences between different passes in the same location. Importantly, we corrected the activity of each pass for the different amount of time spent in each pass by radomly sampling events instead of replacing them in order to reduce artifacts. We verified that the correction does not impact decoding when sampling from the same pass (see Fig. S14).

**Software—**The data analysis has been performed using custom code written in Python (version 2.7.12) and routines from the Scipy (ver. 0.19.0), Numpy (ver. 1.11.3) and the Scikit-learn (0.19.1) (Pedregosa et al., 2012) packages. The source extraction has been performed using Matlab (Mathworks, R2016a) and CNMF-e (Zhou et al., 2018) using the

same parameters across animals and minimal manual intervention only for obvious non-cell like sources based on spatial profile shape and temporal profile dynamics.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

This article is dedicated to Howard Eichenbaum. We thank Loren Frank, Surya Ganguli, Liam Paninski, Daniel Salzman, Attila Losonczy, Misha Tsodyks, Ran Rubin, Mattia Rigotti and Marcus Benna for insightful discussions and comments on an earlier version of this work and Pengcheng Zhou for his help with the CNMF-e software. We are very grateful to the anonymous reviewers for suggesting important controls that significantly improved the quality of this work. F.S. was supported by NSF's NeuroNex program award DBI-1707398, by the Swiss National Science Foundation's Early Postdoc Mobility grant P2EZP2_155561 and by the Kavli Foundation. S.F. was supported by NSF's NeuroNex program award DBI-1707398, the Gatsby Charitable Foundation, the Simons Foundation, the Schwartz foundation, the Kavli foundation. M.A.K. was supported by NIMH (R01 MH108623, R01 MH111754, and R01 MH117961), a One Mind Rising Star Award, the Human Frontier Science Program, the Esther A. and Joseph Klingenstein Fund, the McKnight Memory and Cognitive Disorders Award, and the Pew Charitable Trusts. J.H.J. was supported by the Helen Hay Whitney Foundation.

## References

Abbott LF, Dayan P, 1999. The effect of correlated variability on the accuracy of a population code. Neural computation 11, 91–101. [PubMed: 9950724]

Acharya L, Aghajan ZM, Vuong C, Moore JJ, Mehta MR, 2016. Causal influence of visual cues on hippocampal directional selectivity. Cell 164, 197–207. [PubMed: 26709045]

Agarwal G, Stevenson IH, Berényi A, Mizuseki K, Buzsáki G, Sommer FT, 2014. Spatially distributed local fields in the hippocampus encode rat position. Science 344, 626–630. [PubMed: 24812401]

Allegra M, Posani L, Schmidt-Hieber C, 2019. The hippocampus as a perceptual map: neuronal and behavioral discrimination during memory encoding. bioRxiv, 868794.

Averbeck BB, Lee D, 2006. Effects of noise correlations on information encoding and decoding. Journal of neurophysiology 95, 3633–3644. [PubMed: 16554512]

Benna MK, Fusi S, 2019. Are place cells just memory cells? memory compression leads to spatial tuning and history dependence. bioRxiv, 624239.

Bernardi S, Benna MK, Rigotti M, Munuera J, Fusi S, Salzman D, 2018. The geometry of abstraction in hippocampus and prefrontal cortex. bioRxiv, 408633.

Bishop CM, 2006. Pattern Recognition and machine learning. Springer New York.

Brody CD, 1999. Correlations without synchrony. Neural computation 11, 1537–1551. [PubMed: 10490937]

Cortes C, Vapnik V, 1995. Support-vector networks. Machine learning 20, 273–297.

Danielson NB, Kaifosh P, Zaremba JD, Lovett-Barron M, Tsai J, Denny CA, Balough EM, Goldberg AR, Drew LJ, Hen R, Losonczy A, Kheirbek MA, 2016. Distinct Contribution of Adult-Born Hippocampal Granule Cells to Context Encoding. Neuron 90, 101–112. doi:10.1016/j.neuron.2016.02.019. [PubMed: 26971949]

Danielson NB, Turi GF, Ladow M, Chavlis S, Petrantonakis PC, Poirazi P, Losonczy A, 2017. In vivo imaging of dentate gyrus mossy cells in behaving mice. Neuron 93, 552–559. [PubMed: 28132825]

Dayan P, 1993. Improving generalization for temporal difference learning: The successor representation. Neural Computation 5, 613–624.

van Dijk MT, Fenton AA, 2018. On how the dentate gyrus contributes to memory discrimination. Neuron 98, 832–845. [PubMed: 29731252]

Dipoppa M, Ranson A, Krumin M, Pachitariu M, Carandini M, Harris KD, 2018. Vision and locomotion shape the interactions between neuron types in mouse visual cortex. Neuron 98, 602–615. [PubMed: 29656873]

Dombeck DA, Harvey CD, Tian L, Looger LL, Tank DW, 2010. Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. Nature neuroscience 13, 1433. [PubMed: 20890294]

Dombeck DA, Khabbaz AN, Collman F, Adelman TL, Tank DW, 2007. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. Neuron 56, 43–57. [PubMed: 17920014]

Eichenbaum H, 2018. Barlow versus hebb: When is it time to abandon the notion of feature detectors and adopt the cell assembly as the unit of cognition? Neuroscience letters 680, 88–93. [PubMed: 28389238]

Eyherabide HG, Samengo I, 2013. When and why noise correlations are important in neural decoding. Journal of Neuroscience 33, 17921–17936. [PubMed: 24198380]

Fenton AA, Kao HY, Neymotin SA, Olypher A, Vayntrub Y, Lytton WW, Ludvig N, 2008. Unmasking the ca1 ensemble place code by exposures to small and large environments: more place cells and multiple, irregularly arranged, and expanded place fields in the larger space. Journal of Neuroscience 28, 11250–11262. [PubMed: 18971467]

Fenton AA, Muller RU, 1998. Place cell discharge is extremely variable during individual passes of the rat through the firing field. Neurobiology Communicated by Jan Bures Czech Academy of Sciences 95, 3182–3187.

Fusi S, Miller EK, Rigotti M, 2016. Why neurons mix: High dimensionality for higher cognition. Current Opinion in Neurobiology 37, 66–74. URL: 10.1016/j.conb.2016.01.010, doi:10.1016/j.conb.2016.01.010. [PubMed: 26851755]

Gershman SJ, Moore CD, Todd MT, Norman KA, Sederberg PB, 2012. The successor representation and temporal context. Neural Computation 24, 1553–1568. [PubMed: 22364500]

Gini C, 1912. Variabilità e mutabilità. Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini T). Rome: Libreria Eredi Virgilio Veschi.

Gluck MA, Myers CE, 1993. Hippocampal mediation of stimulus representation: A computational theory. Hippocampus 3, 491–516. [PubMed: 8269040]

GoodSmith D, Chen X, Wang C, Kim SH, Song H, Burgalossi A, Christian KM, Knierim JJ, 2017. Spatial representations of granule cells and mossy cells of the dentate gyrus. Neuron 93, 677–690. [PubMed: 28132828]

Hardcastle K, Maheswaranathan N, Ganguli S, Giocomo LM, 2017. A Multiplexed, Heterogeneous, and Adaptive Code for Navigation in Medial Entorhinal Cortex. Neuron 94, 375–387.e7. URL: 10.1016/j.neuron.2017.03.025, doi:10.1016/j.neuron.2017.03.025. [PubMed: 28392071]

Harvey CD, Collman F, Dombeck DA, Tank DW, 2009. Intracellular dynamics of hippocampal place cells during virtual navigation. Nature 461, 941. [PubMed: 19829374]

Haufe S, Meinecke F, GÃ¶rgen K, DÃ¶hne S, Haynes JD, Blankertz B, BieÃ mann F, 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. NeuroImage 87, 96–110. URL: 10.1016/j.neuroimage.2013.10.067, doi:10.1016/j.neuroimage.2013.10.067. [PubMed: 24239590]

Keinath AT, Wang ME, Wann EG, Yuan RK, Dudman JT, Muzzio IA, 2014. Precise spatial coding is preserved along the longitudinal hippocampal axis. Hippocampus 24, 1533–1548. [PubMed: 25045084]

Kelemen E, Fenton AA, 2010. Dynamic grouping of hippocampal neural activity during cognitive control of two spatial frames. PLoS biology 8, e1000403. [PubMed: 20585373]

Kinsky NR, Sullivan DW, Mau W, Hasselmo ME, Eichenbaum HB, 2018. Hippocampal place fields maintain a coherent and flexible map across long timescales. Current Biology 28, 3578–3588. [PubMed: 30393037]

Kriegeskorte N, Douglas PK, 2019. Interpreting encoding and decoding models. Current opinion in neurobiology 55, 167–179. [PubMed: 31039527]

Latham PE, Nirenberg S, 2005. Synergy, redundancy, and independence in population codes, revisited. Journal of Neuroscience 25, 5195–5206. [PubMed: 15917459]

Leutgeb JK, Leutgeb S, Moser MB, Moser EI, 2007. Pattern separation in the dentate gyrus and ca3 of the hippocampus. science 315, 961–966. [PubMed: 17303747]

Lindsay GW, Rigotti M, Warden MR, Miller EK, Fusi S, 2017. Hebbian learning in a random network captures selectivity properties of prefrontal cortex. Journal of Neuroscience, 1222–17.

Meshulam L, Gauthier JL, Brody CD, Tank DW, Bialek W, 2017. Collective behavior of place and non-place neurons in the hippocampal network. Neuron 96, 1178–1191. [PubMed: 29154129]

Mladenić D, Brank J, Grobelnik M, Natasa Milic-Frayling I, 2004. Feature Selection using Linear Classifier Weights: Interaction with Classification Models. Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, 234–241 doi:10.1145/1008992.1009034.

Moser EI, Kropff E, Moser MB, 2008. Place cells, grid cells, and the brain's spatial representation system. Annual review of neuroscience 31.

Mukamel EA, Nimmerjahn A, Schnitzer MJ, 2009. Automated analysis of cellular signals from large-scale calcium imaging data. Neuron 63, 747–760. [PubMed: 19778505]

Neunuebel JP, Knierim JJ, 2012. Spatial firing correlates of physiologically distinct cell types of the rat dentate gyrus. Journal of Neuroscience 32, 3848–3858. [PubMed: 22423105]

Nitz D, McNaughton B, 2004. Differential modulation of ca1 and dentate gyrus interneurons during exploration of novel environments. Journal of neurophysiology 91, 863–872. [PubMed: 14523073]

Olypher A, Lánskỳ P, Muller R, Fenton A, 2003. Quantifying location-specific information in the discharge of rat hippocampal place cells. Journal of neuroscience methods 127, 123–135. [PubMed: 12906942]

Pachitariu M, Stringer C, Dipoppa M, Schröder S, Rossi LF, Dalgleish H, Carandini M, Harris KD, 2017. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. Biorxiv.

Panzeri S, Senatore R, Montemurro MA, Petersen RS, 2007. Correcting for the Sampling Bias Problem in Spike Train Information Measures. J Neurophysiol 98, 1064–1072. doi:10.1152/jn.00559.2007. [PubMed: 17615128]

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É, 2012. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830. URL: http://dl.acm.org/citation.cfm?id=2078195%5Cnhttp://arxiv.org/abs/1201.0490, doi:10.1007/s13398-014-0173-7.2, arXiv:1201.0490.

Pernía-Andrade AJ, Jonas P, 2014. Theta-gamma-modulated synaptic currents in hippocampal granule cells in vivo define a mechanism for network oscillations. Neuron 81, 140–152. [PubMed: 24333053]

Pfeiffer BE, Foster DJ, 2013. Hippocampal place-cell sequences depict future paths to remembered goals. Nature 497, 74. [PubMed: 23594744]

Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky E, Simoncelli EP, 2008. Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature 454, 995. [PubMed: 18650810]

Pnevmatikakis EA, Soudry D, Gao Y, Machado TA, Merel J, Pfau D, Reardon T, Mu Y, Lacefield C, Yang W, et al. , 2016. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. Neuron 89, 285–299. [PubMed: 26774160]

Recanatesi S, Farrell M, Lajoie G, Deneve S, Rigotti M, Shea-Brown E, 2018. Signatures and mechanisms of low-dimensional neural predictive manifolds. bioRxiv, 471987.

Resendez SL, Jennings JH, Ung RL, Namboodiri VMK, Zhou ZC, Otis JM, Nomura H, McHenry JA, Kosyk O, Stuber GD, 2016. Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. Nature protocols 11, 566. [PubMed: 26914316]

Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S, 2013. The importance of mixed selectivity in complex cognitive tasks. Nature 497, 1–6. URL: http://www.ncbi.nlm.nih.gov/pubmed/23685452 http://dx.doi.org/10.1038/nature12160%5Cnpapers3://publication/doi/10.1038/nature12160, doi:10.1038/nature12160.

Saxena S, Cunningham JP, 2019. Towards the neural population doctrine. Current opinion in neurobiology 55, 103–111. [PubMed: 30877963]

Schneidman E, Berry MJ II, Segev R, Bialek W, 2006. Weak pairwise correlations imply strongly correlated network states in a neural population. Nature 440, 1007. [PubMed: 16625187]

Schneidman E, Bialek W, Berry MJ, 2003. Synergy, redundancy, and independence in population codes. Journal of Neuroscience 23, 11539–11553. [PubMed: 14684857]

Skaggs WE, McNaughton BL, Gothard KM, 1993. An information-theoretic approach to deciphering the hippocampal code, in: Advances in neural information processing systems, pp. 1030–1037.

Stachenfeld KL, Botvinick M, Gershman SJ, 2014. Design principles of the hippocampal cognitive map, in: Advances in neural information processing systems, pp. 2528–2536.

Stachenfeld KL, Botvinick MM, Gershman SJ, 2017. The hippocampus as a predictive map. bioRxiv, 097170.

Talbot ZN, Sparks FT, Dvorak D, Curran BM, Alarcon JM, Fenton AA, 2018. Normal ca1 place fields but discoordinated network discharge in a fmr1-null mouse model of fragile × syndrome. Neuron 97, 684–697. [PubMed: 29358017]

Tang S, Zhang Y, Li Z, Li M, Liu F, Jiang H, Lee TS, 2018. Large-scale two-photon imaging revealed super-sparse population codes in the v1 superficial layer of awake monkeys. eLife 7, e33370. [PubMed: 29697371]

Tibshirani R, 1996. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological) 57, 267–288.

Whittington JC, Muller TH, Mark S, Chen G, Barry C, Burgess N, Behrens TE, 2019. The tolman-eichenbaum machine: Unifying space and relational memory through generalisation in the hippocampal formation. bioRxiv, 770495.

Wilson MA, McNaughton BL, 1993. Dynamics of the hippocampal ensemble code for space. Science 261, 1055–1058. [PubMed: 8351520]

Wood ER, Dudchenko PA, Eichenbaum H, 1999. The global record of memory in hippocampal neuronal activity. Nature 397, 613–616. [PubMed: 10050854]

Zenke F, Poole B, Ganguli S, 2017. Continual learning through synaptic intelligence, in: Precup D, Teh YW (Eds.), Proceedings of the 34th International Conference on Machine Learning, PMLR, International Convention Centre, Sydney, Australia. pp. 3987–3995. URL: http://proceedings.mlr.press/v70/zenke17a.html.

Zhang K, Ginzburg I, McNaughton BL, Sejnowski TJ, 1998. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. Journal of neurophysiology 79, 1017–1044. [PubMed: 9463459]

Zhou P, Resendez SL, Rodriguez-Romaguera J, Jimenez JC, Neufeld SQ, Giovannucci A, Friedrich J, Pnevmatikakis EA, Stuber GD, Hen R, et al. , 2018. Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. eLife 7, e28728. [PubMed: 29469809]

Ziv Y, Burns LD, Cocker ED, Hamel EO, Ghosh KK, Kitch LJ, El Gamal A, Schnitzer MJ, 2013. Long-term dynamics of ca1 hippocampal place codes. Nature neuroscience 16, 264. [PubMed: 23396101]

- Position, direction of motion and speed are encoded in CA1 and DG neurons
- Single-cell tuning only weakly predicts a cell's importance for position encoding
- Non-place cells contribute to position encoding
- Disrupting correlations between neurons leads to decreased decoding accuracy in CA1

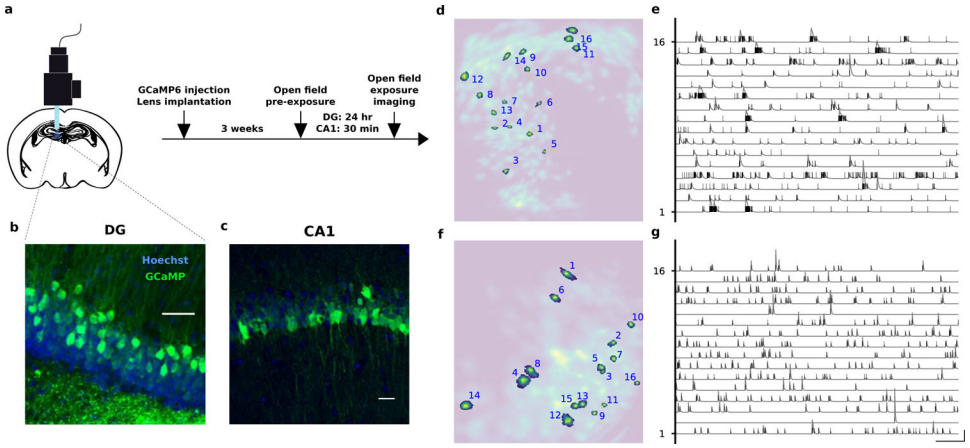


**Figure 1.**

**a**) Experiment protocol. Mice were anesthetized with isoflurane and placed in a stereotactic apparatus. DG mice were then injected in the dorsal DG with a virus encoding GCaMP6m. Ca1 mice were injected with GCaMP6f. Mice were then implanted with a GRIN lens and a baseplate was attached to the skull at the optimal imaging places. Three weeks after surgery they were checked for GCaMP expression with a miniaturized microscope (Inscopix, Palo Alto, CA) and procedures previously described (Resendez et al., 2016). Recording site. The imaging plane was later assessed through histology. **b**) DG recording site. GCL, granule cell layer; SGZ, subgranular zone. **c**) CA1 recording site. **d–g**) Automated signal extraction using CNMF-e (Zhou et al., 2018). The algorithm identifies the spatial (left) and temporal (right) components of the signal sources, i.e., putative cells. It uses a generative model of Calcium traces and non-negative matrix factorization to separate actual signal sources from the background due to diffused neuropil fluorescence. The extracted spatial components are displayed in **d** (DG) and **f** (CA1), where a few representative ones are highlighted. The corresponding signals are shown in panels **e** (DG) and **g** (CA1) in which vertical ticks correspond to the times of the inferred Calcium events and gray lines to the temporal profiles (see Fig. S1). In line with electrophysiology studies, dentate gyrus granule cells are sparsely active but often in bursts (Pernía-Andrade and Jonas, 2014). Scale bars are 1 min and 1 sd.

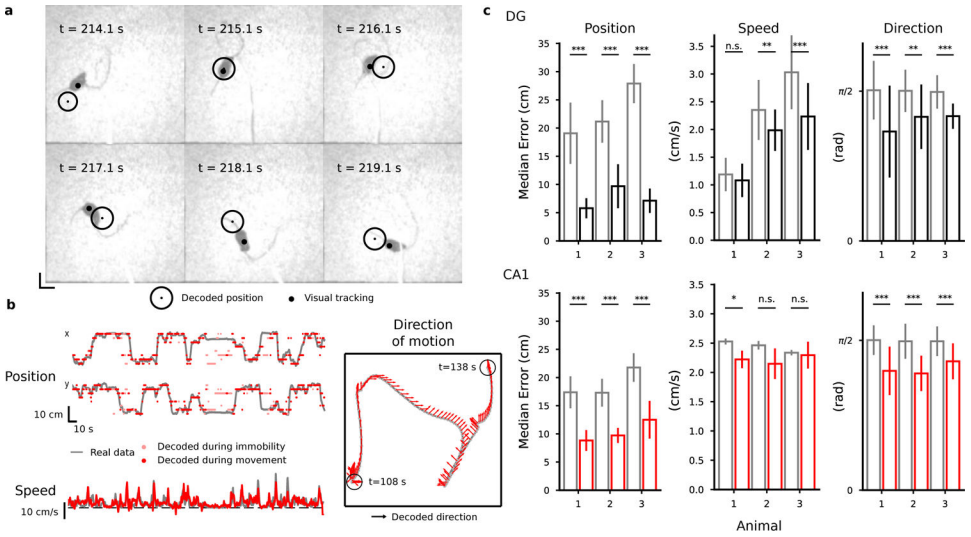


**Figure 2.**

Decoding position, speed and direction of motion. **a**, **b**) Decoding results for a representative DG mouse. See Supplemental Material online for a decoding video. **a**) Selected frames of a video showing the arena and a DG animal from above. The black circle represents the mouse's actual position and red circle is the decoded position, obtained with a probabilistic decoder that reads the activity of 317 DG cells (see Methods). Neural activity has been pre-processed to identify putative calcium events as explained in the Methods. **b**) Examples of decoding position, speed and direction of motion (DG representative mouse). Grey lines correspond to the real values of position and speed variables in the top left and lower left panels respectively while the red dots correspond to their decoded values. The time bins marked in light red in for position and direction of movement correspond to moments of immobility that have been excluded from the training data. The grey line in the right panel corresponds to the position of the mouse and the red arrows correspond to the decoded direction of motion in a 30 s time window. **c**) Decoding accuracy (top: DG; bottom: CA1). The decoding error for position and head direction is computed as the median distance between the decoded value in each time bin and the actual value of the decoded variable in the test data. For the direction of motion, the smallest angle between the decoded and the actual value is considered. The red vertical bars corresponding to the mean over the 10-fold cross-validation (error bars correspond to st. dev.). Grey: chance error obtained by decoding from shuffled data in a way that preserves the autocorrelations in the data (see Methods and Fig. S5). Number of cells: 483, 309, 317 in DG mice; 371, 286, 206 in CA1 mice. See also Fig. S4–S7, S12.

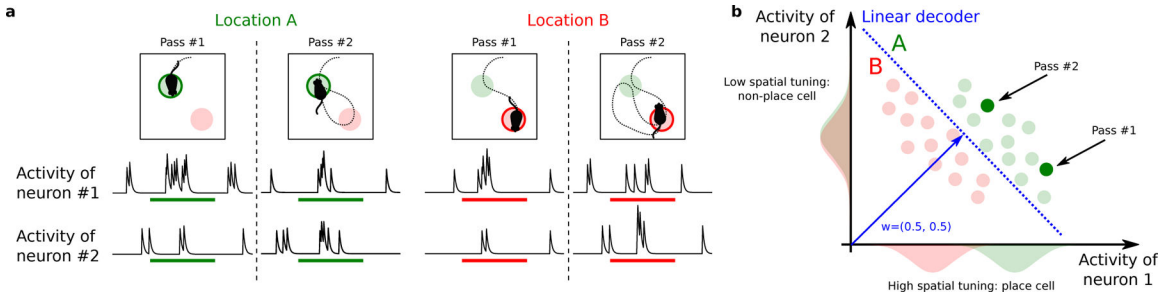


**Figure 3.**

The contribution of untuned cells for encoding position. We show an extreme situation in which one simulated neuron has the same activity distribution when the animal is in two different locations of the arena. Hence the neuron is not selective to position. Nevertheless, for a decoder this neuron can be as important as other selective neurons due to its contribution to the population coding. **a**) Activity of two simulated neurons as a function of time. Top: The simulated animal visits the same discrete location twice (location A in green, location B in red). Bottom: Simulated traces around the time of passage through each location. Different responses for the two neurons are elicited by different experiences, for example due to the different direction of motion. **b**) Example of how place cells and non place-cells can be equally important for encoding the position of the animal. In the scatter plot, the x-axis represents the average activity of the first neuron during one pass and the y-axis is the activity of the second neuron. Each point in the space represents an average population response in a single pass. Their responses are typically highly variable and are scattered around their mean values. The two neurons in the example have very different activity profiles: the first has a strong spatial tuning (place cell) while the second has only a weak tuning. The distributions of their activities in each location, reported along the axis, only partially (neuron 1, place cells) or almost completely overlap (neuron 2). Despite this variability in the single neuron responses, the neural representations at the population level are well separated, making it possible for a linear decoder (blue dashed line) to discriminate them with high accuracy. The resulting decoder's weight vector has two equal components corresponding to the importance of the two neurons in encoding position. In this example both neurons are important for encoding position despite their very different tuning properties.

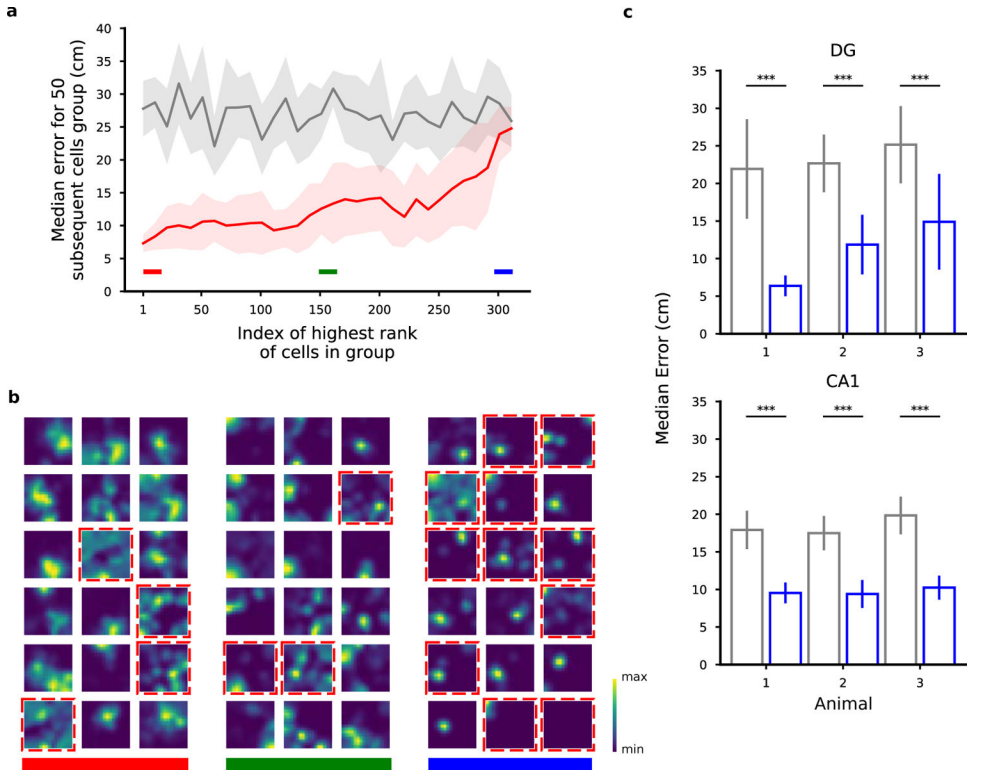


**Figure 4.**

Ranking neurons according to their contribution to the decoding accuracy for position. **a**) Validation of the importance index. In this figure we show the median error for various selections of 50 DG cells from a representative animal ranked by their importance index as obtained using the decoder's weight. Each point in the plot is aligned to the rank of the first cell in the selection (for example, the first dot corresponds to the selection of the first 50 cells from index 1 to index 50; the shaded region represents the standard error for the 10-fold cross-validation). Grey: chance level and standard error. As expected, the median error for the population of the 50 top ranked (best) cells is much smaller than the median error for the worst last (worst) 50 ones. **b**) Spatial tuning maps for groups of 18 cells ordered by importance index. Same cells as in **a**. We ranked the cells using the importance index for position (see Methods). The three groups of best, mid and worst cells are highlighted with the color bands in **a** for reference. The maps are normalized to the peak rate in each map. Dashed red borders indicate cells that don't pass the criteria for place-cells using a commonly used statistical test for tuning (see Methods). Even among the most important cells there appear some non place-cells (and vice versa). Similarly, some place cells appear in the group of cells with medium and low importance. The small fields in the group of low importance cells are due to significantly lower activities (see also Fig. 5). **c**) The position of both DG and CA1 animals can be decoded from the activity of the non-place cells with a performance significantly higher than chance (Mann-Whitney U test, ***$p < 0.001$). Number of cells: 451, 208, 98 in DG mice; 350, 277, 198 in CA1 mice. See also Fig. S2, S3, S8, S10, S24 and Table S1.

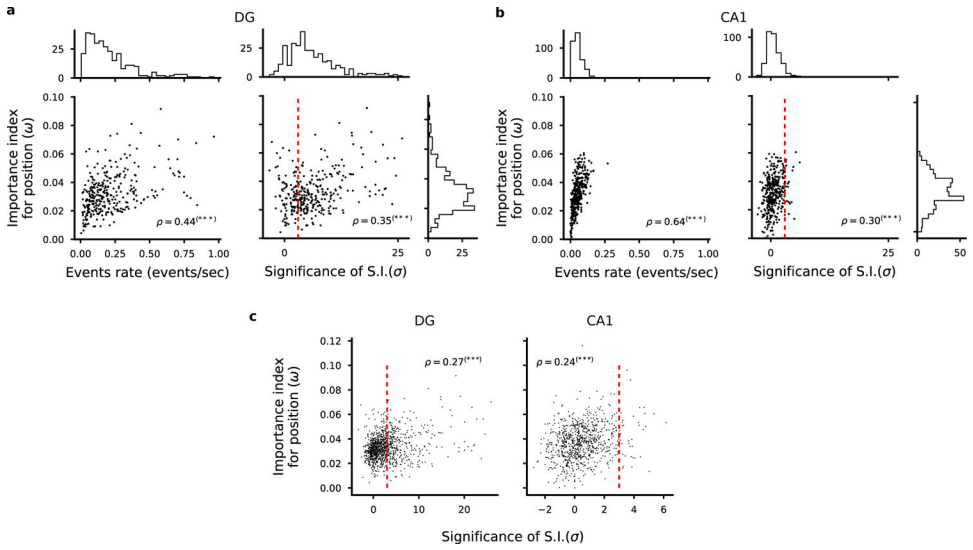


**Figure 5.**

Correlation between importance index and spatial information. **a-b**) Left: Scatter plot of importance index and overall cell activity for each cell in one representative animal. As expected, we found a strong correlation between these quantities because it is unlikely that a weakly active cell can contribute to decoding. Right: Scatter plot of importance index and statistical significance of spatial information with respect to independent random temporal shuffling of each cell's identified calcium events. DG cells in **a**, CA1 cells in **b**. Each dot corresponds to one cell in one representative animal. Pearson's correlation factor $\rho$ between the plotted quantities are reported (***$p < 0.001$). Significant correlations are found between the analysed quantities but single cell statistics only partially capture the information available at the population level. For each quantity, overall histograms are reported on the side of the plot. The dashed red line corresponds to a value of a threshold of 3 used to define place cells (see Methods). **c**) Same plots as in **a** and **b** but for all cells identified in all FOVs in DG (left) and CA1 (right). See also Fig. S9, S13.

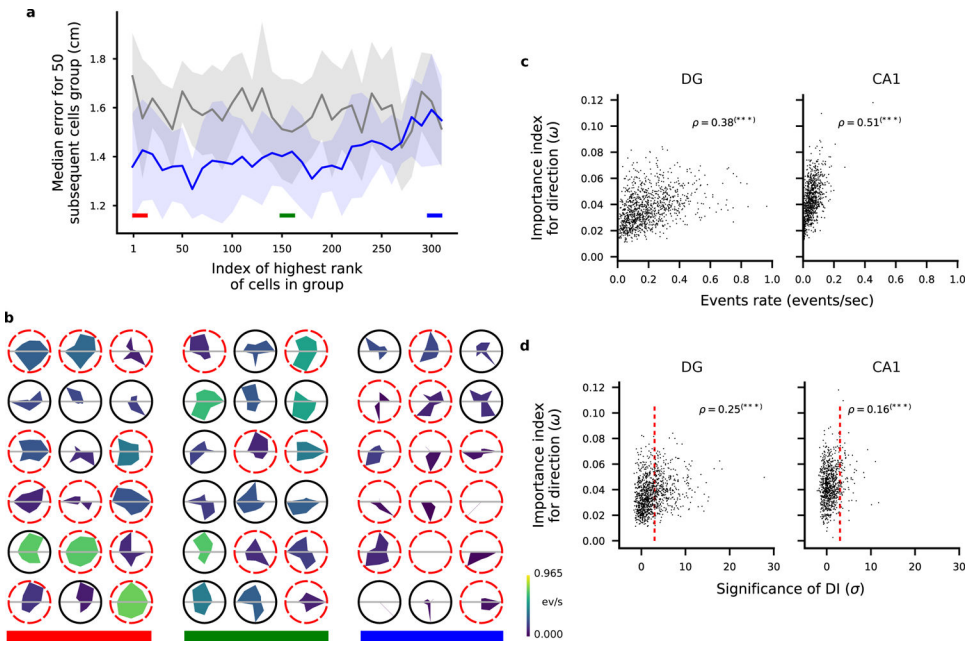


**Figure 6.**

Ranking neurons according to their contribution to the decoding accuracy for head direction. **a**) Validation of the importance index as in Fig. 4a but we ranked the cells according to the importance index for decoding direction of motion (see Methods). **b**) Tuning maps as in Fig. 4b. Here we show the tuning for direction of motion of single cells as polar tuning maps for groups of 18 cells ordered by importance index. The area colour represents the overall activity of the cell throughout the trial. Dashed red borders indicate cells that don't pass the criteria for significant direction tuning using a commonly used statistical test (see Methods). As in the case of position tuning, some untuned cells appear among the most important cells and highly tuned cells appear among the least important. **c**) Scatter plots of cell activity and importance for position decoding for all identified cells combined from all FOVs in DG (left) and CA1 (right). Pearson correlation factor $\rho$ between the plotted quantities are reported (***p < 0.001). **d**) Same as in **c** but for importance index for direction and significance of direction information.

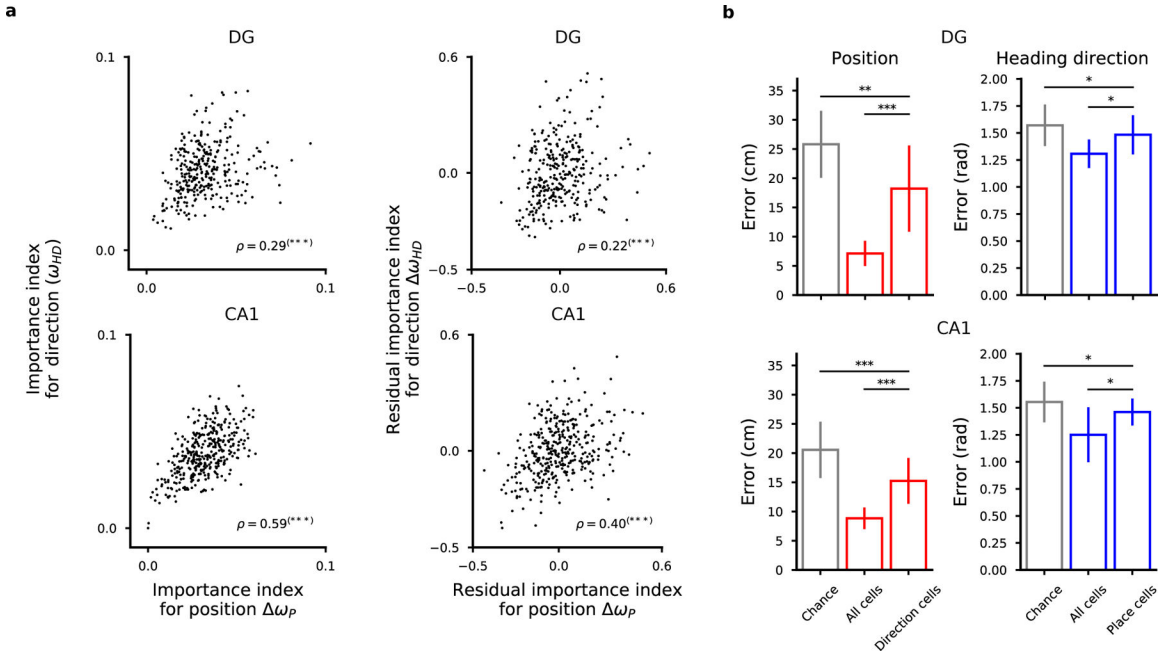


**Figure 7.**

The representations for space and direction of motion are distributed in DG cells and CA1 cells. **a**) Left: Scatter plots of importance index for position and direction of motion (top: DG cells in one representative mouse; bottom: CA1 cells). Each dot corresponds to one cell for which we computed the importance index for the variables we decoded. Pearson's correlation values $\rho$ are reported (***p < 0.001). Right: same as in left but the component of the correlation due to the correlation between importance index and cell activity has been removed from the data. Residuals from linear regression are considered for both quantities. Also the residuals show a positive correlation. **c**, **f**) Even the most important cells for encoding one variable carry information about the other variable. We show the decoding performance of position (left) and direction of motion (right) using the most important cells for direction and position (left and right plots respectively). See also Fig. S7.

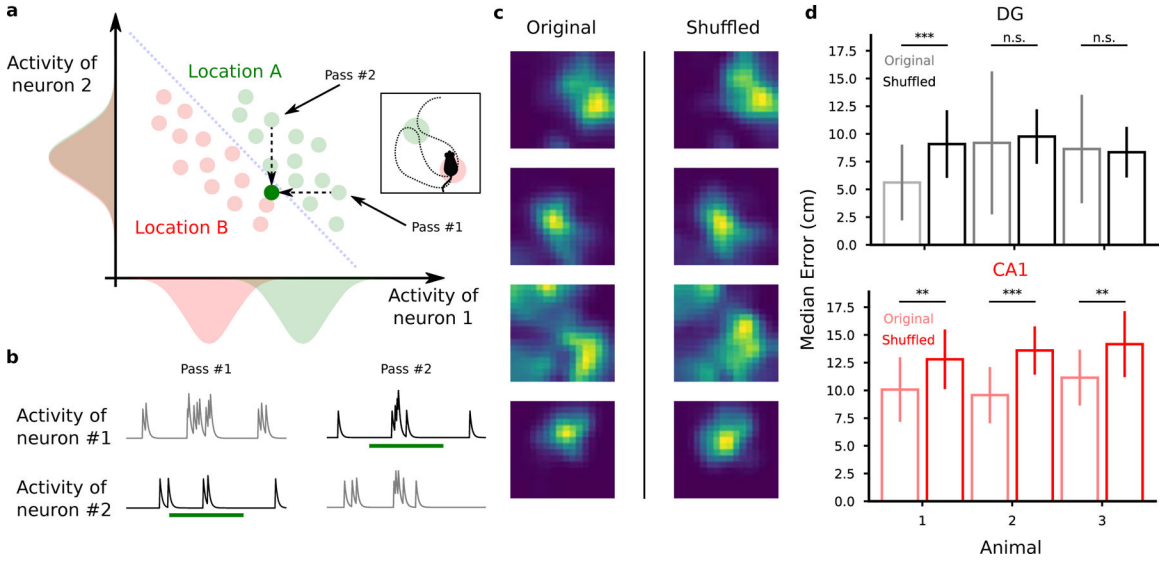


**Figure 8.**

Destroying correlations impacts decoding performance in CA1 but not in DG. **a**) Procedure to test the presence of correlations between cells. We record neural activity during multiple passes through location A (green). Then for we generate a new recording by randomly choosing one of the activities recorded in that location for each cell independently. The green dot below the decoder's discrimination line in the activity plot corresponds to the newly generated activity. We repeat this for all the passes through each location and for each cell independently, therefore destroying the correlations between cells, if any. In the extreme case depicted in the cartoon, this procedure will introduce errors in decoding position since the generated activity will be classified as the wrong location. **b** Cartoon activity traces for the two correlated neurons during the two passes through the same location. As described in **a**, we destroyed correlations by choosing for each neuron the activity during one of the passes through that location and combined them to generate a new activity pattern corresponding to that location. In this example, we chose pass 2 for neuron 1 and pass 1 for neuron 2. **c** Spatial tuning maps of four representative cells before and after the shuffling procedure to destroy correlations. The spatial tuning of the cells remain unaltered after the procedure (mean ± st. dev.; Mann-Withney U test, ***$p < 0.001$, **$p < 0.01$). **d** Decoding performance before (light colors) and after destroying correlations through shuffling (full colors). Top: DG animals. Bottom: CA1 animals. See also Fig. S14–S15, S24.

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Bacterial and Virus Strains | | |
| AAVdj-CaMKII-GCaMP6m | Stanford Vector Core | Cat#GVVC-AAV-89 |
| AAV-DJ-CaMKIIa-GCaMP 6 | Stanford Vector Core | Cat#GVVC-AAV-90 |
| AAV1-Syn-GCaMP6f.WPRE.SV40 | U Penn Vector Core | Cat#AV-1-PV2822 |
| Experimental Models: Organisms/Strains | | |
| C57BL/6J mice | Jackson Laboratory | CAT#000664; RRID:SCR_004633; https://www.jax.org/index.html |
| Software and Algorithms | | |
| Ethovision XT 10 | Noldus | https://www.noldus.com; RRID:SCR_000441 |
| Mosaic | Inscopix | https://www.inscopix.com |
| MATLAB | Mathworks | https://www.mathworks.com/products/matlab; RRID:SCR_001622 |
| CNMF-E | Zhou et al., 2018 | https://github.com/zhoupc/CNMF_E |
| Scikit-learn | Pedregosa et al., 2012 | https://scikit-learn.org |
| Decoding Algorithm | This paper | N/A |
| Spatial information | Skaggs et al., 1993 | N/A |
| | | |
| | | |