Neuron

A Distributed Neural Code in the Dentate Gyrus and in CA1

Highlights

- 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

Authors

Fabio Stefanini, Lyudmila Kushnir, Jessica C. Jimenez, ..., Mazen A. Kheirbek, René Hen, Stefano Fusi

Correspondence

mazen.kheirbek@ucsf.edu (M.A.K.), rh95@cumc.columbia.edu (R.H.), sf2237@columbia.edu (S.F.)

In Brief

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 that cells with easily interpretable responses are not necessarily the most important ones.

Stefanini et al., 2020, Neuron *107*, 1–14
August 19, 2020
https://doi.org/10.1016/j.neuron.2020.05.022

CellPress

Neuron

CellPress

Article

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] Nicholas I. Woods,[7,8] Garret D. Stuber,[9] Mazen A. Kheirbek,[7,10,11,12,14,*] René Hen,[4,5,13,14,*] and Stefano Fusi[1,2,6,13,14,15,*]

[1]Center for Theoretical Neuroscience, College of Physicians and Surgeons, Columbia University, New York, NY, USA
[2]Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA
[3]GNT-LNC, Département 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, USA
[9]Center 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, New York, NY, USA
[14]Senior author
[15]Lead Contact
*Correspondence: mazen.kheirbek@ucsf.edu (M.A.K.), rh95@cumc.columbia.edu (R.H.), sf2237@columbia.edu (S.F.)
https://doi.org/10.1016/j.neuron.2020.05.022

SUMMARY

Neurons are often considered 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 an 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 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.

INTRODUCTION

The hippocampus has been studied extensively in experiments regarding navigation and spatial memory. The responses of some of its cells are easily interpretable because 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 has been shown recently 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 because of 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

Please cite this article in press as: Stefanini et al., A Distributed Neural Code in the Dentate Gyrus and in CA1, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.05.022

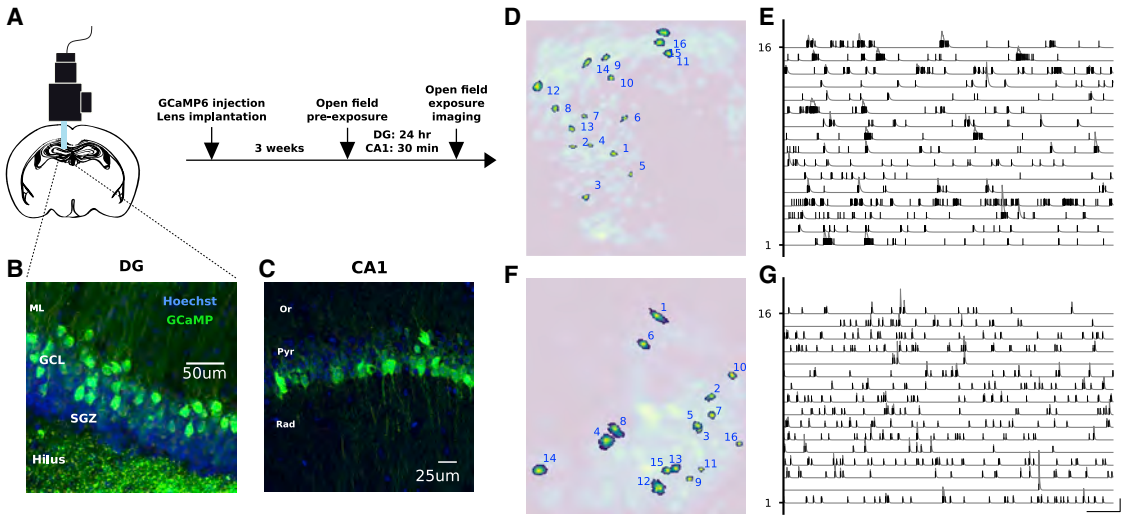

Figure 1. Calcium Image Recordings

(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 described previously (Resendez et al., 2016). The imaging plane was later assessed through histology (Figure S22).

(B) DG recording site. GCL, GC layer; SGZ, subgranular zone.

(C) CA1 recording site. Pyr: pyramidal layer; Or, stratum oriens; Rad, stratum radiatum.

(D–G) Automated signal extraction using CNMF-E (Zhou et al., 2018). The algorithm identifies the spatial (D, F) and temporal (E, G) 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 because of 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 (E) (DG) and (G) (CA1), where vertical ticks correspond to the times of the inferred calcium events and gray lines to the temporal profiles (Figure S1). In line with electrophysiology studies, DG GCs are sparsely active but often in bursts (Pernía-Andrade and Jonas, 2014).

Scale bars, 1 min and 1 SD.

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 studied extensively in relation to spatial navigation using 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 the DG with an accuracy comparable with 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 the 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 effect on decoding performance in CA1 but not in the 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 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 (Figure 1A–1C). 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; Figures 1D–1G). This algorithm separates local background signals resulting from changes in fluorescence in the neuropil from signals resulting from calcium concentration changes in individual cells. This was necessary to identify signal sources in our GC imaging data without introducing spurious distortions or correlations among cells because of artifacts. We identified a total of 1,109 DG cells across 3 animals, of which 352 (32%) were significantly tuned to position,

Please cite this article in press as: Stefanini et al., A Distributed Neural Code in the Dentate Gyrus and in CA1, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.05.022

and a total of 863 CA1 cells across 3 animals, of which 38 (4%) were significantly tuned to position (STAR Methods; Figure S2).

The low fraction of place cells in CA1 seems to be in contrast with reports from previous studies in CA1 (Meshulam et al., 2017; Pfeiffer and Foster, 2013; Talbot et al., 2018; Ziv et al., 2013). 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 these low firing rate cells from the analysis, the fraction of place cells becomes significantly higher (Figure S3; Table S1). Moreover, although calcium signal extraction may miss isolated spikes, this is likely less of an issue for DG GCs than pyramidal cells because the former are often active in burst, as electrophysiology studies have shown (Pernía-Andrade and Jonas, 2014) and as our data also show (Figure 1E). See Figure S3 and Table S1 for a brief review of the literature comparing firing rates in the 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 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 × 8 grid) (50-cm square arena for DG mice and 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 1-min-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 with the animal size, revealing for the first time that instantaneous position can be decoded from DG GC population activity (Figure 2). Our analysis of the CA1 data shows a comparable decoding accuracy in the DG and CA1 after correcting for the number of cells (Figure 2C; Figure 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 (Figure S6). The decoding error was found to weakly correlate to the speed of movement (Figure 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 the DG. Speed was weakly correlated with the overall level of activity in the DG, and we could decode it in two animals of three using linear regression (Figures 2B and 2C). 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 (STAR Methods). To our knowledge, this is also the first time that decoding of direction and speed of motion from populations of the DG and CA1 cells has been reported, although direction tuning has been observed previously 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 (Figure 2C).

To better characterize the neural code, we tried to determine which 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 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 where 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 shown in Figure 3, where 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 (Figure 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 because of 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 (Figures 3B). 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. Although 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 a situation, the linear decoder would assign equal weights to the two neurons, as shown in Figure 3B.

In the real data, there might be a spectrum of different situations that are less extreme than the one illustrated in Figure 3, in which a decoder can take advantage of weakly tuned cells. Cells like the untuned one shown in Figure 3 or weakly tuned cells can "cooperate" with more tuned cells to more precisely encode a variable like position. This is a situation similar to the one shown in Figure 3, where 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, 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 read out 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

Please cite this article in press as: Stefanini et al., A Distributed Neural Code in the Dentate Gyrus and in CA1, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.05.022

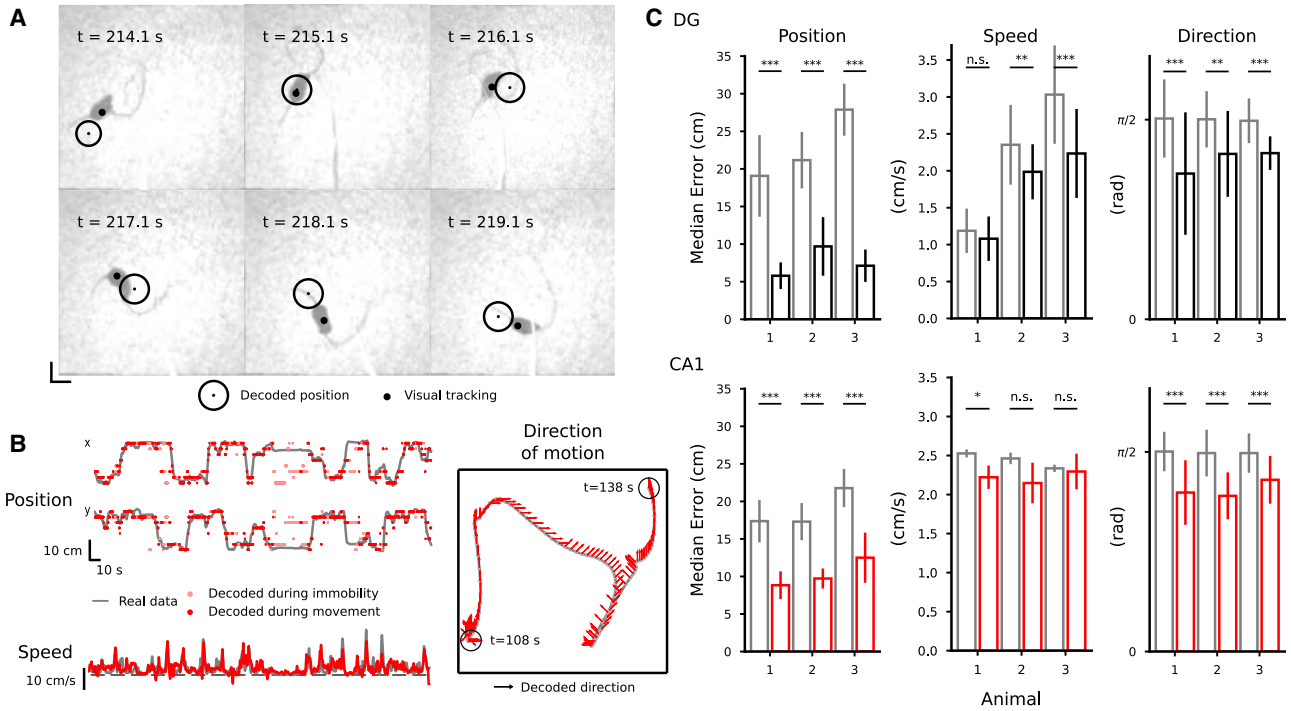

Figure 2. Decoding Position, Speed, and Direction of Motion

(A and B) Decoding results for a representative DG mouse. See also Videos S1 and S2.

(A) Selected frames of a video showing the arena and a DG animal from above. The black filled dot represents the mouse's actual position, and the black circle with the dot in the center is the decoded position, obtained with a probabilistic decoder that reads the activity of 317 DG cells (STAR Methods). Neural activity was pre-processed to identify putative calcium events, as explained in STAR 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 bottom left panels, respectively, whereas the red dots correspond to their decoded values. The time bins marked in light red for position and direction of movement correspond to moments of immobility that were excluded from the training data. The gray 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 of the distances computed 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 black and red vertical bars correspond to the mean over the 10-fold cross-validation (error bars correspond to SD). Gray, chance error obtained by decoding from shuffled data in a way that preserves the autocorrelations in the data (*p<0.05, **p<0.01, ***p<0.001, STAR Methods; Figure S5).

Number of cells: 483, 309, and 317 in DG mice; 371, 286, and 206 in CA1 mice.

See also Figures S4–S7 and S12.

first train the position decoder on each pair of locations and then to combine the resulting weights to obtain a single importance index (ω) for each cell (STAR 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 used recently 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 (Figures 4A and S17) to assess the validity of our approach. The 50 neurons with the largest importance index indeed performed significantly better than the worst 50 neurons, although 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 (Figure S8) and that it was not due to poor cell segmentation (Figure S24).

The observation that most neurons could contribute to 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 where 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 in the 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 where only a few cells are important (Figure S13).

Not too surprisingly, one important feature of an individual neuron is its average activity, which is strongly correlated with

Please cite this article in press as: Stefanini et al., A Distributed Neural Code in the Dentate Gyrus and in CA1, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.05.022

CellPress

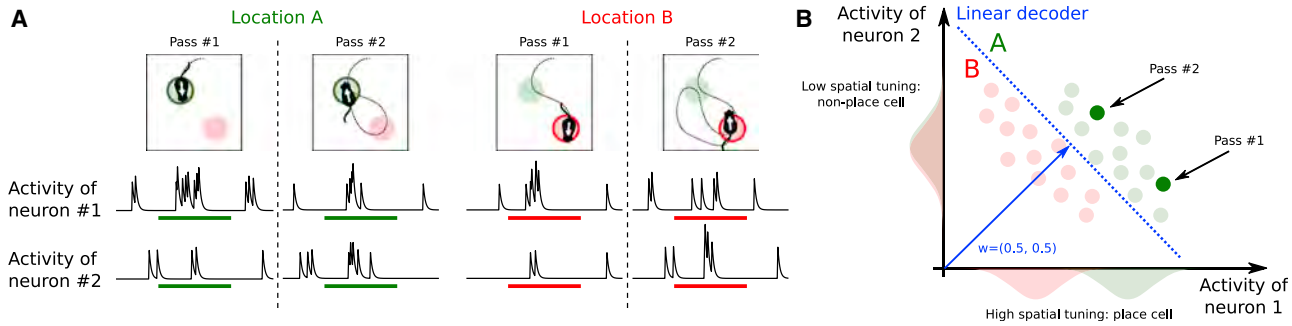

Figure 3. The Contribution of Untuned Cells for Encoding Position

We show an extreme situation where 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 because of 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 pass through each location. Different responses for the two neurons are elicited by different experiences; for example, because of 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 scatterplot, the x axis represents the average activity of the first neuron during one pass, and the y axis represents 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 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), whereas the second has only a weak tuning. The distributions of their activities in each location, reported along the axis, overlap only partially (neuron 1, place cells) or almost completely (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.

the importance index and, hence, to the overall ability to encode position (Figures 5A and 5B; Figures S9 and S19). However, inspection of the firing fields of Figure 4B indicated that there were no other obvious properties that predicted whether a neuron is important in the DG and CA1 neuron populations. We then identified which neurons were spatially tuned and called them place cells when the spatial information contained in their activity was statistically significant (see STAR 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 relative to a null distribution (Allegra et al., 2019; Danielson et al., 2017; Meshulam et al., 2017; Panzeri et al., 2007; Skaggs et al., 1992).

From Figure 4B, it is clear that, in our data, there are non-place cells that have a large importance index. The animal's position could be decoded from these cells alone in the DG and CA1 (Figures 4C and 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 we adopted to characterize place cells.

Although the SSI is a property of a 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 (Figures 5A–5C; Figure 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 (Figure S9A; Kelemen and Fenton, 2010). However, the baseline was also negatively correlated with the importance index (Figure 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 unreliable (e.g., they fire at a particular location only during one pass); hence, their importance index is low.

We performed a similar analysis of importance for the direction of movement. In Figure 6 (see also Figure S18), we show that we could rank the cells according to their contribution to decoding (Figure 6A; Figure S18) and that the important cells were highly heterogeneous in their direction tuning (Figure 6B). Considering all recorded cells, we also found that a cell's activity correlated with the importance index for direction of movement in the DG and in CA1 (Figure 6C). We defined the significance of directional information (SDI) in a similar way as 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 in the DG and CA1 (Figure 6D; Figure S20).

All of 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

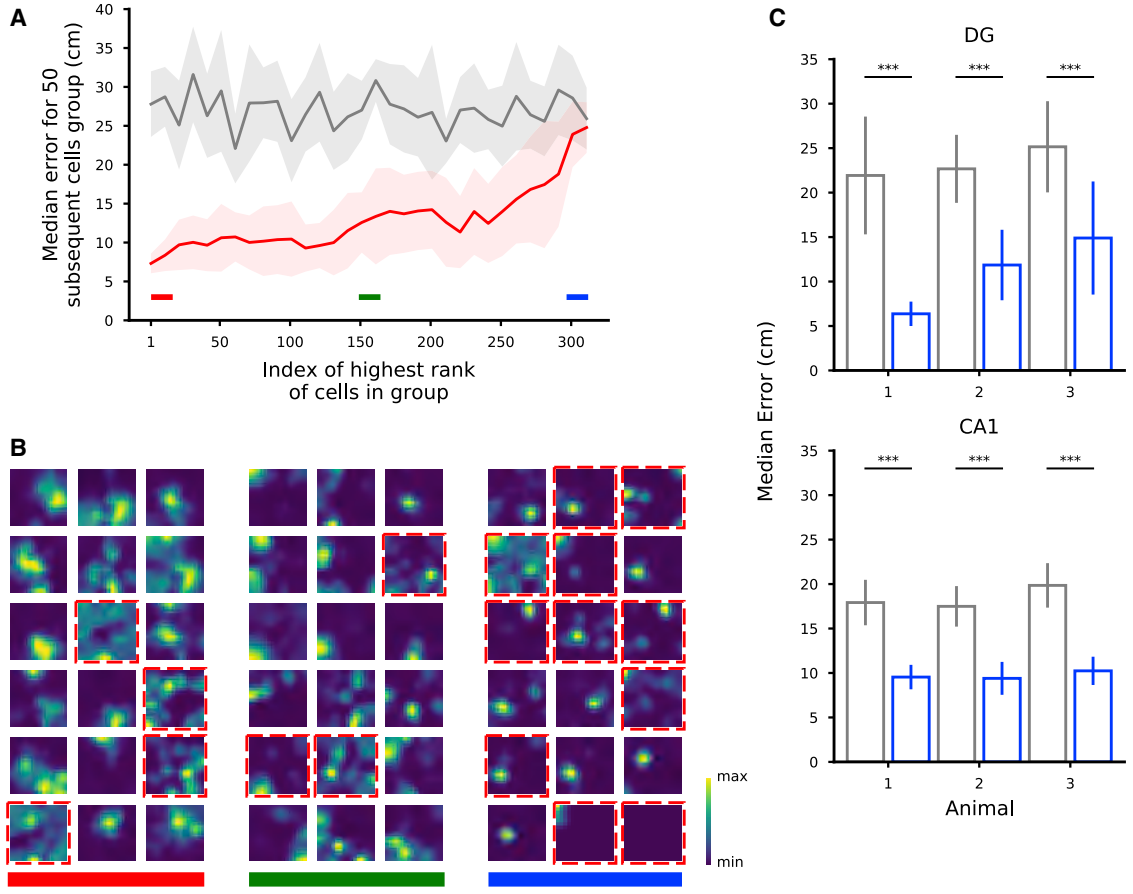

Figure 4. Ranking Neurons According to Their Contribution to the Decoding Accuracy for Position

(A) Validation of the importance index. 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 SD for the 10-fold cross-validation). Gray: chance level and SD. As expected, the median error for the population of the 50 top-ranked (best) cells is much smaller than the median error for the last (worst) 50 ones.

(B) Spatial tuning maps for groups of 18 cells ordered by importance index (the same cells as in A). We ranked the cells using the importance index for position (STAR 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 do not pass the criteria for place cells using a commonly used statistical test for tuning (STAR 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 (Figure 5).

(C) The position of DG and CA1 animals can be decoded from the activity of the non-place cells with a performance significantly higher than chance (vertical bars correspond to mean and SD, ***$p<0.001$, STAR Methods).

Number of cells: 451, 208, and 98 in DG mice; 350, 277, and 198 in CA1 mice.

See also Figures S2, S3, S8, S10, and S24 and Table S1.

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 Figure 3, where the untuned cell is important when combined with the cell represented on the horizontal axis, but it would be useless when combined with another untuned cell or with an uncorrelated tuned cell.

Because we could decode at least two variables from the neural activities, we wondered whether we could identify some form of specialization where segregated groups of neurons encoded different variables. In Figure 7, we report the importance index for the direction of movement versus the importance index for position (Figure 7A; Figure S21). The situation where different variables would be encoded by segregated populations of neurons would predict a negative correlation between these two importance indices: cells with a large importance index for position should have a small importance for the direction of movement and vice versa. Instead, for both 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 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 (Figure 7A). In addition, this could not be explained

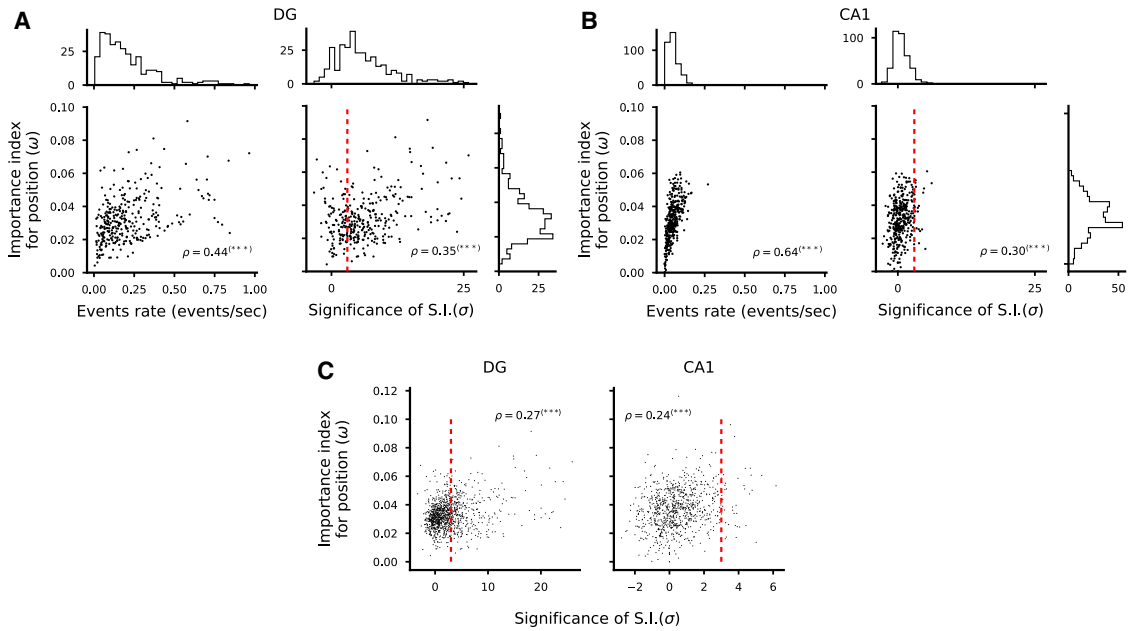

Figure 5. Correlation between Importance Index and Spatial Information

(A and B) Left: scatterplot of the 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: scatterplot of the importance index and statistical SSI with respect to independent random temporal shuffling of each cell's identified calcium events. DG cells are shown in (A) and CA1 cells in (B). Each dot corresponds to one cell in one representative animal. Pearson's correlation factor ρ between the plotted quantities are reported (Pearson's correlation significance, ***p<0.001). Significant correlations are found between the analyzed 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 (STAR Methods).

(C) The same plots as in (A) and (B) but for all cells identified in all fields of view (FOVs) in DG (left) and CA1 (right) (Pearson's correlation significance, ***p<0.001). See also Figures S9 and S13.

by a correlation between direction of movement and position (Figure S11).

We then focused on cells that had a high importance for one variable but 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 (Figure 7B). We conclude that, in CA1 and the DG, neurons have mixed selectivity to the variables we decoded, in line with recent studies 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; 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 whether correlations between the activities of different neurons contribute to the decoding performance in a similar way as what we described in Figure 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 noise (i.e., not related to 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 because 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 (Discussion). For example, the different points that encode the same position in Figure 3 might correspond to visitations where 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 (Figure 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 (STAR Methods; Figure 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 (Figure 8B). We then corrected for the different time spent in each pass at the same location and repeated the procedure for all cells independently.

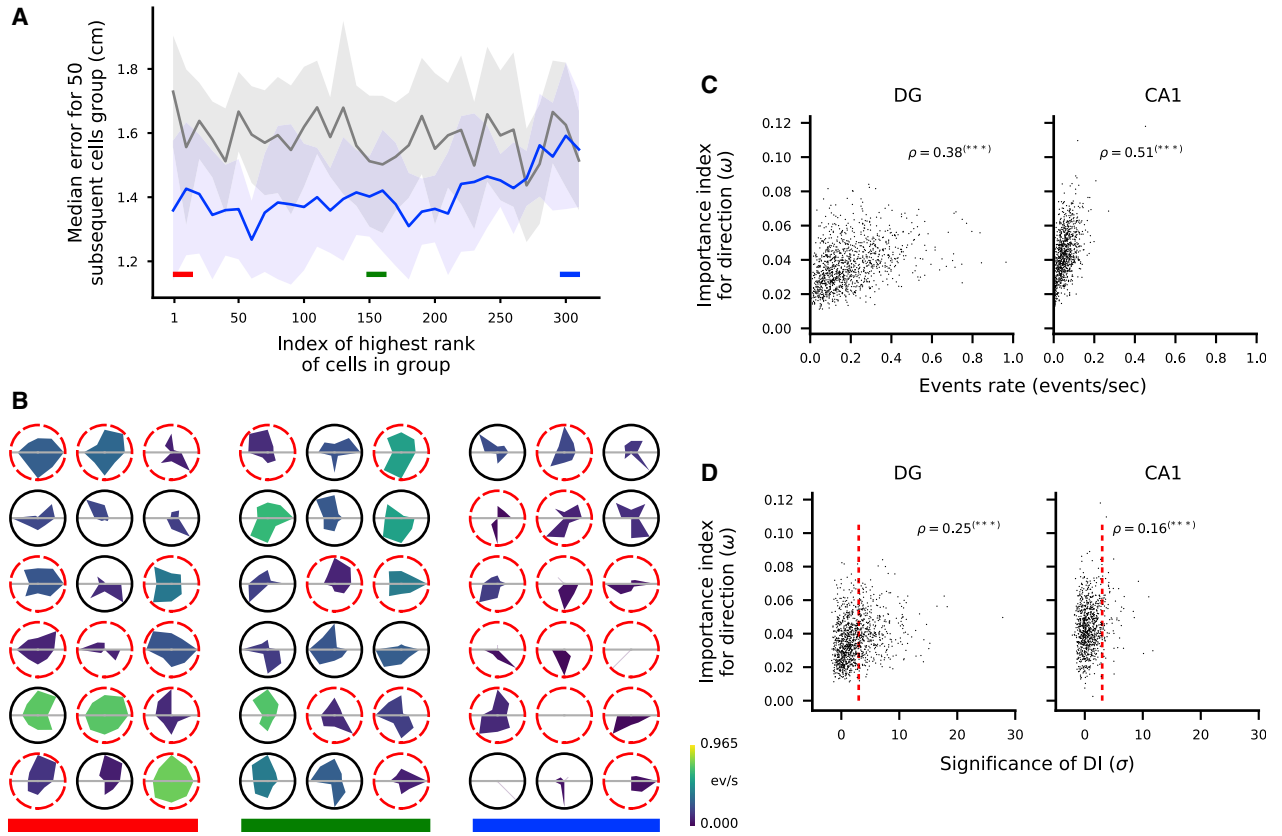

Figure 6. Ranking Neurons According to Their Contribution to the Decoding Accuracy for Head Direction

(A) Validation of the importance index as in Figure 4A, but we ranked the cells according to the importance index for decoding direction of motion (STAR Methods).

(B) Tuning maps as in Figure 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 color represents the overall activity of the cell throughout the trial. Dashed red borders indicate cells that do not pass the criteria for significant direction tuning using a commonly used statistical test (STAR 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) Scatterplots of cell activity and importance for position decoding for all identified cells combined from all FOVs in the DG (left) and CA1 (right). The Pearson correlation factor ρ between the plotted quantities is reported (Pearson's correlation, ***p<0.001).

(D) Same as in (C) but for importance index for direction and significance of direction information.

By using this procedure, we effectively destroyed the noise correlations between neurons because, after this manipulation, each cell'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 (Figure 8C) or the signal correlations among neurons induced by their tuning profiles. By comparing the performance of the decoder on the modified data with 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 (Figure 8D). The effect was very consistent in CA1 neurons, where performance was reduced by about 20%, whereas almost no effect was observed in the DG (Figure S14). Pairwise correlations were found to be lower in the DG than in CA1 (Figure 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 the DG, and those we observed certainly changed after disrupting the noise correlations (for an analysis of how destroying correlations affects pairwise correlations and the importance index, see Figures S16 and S23).

DISCUSSION

Neurons in the DG and 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, speed, and direction of motion of the animal. Neurons respond to mixtures of the decoded variables, as observed in other highly cognitive brain areas (Rigotti et al., 2013; Fusi et al., 2016). The information

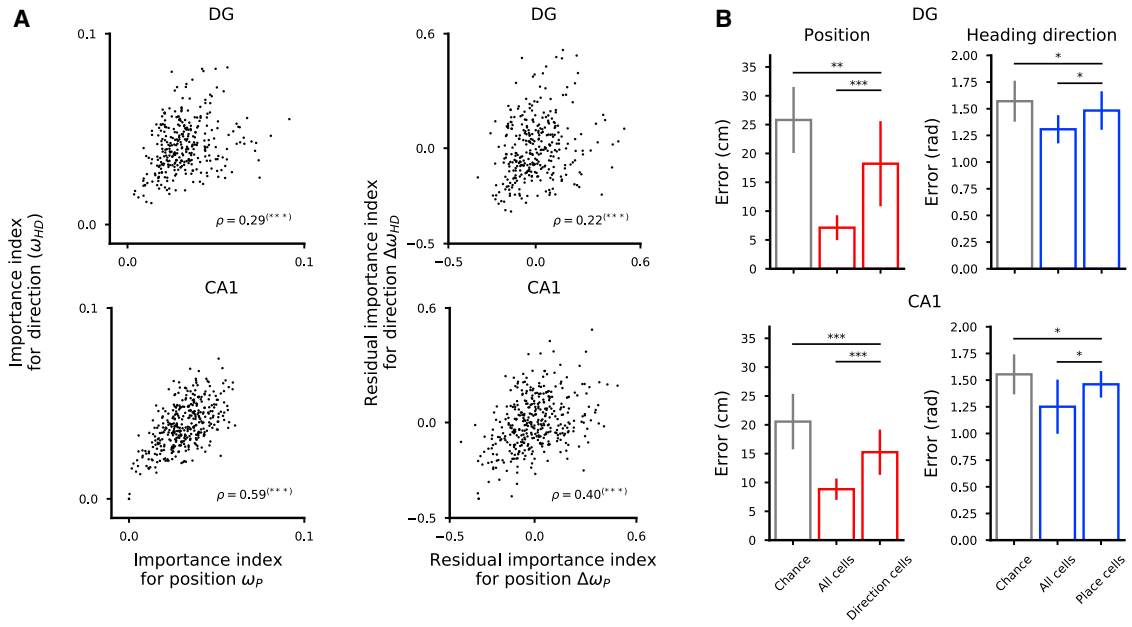

Figure 7. The Representations for Space and Direction of Motion Are Distributed in DG Cells and CA1 Cells

(A) Left: scatterplots 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 ρ are reported (Pearson's correlation, ***p<0.001). Right: same as left, but the component due to the correlation between importance index and cell activity was removed from the data. Residuals from linear regression are considered for both quantities. The residuals also show a positive correlation.

(B) 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). Vertical bars correspond to mean and SD (*p<0.05, **p<0.01, ***p<0.001, STAR Methods).

about these variables is highly distributed across neurons to the point where 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 a 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 the 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 a few tens of cells encode position with high precision in both analyzed areas. Furthermore, the decoding was accurate even when the model training and model test periods were separated by up to 18 min, indicating that, at the population level, the representations were stable despite the elevated variability of individual cells (Figure S12).

Our findings are in line with studies suggesting that session-averaged, single-cell statistics fall short in describing the activities of hippocampal cells. For example, although 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 the CA1 and DG areas of the hippocampus. A recent work by 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 other neurons in the network, regardless of how well that neuron codes for position. They conclude that correlation patterns in the 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 et al. (2017). Our analysis also shows directly that non-place cells contribute to encoding of the position of the animal. This is partially due to the fact that some of the cells,

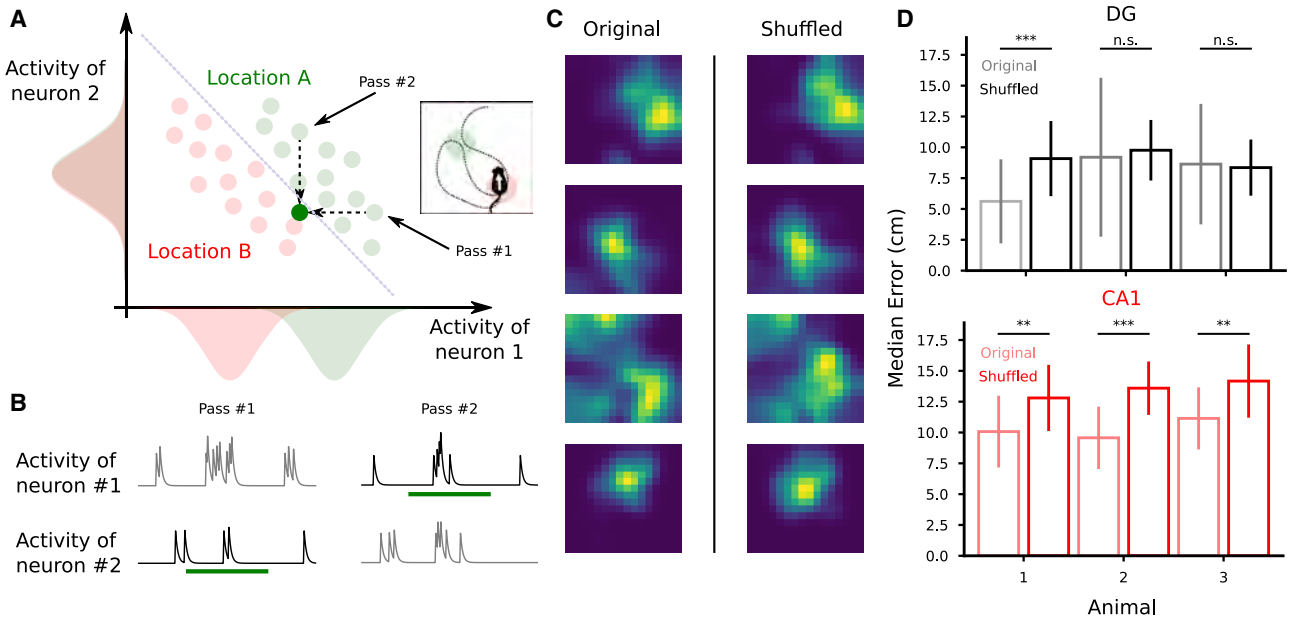

Figure 8. Destroying Correlations Affects Decoding Performance in CA1 but Not in the DG

Figure360▷ For a Figure360 author presentation of this figure, see https://doi.org/10.1016/j.neuron.2020.05.022.

(A) Procedure to test the presence of correlations between cells. We recorded neural activity during multiple passes through location A (green). Then we generateed 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 repeated this procedure for all the passes through each location and for each cell independently, destroying the correlations between cells, if any. In the extreme case depicted in the cartoon, this procedure will introduce errors in decoding position because 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 (left) and after (right) applying the shuffling procedure to destroy correlations. The spatial tuning of the cells remain unaltered after the procedure.

(D) Decoding performance before (light colors) and after destroying correlations through shuffling (full colors). Top: DG animals. Bottom: CA1 animals. Vertical bars correspond to mean and SD (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, STAR Methods).

See also Figures S14, S15, and S24.

when 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 et al. (2017) to select place cells is similar to ours. However, non-place cells contribute to position encoding also because of 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 because tuning properties that are based on single-cell statistics might not be sufficient to understand how task-relevant variables are encoded (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 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 where 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; hence, they greatly limit the number of combinations of input variables to which a linear readout or a downstream neuron can respond; 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 all relevant variables non-linearly 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 also showed that our data support 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 the 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 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 in Figure 3, for instance, we show how 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 that include position, speed, movement direction, and possibly other variables that were not under our control (Allegra et al., 2019). We modeled such a situation in Figure 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 effect on 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 pairwise correlations only accounted for about 10% of the information contained in neural activities (Averbeck and Lee, 2006; Latham and Nirenberg, 2005; Schneidman et al., 2006), whereas 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 a relatively modest but statistically significant decrease in decoding accuracy in CA1 but not in the DG.

These observations indicate that the correlations we are destroying should be considered 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 Figure 3 (e.g., passes 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 Figure S15, we show, in simulations, that this is indeed the case. We considered a model where 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, where different conditions are represented by different random vectors in the activity space, and one with a kind of structured geometry that 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 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 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 the DG are compatible with a scenario in which the representations in CA1 are unstructured, similar to the simulated representations obtained with the random model. Our results also show that the representations in the DG are compatible with at least two scenarios: (1) they could be structured, as we described them in Figures S15D–S15F, 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 Figure 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 we did not consider may be more appropriate to describe DG and CA1 representations.

One alternative explanation for the difference between CA1 and the 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 the 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 the DG in the effect of destroying correlations because we did not observe any performance reduction in the DG when the level of activity was matched to the one observed in CA1 (Figure S14).

Our simulations where 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 of these models, the neural representations in the hippocampus are constructed by learning the statistics of the sensory experiences to generate 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 the DG Is Larger Than in CA1

One of the observations that requires some discussion is that the average activity in the DG is larger than that 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 (Figure S3; 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 were employed. Our conclusion is that our results fall within the range of values reported in the existing literature. The review reported in the Table S1 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 CA1 area of the hippocampus is highly distributed and that it is important to analyze it using a population approach (Fenton et al., 2008; Meshulam et al., 2017; van Dijk and Fenton, 2018). 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 CA1 area of the hippocampus should be revisited in light of our observations. The methods 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

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
  - Viral Constructs
- METHOD DETAILS
  - Calcium imaging
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavior data pre-processing
  - Signal extraction and spike deconvolution
  - Place fields and heading direction tuning
  - Spatial information statistics
  - Decoding position
  - Chance level decoding performance
  - Decoding the direction of motion
  - Decoding speed
  - Bayesian decoder
  - Importance index
  - Procedure to destroy correlations
  - Software

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.neuron.2020.05.022.

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 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 NeuroNex Program award DBI-1707398, Swiss National Science Foundation (SNSF) Early Postdoc Mobility fellowship P2EZP2_155561, and by the Kavli Foundation. S.F. was supported by NSF NeuroNex Program award DBI-1707398, the Gatsby Charitable Foundation, the Simons Foundation, the Swartz Foundation, and 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.

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., and S.F.; Investigation, J.H.J., M.A.K., J.C.J., and N.I.W.; Resources, J.H.J., G.D.S., M.A.K., and R.H.; Data Curation, J.C.J., F.S., and M.A.K.; Writing – Original Draft, F.S. and S.F.; Writing – Review & Editing, F.S., S.F., M.A.K., and R.H.; Visualization, F.S.; Supervision, M.A.K., S.F., and R.H.; Funding Acquisition, F.S., M.A.K., S.F., and R.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 28, 2019
Revised: April 1, 2020
Accepted: May 15, 2020
Published: June 9, 2020

SUPPORTING CITATIONS

The following references appear in the Supplemental Information: Dombeck et al. (2007); Gini (1912); GoodSmith et al. (2017); Kinsky et al. (2018); Mukamel et al. (2009); Neunuebel and Knierim (2012); Nitz and McNaughton (2004); Pachitariu et al. (2017); Pnevmatikakis et al. (2016).

REFERENCES

Abbott, L.F., and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. Neural Comput. *11*, 91–101.

Acharya, L., Aghajan, Z.M., Vuong, C., Moore, J.J., and Mehta, M.R. (2016). Causal influence of visual cues on hippocampal directional selectivity. Cell *164*, 197–207.

Agarwal, G., Stevenson, I.H., Berényi, A., Mizuseki, K., Buzsáki, G., and Sommer, F.T. (2014). Spatially distributed local fields in the hippocampus encode rat position. Science *344*, 626–630.

Allegra, M., Posani, L., and Schmidt-Hieber, C. (2019). The hippocampus as a perceptual map: neuronal and behavioral discrimination during memory encoding. bioRxiv. https://doi.org/10.1101/868794.

Averbeck, B.B., and Lee, D. (2006). Effects of noise correlations on information encoding and decoding. J. Neurophysiol. *95*, 3633–3644.

Benna, M.K., and Fusi, S. (2019). Are place cells just memory cells? memory compression leads to spatial tuning and history dependence. bioRxiv. https://doi.org/10.1101/624239.

Bernardi, S., Benna, M.K., Rigotti, M., Munuera, J., Fusi, S., and Salzman, D. (2018). The geometry of abstraction in hippocampus and prefrontal cortex. bioRxiv. https://doi.org/10.1101/408633.

Bishop, C.M. (2006). Pattern Recognition and Machine Learning (Springer New York).

Brody, C.D. (1999). Correlations without synchrony. Neural Comput. *11*, 1537–1551.

Cortes, C., and Vapnik, V. (1995). Support-vector networks. Mach. Learn. *20*, 273–297.

Danielson, N.B., Kaifosh, P., Zaremba, J.D., Lovett-Barron, M., Tsai, J., Denny, C.A., Balough, E.M., Goldberg, A.R., Drew, L.J., Hen, R., et al. (2016). Distinct Contribution of Adult-Born Hippocampal Granule Cells to Context Encoding. Neuron *90*, 101–112.

Danielson, N.B., Turi, G.F., Ladow, M., Chavlis, S., Petrantonakis, P.C., Poirazi, P., and Losonczy, A. (2017). In vivo imaging of dentate gyrus mossy cells in behaving mice. Neuron *93*, 552–559.e4.

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

Dipoppa, M., Ranson, A., Krumin, M., Pachitariu, M., Carandini, M., and Harris, K.D. (2018). Vision and locomotion shape the interactions between neuron types in mouse visual cortex. Neuron *98*, 602–615.e8.

Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging large-scale neural activity with cellular resolution in awake, mobile mice. Neuron *56*, 43–57.

Dombeck, D.A., Harvey, C.D., Tian, L., Looger, L.L., and Tank, D.W. (2010). Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. Nat. Neurosci. *13*, 1433–1440.

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? Neurosci. Lett. *680*, 88–93.

Eyherabide, H.G., and Samengo, I. (2013). When and why noise correlations are important in neural decoding. J. Neurosci. *33*, 17921–17936.

Fenton, A.A., and Muller, R.U. (1998). Place cell discharge is extremely variable during individual passes of the rat through the firing field. Proc. Natl. Acad. Sci. USA *95*, 3182–3187.

Fenton, A.A., Kao, H.Y., Neymotin, S.A., Olypher, A., Vayntrub, Y., Lytton, W.W., and 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. J. Neurosci. *28*, 11250–11262.

Fusi, S., Miller, E.K., and Rigotti, M. (2016). Why neurons mix: High dimensionality for higher cognition. Curr. Opin. Neurobiol. *37*, 66–74.

Gershman, S.J., Moore, C.D., Todd, M.T., Norman, K.A., and Sederberg, P.B. (2012). The successor representation and temporal context. Neural Comput. *24*, 1553–1568.

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

Gluck, M.A., and Myers, C.E. (1993). Hippocampal mediation of stimulus representation: a computational theory. Hippocampus *3*, 491–516.

GoodSmith, D., Chen, X., Wang, C., Kim, S.H., Song, H., Burgalossi, A., Christian, K.M., and Knierim, J.J. (2017). Spatial representations of granule cells and mossy cells of the dentate gyrus. Neuron *93*, 677–690.e5.

Hardcastle, K., Maheswaranathan, N., Ganguli, S., and Giocomo, L.M. (2017). A Multiplexed, Heterogeneous, and Adaptive Code for Navigation in Medial Entorhinal Cortex. Neuron *94*, 375–387.e7.

Harvey, C.D., Collman, F., Dombeck, D.A., and Tank, D.W. (2009). Intracellular dynamics of hippocampal place cells during virtual navigation. Nature *461*, 941–946.

Haufe, S., Meinecke, F., Görgen, K., Döhne, S., Haynes, J.D., Blankertz, B., and Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. NeuroImage *87*, 96–110.

Keinath, A.T., Wang, M.E., Wann, E.G., Yuan, R.K., Dudman, J.T., and Muzzio, I.A. (2014). Precise spatial coding is preserved along the longitudinal hippocampal axis. Hippocampus *24*, 1533–1548.

Kelemen, E., and Fenton, A.A. (2010). Dynamic grouping of hippocampal neural activity during cognitive control of two spatial frames. PLoS Biol. *8*, e1000403.

Kinsky, N.R., Sullivan, D.W., Mau, W., Hasselmo, M.E., and Eichenbaum, H.B. (2018). Hippocampal place fields maintain a coherent and flexible map across long timescales. Curr. Biol. *28*, 3578–3588.e6.

Kriegeskorte, N., and Douglas, P.K. (2019). Interpreting encoding and decoding models. Curr. Opin. Neurobiol. *55*, 167–179.

Latham, P.E., and Nirenberg, S. (2005). Synergy, redundancy, and independence in population codes, revisited. J. Neurosci. *25*, 5195–5206.

Leutgeb, J.K., Leutgeb, S., Moser, M.B., and Moser, E.I. (2007). Pattern separation in the dentate gyrus and ca3 of the hippocampus. Science *315*, 961–966.

Lindsay, G.W., Rigotti, M., Warden, M.R., Miller, E.K., and Fusi, S. (2017). Hebbian learning in a random network captures selectivity properties of the prefrontal cortex. J. Neurosci. *37*, 11021–11036.

Meshulam, L., Gauthier, J.L., Brody, C.D., Tank, D.W., and Bialek, W. (2017). Collective behavior of place and non-place neurons in the hippocampal network. Neuron *96*, 1178–1191.e4.

Mladenić, D., Brank, J., Grobelnik, M., and Natasa Milic-Frayling, I. (2004). Feature Selection using Linear Classifier Weights: Interaction with Classification Models. In Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '04) (Association for Computing Machinery), pp. 234–241.

Moser, E.I., Kropff, E., and Moser, M.B. (2008). Place cells, grid cells, and the brain's spatial representation system. Annu. Rev. Neurosci. *31*, 69–89.

Mukamel, E.A., Nimmerjahn, A., and Schnitzer, M.J. (2009). Automated analysis of cellular signals from large-scale calcium imaging data. Neuron *63*, 747–760.

Neunuebel, J.P., and Knierim, J.J. (2012). Spatial firing correlates of physiologically distinct cell types of the rat dentate gyrus. J. Neurosci. *32*, 3848–3858.

Nitz, D., and McNaughton, B. (2004). Differential modulation of CA1 and dentate gyrus interneurons during exploration of novel environments. J. Neurophysiol. *91*, 863–872.

Olypher, A.V., Lánský, P., Muller, R.U., and Fenton, A.A. (2003). Quantifying location-specific information in the discharge of rat hippocampal place cells. J. Neurosci. Methods *127*, 123–135.

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L.F., Dalgleish, H., Carandini, M., and Harris, K.D. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. bioRxiv. https://doi.org/10.1101/061507.

Panzeri, S., Senatore, R., Montemurro, M.A., and Petersen, R.S. (2007). Correcting for the sampling bias problem in spike train information measures. J. Neurophysiol. *98*, 1064–1072.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2012). Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

Pernía-Andrade, A.J., and Jonas, P. (2014). Theta-gamma-modulated synaptic currents in hippocampal granule cells in vivo define a mechanism for network oscillations. Neuron *81*, 140–152.

Pfeiffer, B.E., and Foster, D.J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. Nature *497*, 74–79.

Pillow, J.W., Shlens, J., Paninski, L., Sher, A., Litke, A.M., Chichilnisky, E.J., and Simoncelli, E.P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature *454*, 995–999.

Pnevmatikakis, E.A., Soudry, D., Gao, Y., Machado, T.A., 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.

Recanatesi, S., Farrell, M., Lajoie, G., Deneve, S., Rigotti, M., and Shea-Brown, E. (2018). Signatures and mechanisms of low-dimensional neural predictive manifolds. bioRxiv. https://doi.org/10.1101/471987.

Resendez, S.L., Jennings, J.H., Ung, R.L., Namboodiri, V.M.K., Zhou, Z.C., Otis, J.M., Nomura, H., McHenry, J.A., Kosyk, O., and Stuber, G.D. (2016). Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. Nat. Protoc. *11*, 566–597.

Rigotti, M., Barak, O., Warden, M.R., Wang, X.J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature *497*, 585–590.

Saxena, S., and Cunningham, J.P. (2019). Towards the neural population doctrine. Curr. Opin. Neurobiol. *55*, 103–111.

Schneidman, E., Bialek, W., and Berry, M.J., 2nd (2003). Synergy, redundancy, and independence in population codes. J. Neurosci. *23*, 11539–11553.

Schneidman, E., Berry, M.J., 2nd, Segev, R., and Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. Nature *440*, 1007–1012.

Skaggs, W.E., McNaughton, B.L., Gothard, K.M., and Markus, E.J. (1992). An information-theoretic approach to deciphering the hippocampal code. Advances in neural information processing systems 5 (Morgan Kaufmann), pp. 1030–1037.

Stachenfeld, K.L., Botvinick, M., and Gershman, S.J. (2014). Design principles of the hippocampal cognitive map. In Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'14) (MIT Press), pp. 2528–2536.

Stachenfeld, K.L., Botvinick, M.M., and Gershman, S.J. (2017). The hippocampus as a predictive map. Nat. Neurosci. *20*, 1643–1653.

Talbot, Z.N., Sparks, F.T., Dvorak, D., Curran, B.M., Alarcon, J.M., and Fenton, A.A. (2018). Normal CA1 place fields but discoordinated network discharge in a Fmr1-null mouse model of fragile x syndrome. Neuron *97*, 684–697.e4.

Tang, S., Zhang, Y., Li, Z., Li, M., Liu, F., Jiang, H., and Lee, T.S. (2018). Large-scale two-photon imaging revealed super-sparse population codes in the V1 superficial layer of awake monkeys. eLife *7*, e33370.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. J. R. Stat. Soc. Series B Stat. Methodol. *57*, 267–288.

van Dijk, M.T., and Fenton, A.A. (2018). On how the dentate gyrus contributes to memory discrimination. Neuron *98*, 832–845.e5.

Whittington, J.C., Muller, T.H., Mark, S., Chen, G., Barry, C., Burgess, N., and Behrens, T.E. (2019). The tolman-eichenbaum machine: Unifying space and relational memory through generalisation in the hippocampal formation. bioRxiv. https://doi.org/10.1101/770495.

Wilson, M.A., and McNaughton, B.L. (1993). Dynamics of the hippocampal ensemble code for space. Science *261*, 1055–1058.

Wood, E.R., Dudchenko, P.A., and Eichenbaum, H. (1999). The global record of memory in hippocampal neuronal activity. Nature *397*, 613–616.

Zenke, F., Poole, B., and Ganguli, S. (2017). Continual learning through synaptic intelligence. In Proceedings of the 34th International Conference on Machine Learning, D. Precup and Y.W. Teh, eds., pp. 3987–3995.

Zhang, K., Ginzburg, I., McNaughton, B.L., and Sejnowski, T.J. (1998). Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. J. Neurophysiol. *79*, 1017–1044.

Zhou, P., Resendez, S.L., Rodriguez-Romaguera, J., Jimenez, J.C., Neufeld, S.Q., Giovannucci, A., Friedrich, J., Pnevmatikakis, E.A., Stuber, G.D., Hen, R., et al. (2018). Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. eLife *7*, e28728.

Ziv, Y., Burns, L.D., Cocker, E.D., Hamel, E.O., Ghosh, K.K., Kitch, L.J., El Gamal, A., and Schnitzer, M.J. (2013). Long-term dynamics of CA1 hippocampal place codes. Nat. Neurosci. *16*, 264–266.

Please cite this article in press as: Stefanini et al., A Distributed Neural Code in the Dentate Gyrus and in CA1, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.05.022

CellPress

STAR★METHODS

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 6f	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; http://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.html; 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., 1992	N/A

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 $\sim 4X10^{12}$vg/ml, and AAV1-Syn-GCaMP6f.WPRE.SV40 was packaged and supplied by U Penn Vector Core at titers of $\sim 2X10^{12}$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, ∼ 90nl per site) and a ∼ 1.0mm diameter, ∼ 4mm 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 (Figure 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 behavior 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

Behavior data pre-processing

The behavior 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 center 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 analyzed 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 Figure 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 x 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 (Allegra et al., 2019; Danielson et al., 2017; Skaggs et al., 1992). We used a 16x16 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 Figures 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 8x8 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 center 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 Figure S3 and Table S1).

Please cite this article in press as: Stefanini et al., A Distributed Neural Code in the Dentate Gyrus and in CA1, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.05.022

Decoding the direction of motion

One behaviorally 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.

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:

$$P(x|r_t) = P(r_t|x)P(x) \,/\, P(r_t)$$

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

Neuron, Volume *107*

Supplemental Information

A Distributed Neural Code in the Dentate Gyrus and in CA1

Fabio Stefanini, Lyudmila Kushnir, Jessica C. Jimenez, Joshua H. Jennings, Nicholas I. Woods, Garret D. Stuber, Mazen A. Kheirbek, René Hen, and Stefano Fusi

Supplemental Material

Supplemental Material S1: Preprocessing; Related to Fig. 1

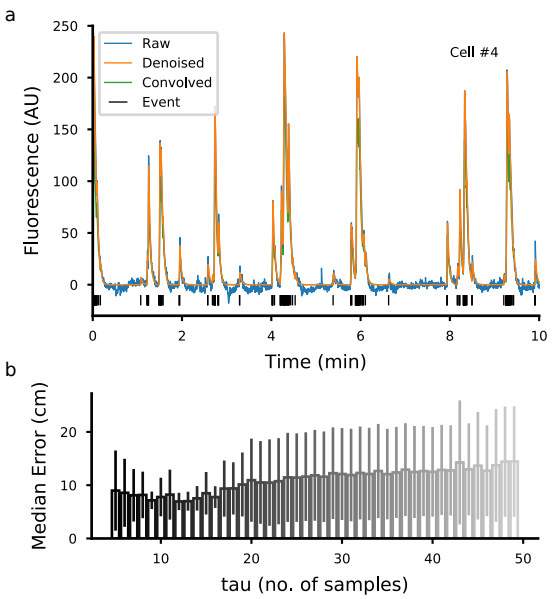

Figure S1: Preprocessing. Here we illustrate the type of signals extracted using CNMF-e that we used for our analysis. The algorithm's output is composed of a spatial footprint, the raw calcium signal, the denoised signal, the calcium events timings for each cell and their magnitude. **a**) Example of fluorescence signal extracted using CNMF-e (Zhou et al., 2018). The raw, denoised and event signals correspond to the output of the algorithm. The raw signal corresponds to the fluorescence levels attributed to the single sources when the contaminating signal due to neuropil in the background and to neighbouring sources have been eliminated. This signal includes noise due to the recording apparatus. The convolved signal is the one used in this work and is obtained using an auto-regressive model that uses the estimated calcium event timings. The parameter of the model as well as the calcium event times, the spatial footprints and the background estimation are all part of the minimization process in CNMF-e (Zhou et al., 2018). In our analysis we used the calcium events times because these gave us the closest estimates of each cell's activity. We then convolved the calcium events in time with a decaying temporal profile to cumulate information in time. Similar procedures have been used in the past to amplify the signal with temporally sparse data (Ziv et al., 2013). **b**) Median Error (mean $\pm$ s.d.) of the position decoding for different time scales of the exponential kernel used to convolve the calcium events. The chosen time scale corresponds to the one that minimizes the position decoding error (2.2 s in this example, 12 samples).

Supplemental Material S2: Place-cells; Related to Fig. 4

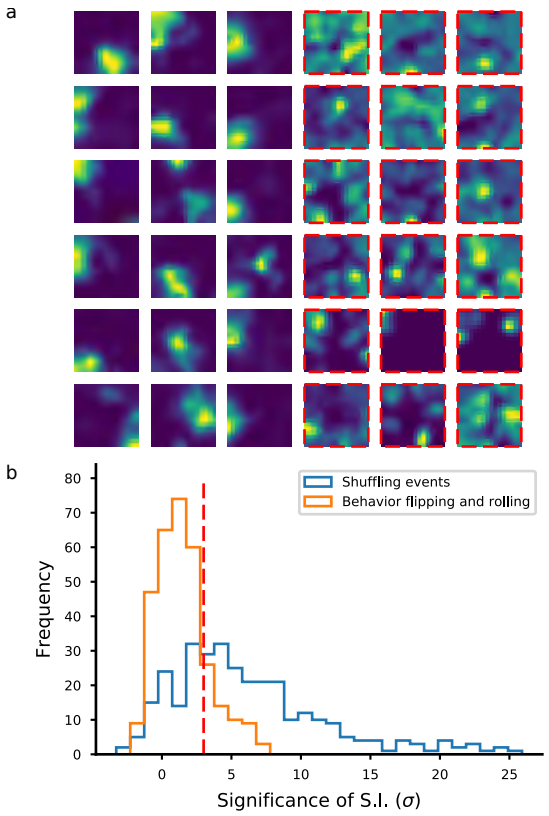

Figure S2: Place-cells. **a**) Gaussian-smoothed event density maps for a selection of cells, normalized by the mouse's occupancy time per 16x16 unit area and the cell's maximum response in the 50 cm x 50 cm arena (Ziv et al., 2013). The cells have been ordered by the statistical significance of their spatial information, from the most to the least significant. Here we show the group of 18 cells with the most significant SSI (place cells) and the 18 with the least significant SSI (non place-cells, red-borders) from the dentate gyrus of a representative mouse (DG3). **b**) Histogram of SSIs for all the DG cells recorded for mouse DG3 computed with the event shuffling method (blue) and the trajectory shuffling method (orange). In order to quantify the spatial tuning of the cells, we applied standard methods used in electrophysiology to measure the spatial information contained in the activity of each cell using the calcium events (Skaggs et al., 1993). We corrected for the sampling bias problem (Panzeri et al., 2007) by using a shuffling approach. We quantified the tuning of the cells by the statistical significance of the spatial information content (SSI). Similar to what we did for assessing the chance level decoding (see Fig. S5), we used two shuffling methods for assessing the SSI of each cell. The first method (blue histogram) consisted in shuffling the event timings, as it is commonly used in the definition of a place cell. We identified place cells as those with an SSI>3. For comparison, here we report a second method which consisted in time reversing the animal's trajectory ("flip") and then shifting it in time by a random amount ("roll"). This second method preserves the activity of all cells while destroying its association to the position of the animal and therefore it's more conservative (see also Fig. S3). The latter method highlights the importance of adopting the correct null hypothesis to assess the spatial information of the cells. In our recordings, many of the cells that are classically classified as place cells have a spatial information content that is not statistically different from the one obtained by a conservative shuffling of the trajectory which still disrupts the association between a cell's activity and the mouse position at each time point. In the text, we

used the former definition for consistency with the literature.

Supplemental Material S3: Fraction of place cells and activity levels in DG and CA1; Related to Fig. 4

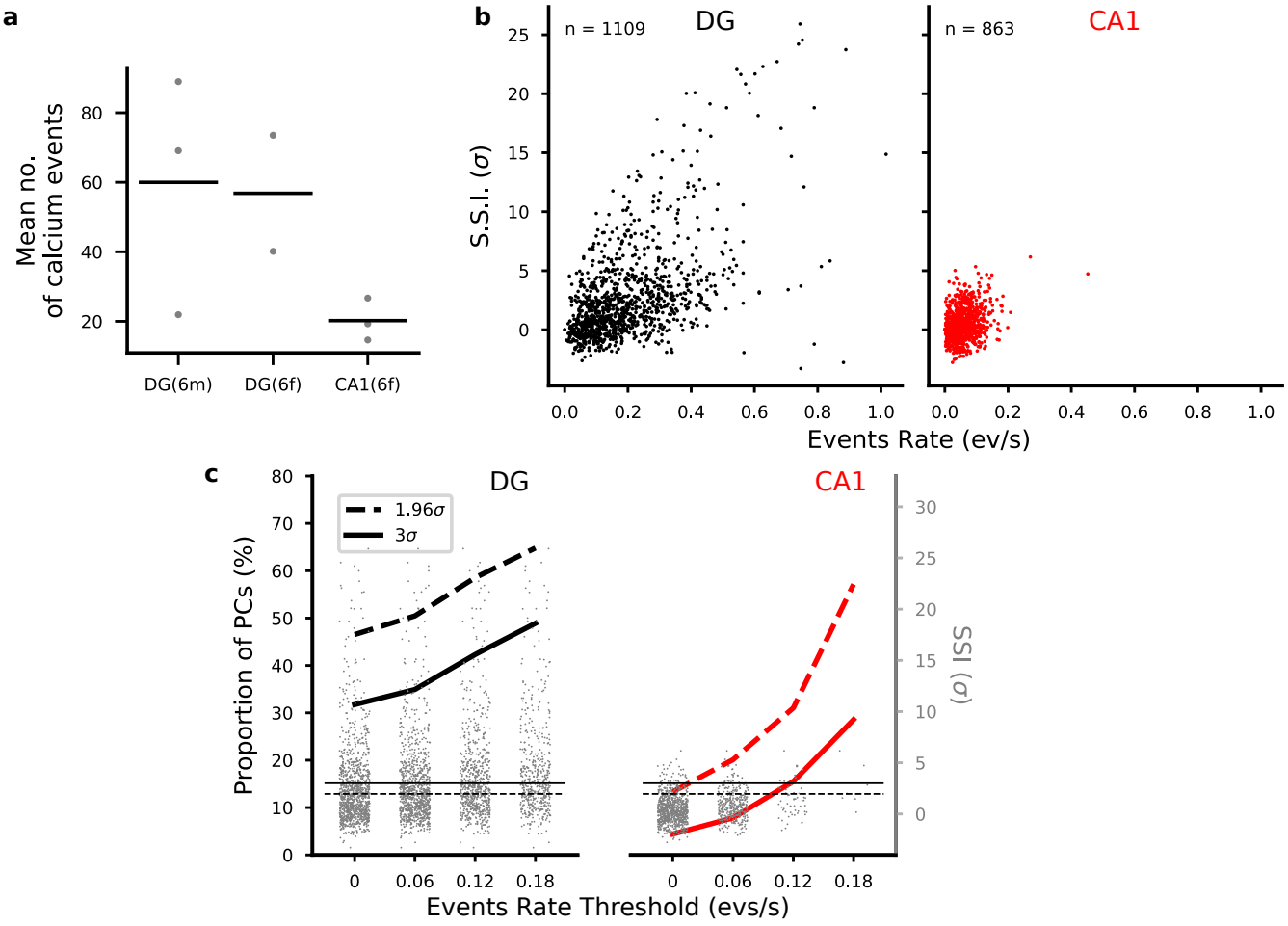

Figure S3. Reported values of overall activity levels and hence place cells ratios in CA1 and DG vary considerably across labs. This is due to a combination of factors, including different recording techniques, different statistical methods used to assess spatial tuning and differences in experimental protocol. Here we compare the significance of spatial information in the populations of CA1 and DG cells we identified and relate it to the activity levels in these areas in general, and then compare the levels of activity we observed with the existing literature. In both CA1 and DG areas, place cells were the minority of cells, however place cell ratios where very different (36% in DG, 4.2% in CA1). One reason for this difference could be in the lower activity levels we found in CA1 cells compared to DG. Place cell studies in rodents hippocampus have found higher place cell ratios when animals run through a linear track (Ziv et al., 2013). Talbot and colleagues (Talbot et al., 2018) found that 47.8% of the recorded CA1 cells were place cells but they used spatial information measure alone (Skaggs et al., 1993). The lower ratio of place cells in our report could be due to a higher sensitivity for sparsely active cells of our method with respect to standard extracellular recording approaches. Despite the likely presence of some false negatives in the single calcium event detection, even cells with a very small number of sufficiently strong calcium events are detected by the source extraction algorithm we used. Also in other recent studies, calcium imaging has revealed larger proportions of sparsely active cells suggesting these populations have been largely underestimated in electrophysiology studies (Dipoppa et al., 2018; Tang et al., 2018). In

particular, the CNMF-e algorithm uses both spatial footprints and temporal components to separate cells from background contamination and from other sources (Zhou et al., 2018). Instead, typical clustering methods for extracellular recordings may bias identification towards more densely active cells because these are more likely to be pulled out from the noise through clustering. This bias ultimately results in an overabundant representation of place cell ratios in the population. To validate our hypothesis, we first confirmed that the difference in activity levels between DG and CA1 cells was not due to the different calcium indicator used. **a**) Control that differences in activity levels between DG and CA1 are not due to the different calcium indicator used. We analyzed imaging data from granule cells in dentate gyrus after viral injection of GCaMP6f in two additional mice and found activity levels perfectly compatible with our GCaMP6m recordings, hence we conclude that differences in activity levels were not due to the calcium indicator used (**a**). Mean number of calcium events as extracted by CNMF-e across the whole session in all the DG(6m), DG(6f) and CA1(6f) mice. Each dot corresponds to one mouse. Bars correspond to mean values across mice. We then looked at the relation between activity levels in terms of event rates and spatial information (see (**b**). **b**) Scatter plot of the significance of spatial information (SSI) as a function of event rates across a session for all cell populations in DG (left) and CA1 mice (right). **c**) Proportion of place cells in DG (black, left axis) and CA1 (red) after removing cells based on their mean activity. We removed weakly active cells in both CA1 and DG populations with a varying threshold to see if we could artificially bias either population to artificially high place cell ratios. For each threshold on activity, we report the SSI for the selected cells (gray dots) and the corresponding proportion of place cells in the remaining population (black: DG; red: CA1). In our study we used a strict threshold of three standard deviations for significance (solid lines) however we also report the same proportion for a threshold of 1.96σ, corresponding to a 95% significance for spatial information as considered in some place cells studies (Kinsky et al., 2018). In line with our hypothesis, we could obtain higher place-cells ratios by choosing an appropriate threshold on activity. For example, if we didn't record any signal from cells with less than 0.12 events per second, we would have observed about 30% of place cells in CA1 and 40% of place cells in DG across all mice. The horizontal dashed and solid lines correspond to the significance thresholds for SSI for the corresponding curves (gray scale on the right-hand side of the plot). Solid lines: proportions computed using 3 standard deviations of the shuffled SSI distribution to assess whether a cell's SSI was statistically significant. Dashed lines: proportions computed on 95-th percentile threshold, i.e., using 1.96 standard deviations. The dots represent the SSIs for all cells with an activity that was larger than the given activity threshold on the x-axis. In conclusion, although differences between DG and CA1 place cell number may depend on several other factors such exposure times during habituation, our analysis suggests that the main drive for place cell counts and differences between these two areas comes from the different activity levels. In particular, a non-negligible proportion of CA1 cells are more sparsely active than previously reported and therefore CA1 place cells may have been over represented in place cells studies. We encourage researchers to adopt a more systematic assessment of the significance of the spatial information as we suggest in our work and to always report at least the proportion of place and non-place cells that are found. See also S1 for a discussion on the variability of reported rates in the literature.

Supplemental Material S4: Dependence of decoding performance from the number of cells in CA1 and DG; Related to Fig. 2

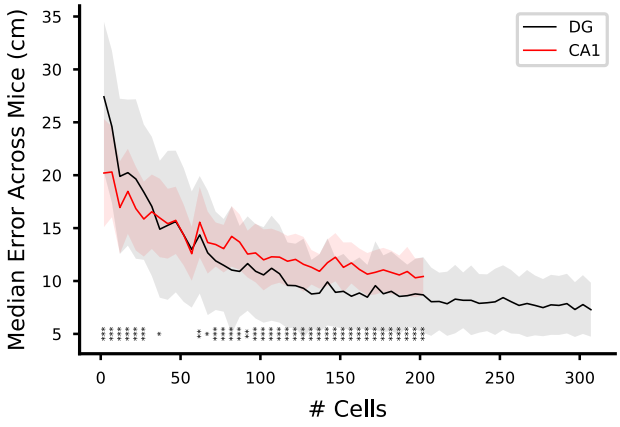

Figure S4: Dependence of decoding performance from the number of cells in CA1 and DG. For each given number of cells, a random selection of cells from the pool of identified cells is used to decode the animal's position using 10-fold cross-validation. The selection is repeated 10 times for each animal and for each number of cells. We then pooled all the results from all the animals to report the mean (solid line) and standard deviations (shaded area) of the median errors for position decoding for each given number of cells (independent samples t-test for significance, *p<0.05, **p<0.01, ***p<0.001). The maximum number of cells we used corresponds to the minimum number of cells we could record from any animal in DG or CA1.

Supplemental Material S5: Chance level decoding performance; Related to Fig. 2

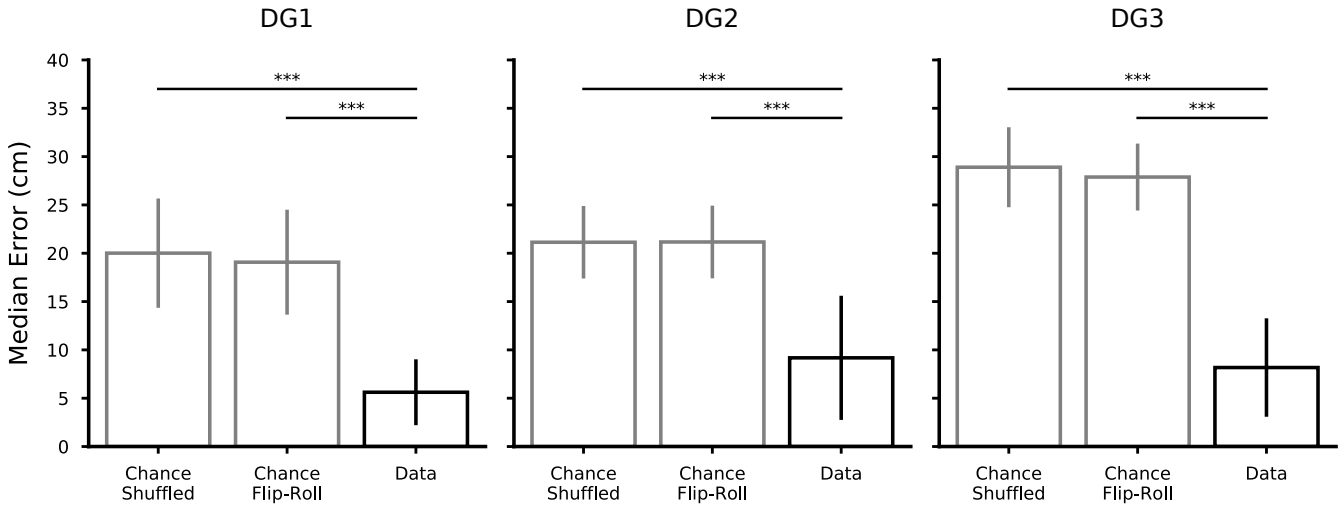

Figure S5: Chance level decoding performance. It is important to assess the performance of the decoder to a chance performance, i.e., the performance of a prediction that is not based on the recorded activities but solely on the behaviour of the mouse. The important factors that must be considered in the case of decoding continuous behaviour, are represented by the auto-correlations of the neural activities and of the behaviour. Therefore, for evaluating the chance performance one must take into account these factors in order to avoid underestimating the performance obtained by chance and therefore overestimate the ability of the decoder to extract the information under consideration. We computed the chance performance by training a position decoder on artificially manipulated data. We used two methods: the first method consisted in shuffling the calcium events in time by keeping the overall instantaneous event rate across the population constant ('chance shuffled'). The second method consisted in applying a time reversal operation to the animal x-y trajectory and then shifting it in time by a random amount. Data points falling outside the temporal window due to the shifting where reinserted from the other end of the period as in a torus.We then trained the decoder on a cross-validated manner in which the data was split into 10 chunks of contiguous data, 9 of which were used for training and the remaining one for testing the decoder. For both methods of shuffling, therefore, we obtained the 10-fold cross validated performance which we aggregated all together in each animal and applied a Mann-Whitney U non-parametric test to compare the resulting distribution to the original data (***$p<0.001$). In this figure we report the performance (mean $\pm$ s.d.) for the position decoders for all the animals for the two chance levels in grey and for the original data in black. The decoding performance is significantly above chance for both choices of chance level and the difference between the two methods is negligible. Gray bars: chance level performance (mean $\pm$ st. dev.) for shuffling calcium events in time ('chance shuffled') and for the 'flip-roll' strategy, which keeps the original correlation structure of the data intact while disrupting the association between behavior and calcium recordings. Black: decoding performance on the original data (Mann-Whitney U, ***$p<0.001$, 10-fold cross-validation)

Supplemental Material S6: Comparison of different decoding strategies; Related to Fig. 2

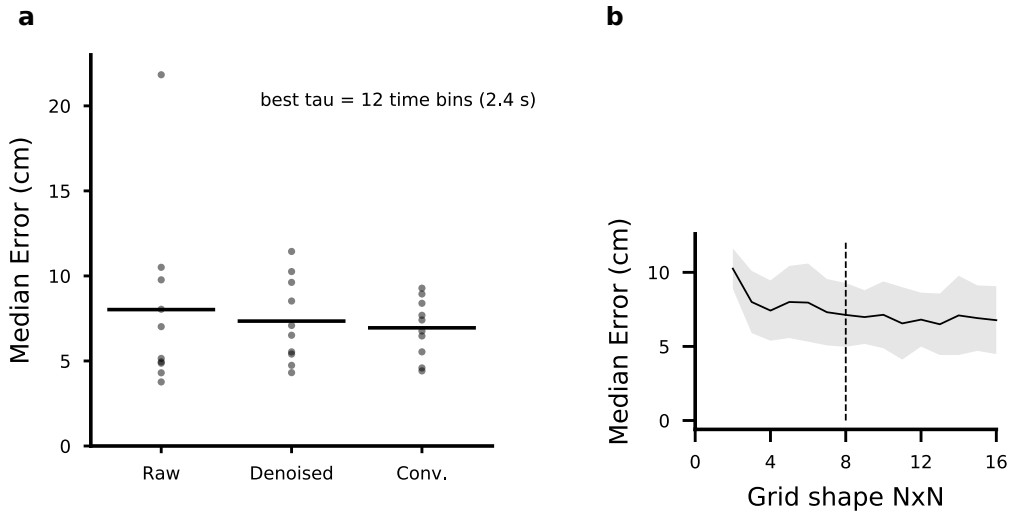

Figure S6: Comparison of different decoding strategies. To decode the animal's position and direction of motion, we used a battery of linear decoders trained on pairs of locations obtained by dividing the arena into a series of discrete locations and assigning the location label to each time bin in the data (see Methods). In this way we could use the weights of the single decoders, properly combined, to obtain an overall importance index for the cells in the population. To verify that our decoding performance did not depend on the specific choices we made, we performed the following comparisons. First, we verified that the choice of signals used for decoding did not have an impact on the position decoding. Using the convolved calcium events instead of the raw traces as described in Fig. S1 does not have a significant effect on the performance of our decoder (data for a representative mouse DG3 in **a**). We also verified that the choice of the binning to discretize the arena did not have a significant impact on the decoding performance, which allowed us to keep the same bin size on all analyzed sessions (**b**). **a**) Median error for decoding position on 10-fold cross validation for raw traces as extracted from CNMF-e, denoised traces, and convolved traces using calcium events convolved with an exponential kernel. **b**) Decoding error as a function of the number of spatial bins in which the arena is discretized (representative DG mouse, DG3, mean and st.dev. over 10-fold cross-validation). For each data point, the arena is divided into $N \times N$ squares of same size, and the size N is reported on the x-axis. The dashed line corresponds to the choice of an 8×8 grid of locations of approximately 6cm side as used in the article. Although smaller bins (larger N) might allow for a more precise determination of the position of the animal, they would also contain a smaller number of data points and therefore limit the accuracy of the decoders. The median error decreases rapidly when N is small, but then it decreases slowly and it becomes almost constant for $N > 8$.

Supplemental Material S7: Position decoding error and speed are negatively correlated; Related to Fig. 2

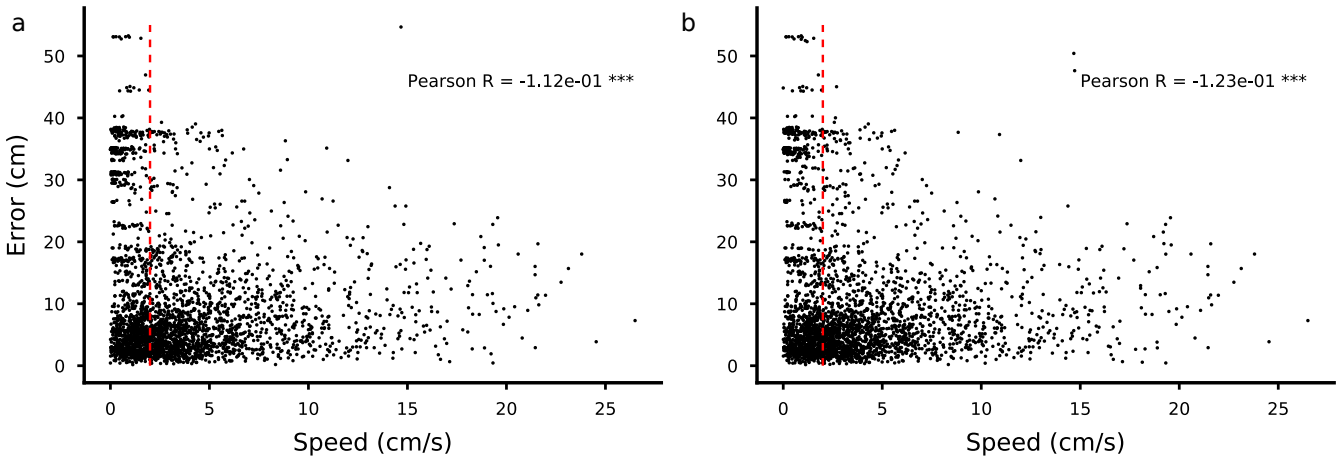

Figure S7: Position Decoding error and speed are negatively correlated. A mouse's movement and immobility states are often considered two distinct behavioral conditions, characterized by distinct neural activities, and have been therefore considered separately for analysis in the literature. In our work, we focused our analysis on the data collected during movement following a typical procedure in the hippocampal research, see for example (Leutgeb et al., 2007). We then asked whether the accuracy the decoded position was related to the speed of movement. First, we decoded the position of the animal using only the datapoints during movement for training and testing on all datapoints. **a** Decoding error at different decoding speed using only datapoints detected as movement for training. Each point corresponds to one time bin (***$p<0.001$). On the x-axis, we plot the instantaneous speed of movement in one animal (data from the dentate gyrus of a representative animal DG3). The decoded position is taken as the centre of the discrete location that is selected by the decoder in each time bin. The red vertical line corresponds to the value we used as a threshold to distinguish between movement and immobility on the training data. **b**) Same as in **a** but all the datapoints have been used for training (***$p<0.001$). To verify that the negative correlation was not due to the fact that the datapoints corresponding to immobility were not used in the training set, we repeated the procedure using all the datapoints during training. In both cases we found a significant negative Pearson-R correlation, suggesting that the animal's position was encoded more accurately during locomotion.

Supplemental Material S8: Temporal stability of the importance index; Related to Fig. 4

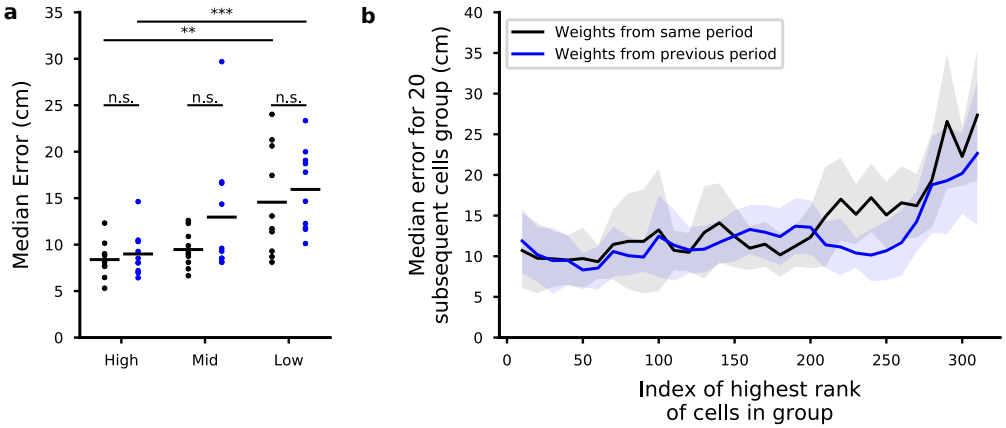

Figure S8: Temporal stability of the importance index. We introduced an importance index that quantifies the contribution of each cell to the population code. We know from the analysis in the main text that the ranking of the cells based on the importance index depends on the cells that are considered for decoding. So we know it is not an intrinsic property of the cell. However, given a population of cells that is used to estimate the importance index, it is interesting to ask whether the rank of a cell changes over time. To answer this question, we ranked the cells from the dentate gyrus using the first 10 minutes of data from the 30 minutes trials. We computed the decoding performance using the 20 best cells (ranked high), 20 cells that ranked in the middle and the 20 cells in the lowest part of the ranking. We then decoded position using each group of cells in different chunks of data of 10 minutes at the beginning of the session and in the next 10 minutes of data. If the cells' ranking changed, we'd expect a different performance for each group of cells since they were defined using the ranking in the first 10 minutes. Instead, we saw no significant changes in decoding performance for the high, middle and low-ranked cells. This indicates that the importance of each cell is relatively stable over time, at least when the decoding performance is considered. **a**) Decoding performance for three groups of cells 20 cells divided by their ranking in importance index. Black: decoding performance for 10-fold cross-validation within the first 10 minutes of data. Blue: decoding performance in the following 10 minutes of data after training on the first 10 minutes (Mann-Whitney U, n.s. $p>0.05$, **$p<0.01$, ***$p<0.001$). **b**) Decoding performance for subsequent groups of 20 cells ranked by their importance index. Black: cells were ranked using the importance index computed on the first 10 minutes of data. Blue: cells were ranked based on the weights computed on the next 10 minutes of data. We applied a procedure like that of Fig.4 of the main text to better assess whether the ranking in one time-period could be used to rank cells in another time period. We first computed the decoding performance in the first 10 minutes of the trial using subsequent groups of cells ranked by their importance index computed in the same time window. We then compared these results with the ones obtained by decoding position during the next 10 minutes of the trial. The two plots largely overlaps, therefore the importance indices obtained in one temporal window can be used to select the important cells in another temporal window.

Supplemental Material S9: Importance index and spatial information; Related to Fig. 5

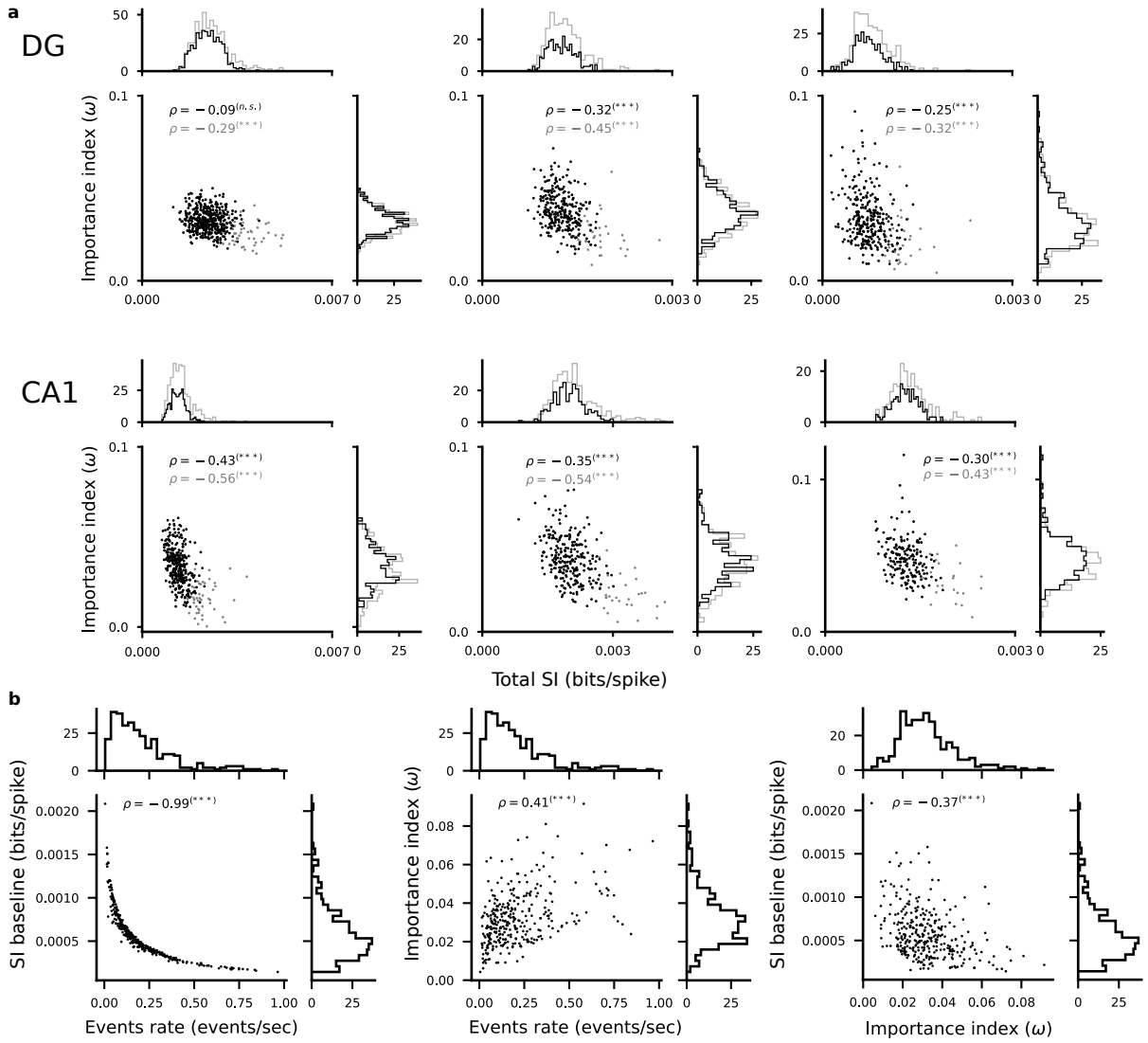

Figure S9: Importance index and spatial information. Limited data, such as low firing rate, introduce a systematic bias in the estimation of the information content in neural activities (Panzeri et al., 2007). Our results show that a more sensible choice is the one of the SSI, a measure of the strength of the spatial information after compensating for this bias. Furthermore the importance index was weakly correlated the SSI. We also showed that importance index and SSI are correlated. Here we additionally show that importance index and spatial information have a low and even negative correlation factor (Pearson-R correlation, ***p<0.001). **a)** Relation between importance index and spatial information for all recorded animals (grey: all cells, black: cells with more than 10 calcium events identified, Pearson-R reported in each plot, ***p<0.001). Each dot in the plots represents one cell. The histograms for each of the quantities are shown on the sides of each plot. This result is further evidence that the spatial information alone may be a misleading factor in estimating the contribution of a cell to encoding position if not validated through an assessment of its statistical significance. **b)** Left: Relation between the baseline spatial information per cell (obtained by shuffling events in time for each cells) and events rates (left), importance index and events rate (middle) and spatial information baseline and importance index (right). In each panel,

each dot corresponds to one cell in a representative animal (DG3). The spatial information baseline is computed as the mean spatial information over the shuffled distribution. Spearman-ρ correlation values and significance are reported in each plot. The baseline and the importance index are negatively correlated due to their relation to event rate. Intuitively, the negative correlation between spatial information baseline and importance index is due to the fact that low activity introduces spurious stimulus-dependent differences in the response probabilities used to compute information (Panzeri et al., 2007; Kelemen and Fenton, 2010). These make the neuron seem more informative (high spatial information) despite its low importance in the population for position encoding.

Supplemental Material S10: Decoding performance using only place cells; Related to Fig. 4

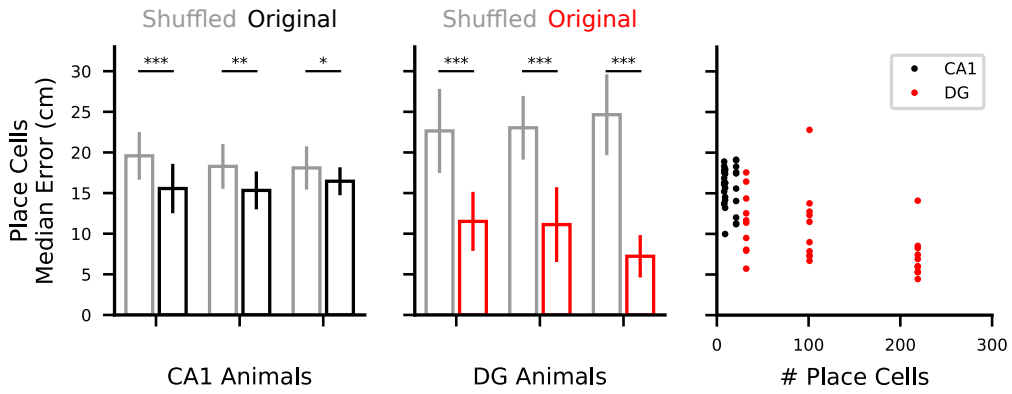

Figure S10: Decoding performance using only place cells. Using only place cells, the decoding performance is close to the best performance we obtained in DG (left, chance performance in grey) but rather limited in CA1 (middle panel). This is because the decoding performance depends on the number of place cells used, which is limited in CA1 (right). (Mann-Whitney U, *p<0.01, **p<0.005, ***p<0.001.) Left and Middle: Median decoding position error using only place cells for CA1 (left) and DG (middle) animals (each bar represents mean and st.dev.; Mann-Whitney U test, *p<0.05, **p<0.005, ***p<0.001, see Methods). Right: Median decoding error for all CA1 (black) and DG (red) animals for each 10-fold cross-validation within the 10 minutes sessions aligned to the corresponding number of cells isolated in the corresponding FOV.

Supplemental Material S11: Control for correlation between movement direction and position; Related to Fig. 7

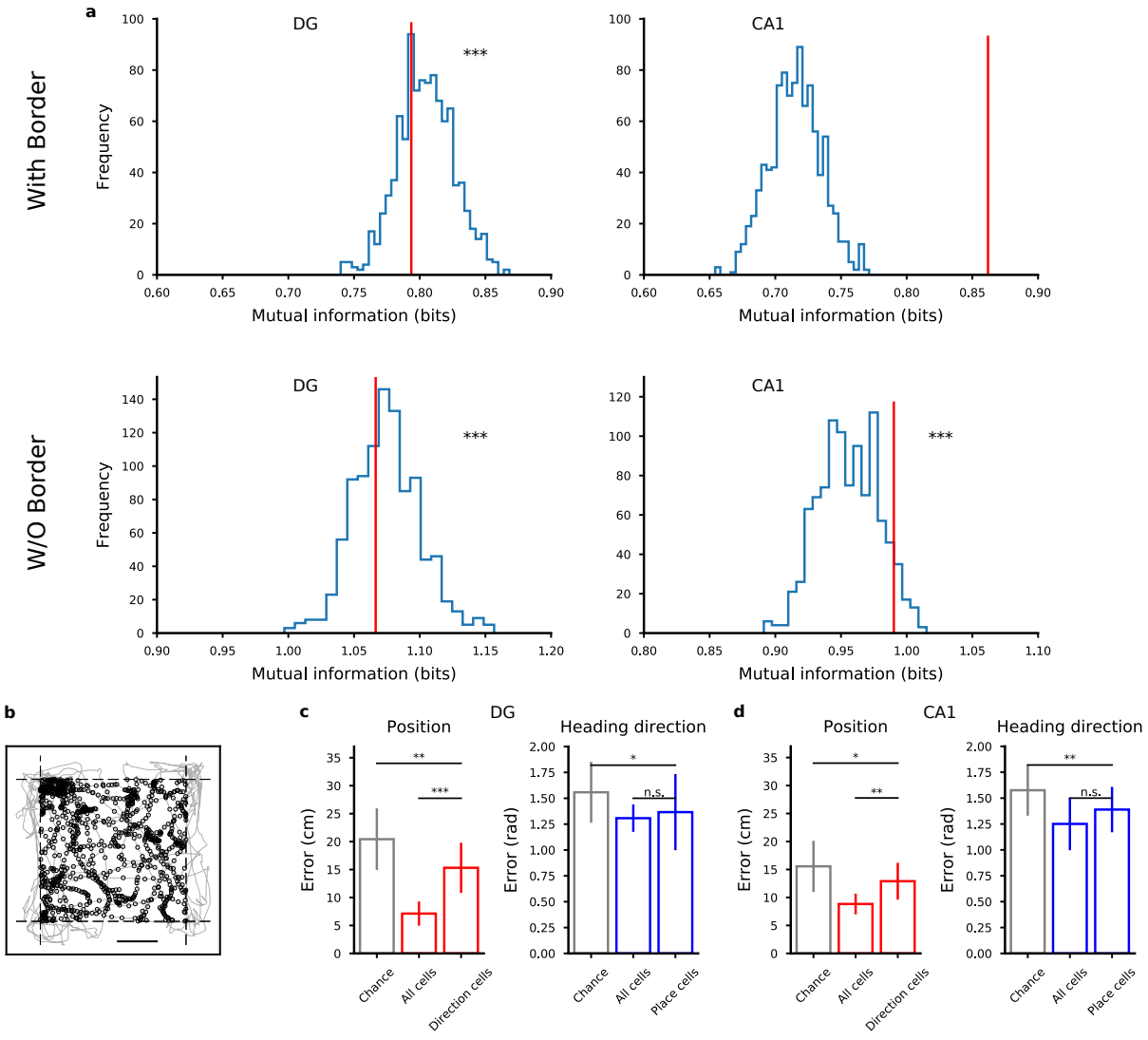

Figure S11: Control for correlation between movement direction and position. We showed in the main text that we can separately decode position and direction of motion and that the neural code for these two variables appears distributed across the neural population (Fig. 7). Our strategy was to show that the important cells for decoding position could be used to decode the direction of motion. However, it could be that the reason why direction could be decoded using the important cells for position was that direction and position were correlated. Indeed, when the animal is very close to one the border of the arena, the direction of movement can only assume values that point towards the inside of it. The decoder could in principle use this information to decode direction of motion from the important cells for position, therefore confounding our results. To exclude this possibility, we computed the mutual information between direction of movement and position and compared it to the distribution obtained with 500 shufflings of data from the same session from the animal with the most homogeneous coverage of the environment. **a**) Mutual information between position and direction of motion for shuffled data in the same session (blue histogram) and real data (red vertical line). Top panels: original data with positions close to the arena walls. Bottom: filtered data as in **b**. DG in top left panel, CA1 in top right panel, same animals

as in Fig. 7. In the shuffled data, position was time reversed and shifted similarly to what was done for evaluating the chance performance (see also Methods in the main text). The mutual information in the real data (red line) lies within the distribution of mutual information for shuffled data (blue histogram, one sample t-test, ***$p<0.001$) for DG, hence the knowledge of one variable is not enough to predict the value of the other variable, but not for CA1. We attribute this difference to the effect of the borders in the CA1 experiments, in which a rectangular arena of half the size than the one for DG mice was used. **b**) To further verify that the correlation between direction of motion and position did not explain the results of Fig. 7b, we repeated the decoding analysis after removing from the data all the positions that were recorded in the spatial bins closest to the walls of the arena. Here we show the trajectory of a representative DG mouse (DG3). Positions close to the walls of the arena, i.e., beyond the dashed lines, were eliminated from the data. Grey: mouse trajectory. Black dots: positions included in the analysis. Same DG mouse as in Fig. 7. After this manipulation, the mutual information between position and head direction becomes indistinguishable from the chance distribution in both DG and CA1 (bottom panels in **a**) and we were still able to decode position from the most important cells for direction and vice versa, as we did for the original data in the text, for both DG and CA1 mice (see **c** and **d**). **c**) Decoding error for important cells for position (left) and direction (right) from a representative DG mouse, same procedure as in Fig. 7b in the main text but after removing border data as in b. **d**) Same as in **c** but for the CA1 mouse of Fig. 7b.

Supplemental Material S12: Temporal stability of the decoder; Related to Fig. 2

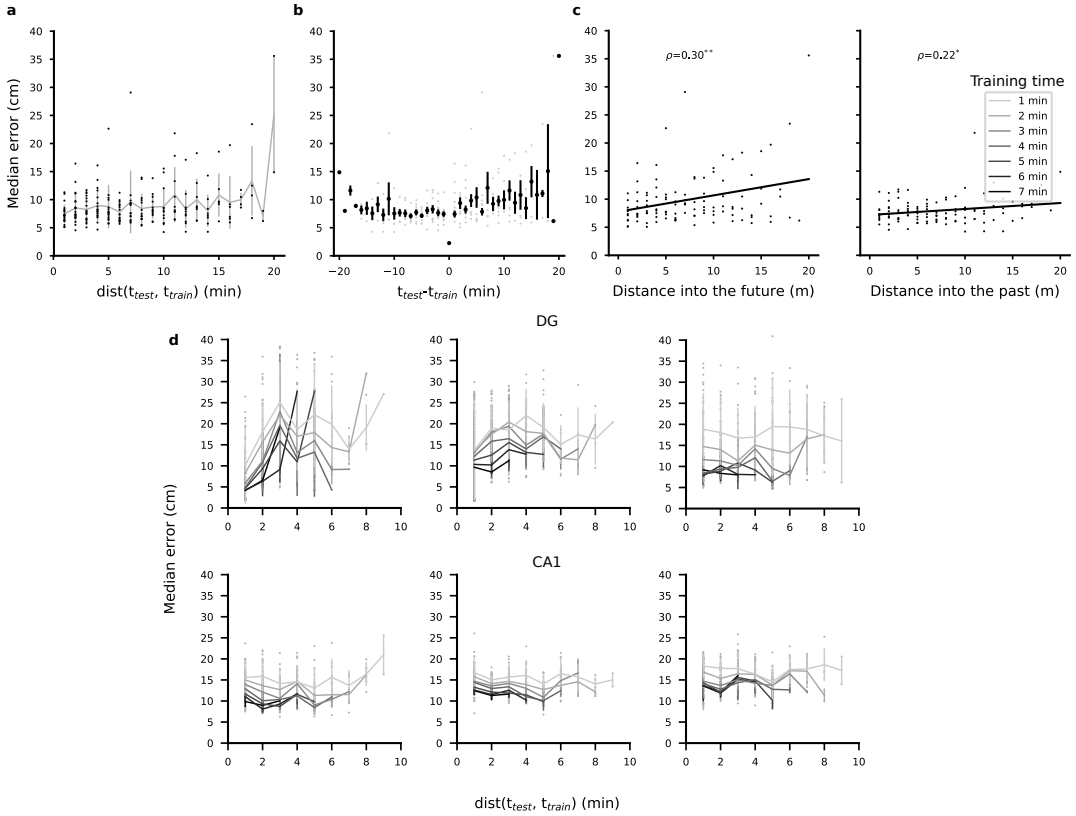

Figure S12: Temporal stability of the decoder. We asked whether the representations for position changed within a recording session. To do this, we used the 30 minutes trial available in our data (animal DG3, data from the dentate gyrus). We trained a position decoder on a shifting window of 10 minutes of data and tested in on a shifting window of 1 minute of data. **a**) Position decoding performance as a function of the distance in time between training set and test set. Each dot corresponds to one choice of training and test set, the grey line joins the mean values on pooled data for each bin of distance in minutes, bars are standard deviations within the binned distances. The performance of the decoder is remarkably strong and stable for a period of time extending to the available 20 minutes in the trial, although with some slow degradation towards the more distant time windows (**a-c**). This result shows that position is encoded with strong accuracy on the populations of cells whose coding properties extend long in the trial and it seems compatible with results in literature showing that the representation of position is stable across sessions despite the large degree of variability on the tuning properties of single cells (Ziv et al., 2013). **b**) Data as in (**a**) but datapoints for which the test precedes the training set are on the negative side of the x-axis. We looked for evidence of a difference between testing the decoder in past time periods and future time periods with respect to the training period. To determine whether decoding data temporally preceding the training data (past) was different from decoding following data (future), we linearly interpolated the decoding performance and obtained two weakly significant Pearson correlation factors of 0.3 for the future (**p<0.01) and 0.22 for the past (*p<0.05). Given the weakness of our statistical test we couldn't detect a difference between these two trends. Therefore, the performance on the test data does not depend on whether the decoder was trained on a past or on a future interval. **c**) Data as in (**b**) where the points have been fitted with a linear regression for test data that is in the future with respect to training data (left) or in the past (right). Reported correlation values correspond to the Pearson's R correlation (*p<0.05,

p<0.01). **d) We compared the stability of the decoder in DG and CA1 animals by repeating the analysis within 10 minutes of the session for all animals. For short training times (down to 1 minute) the decoding error was relatively stable on most animals but was much lower that the best decoding performance we obtained for those animals. We then increased the training time to verify that this was not due to a low temporal stability of the code but rather to the small size of the training set. In all animals the performance drastically improved with training set size. These results further suggest that both in DG and in CA1 the code for space is relatively stable within a session and further studies will further investigate this important aspect of the spatial code. Top: DG animals. Bottom: CA1 animals. Each panel corresponds to one animal. Each line in different shades of gray corresponds to a different choice of training time (test time is always 1 minute). Each dot represents the median error for one choice of training and test time. The decoding performance for different test periods (y-axis) are sorted according to the temporal distance between the test period and the closest datapoint in the training period (x-axis), regardless of whether the test period preceded or followed the training period (t_{test} is the time of the first time bin of the test period, analogously for t_{train}). In each one-minute bin, mean and st.dev. are computed and plotted.

Supplemental Material S13: The information about position is distributed: distributions of importance indices in DG and in CA1; Related to Fig. 5

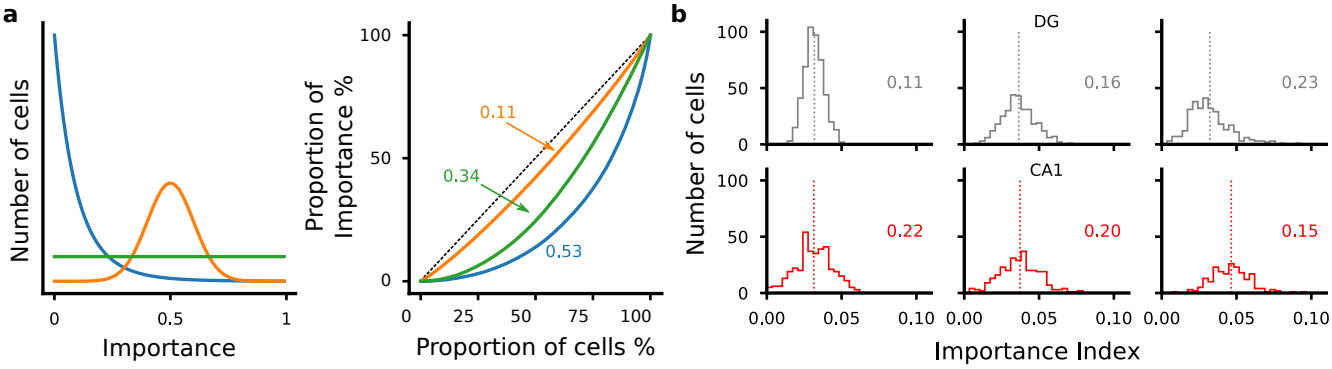

Figure S13: The information about position is distributed. One way to compare how populations of cells represent information is to compare the distribution of their importance indices. We computed the Gini index on the distribution of importance indices to assess whether the information about position was distributed across multiple cells or instead carried by a small minority. The Gini coefficient (Gini, 1912) is a measure of statistical dispersion among values of a certain frequency distribution. It is often used as a measure of inequality of income levels. Low values correspond to a large degree of equality, i.e., people share large proportions of the total wealth of the population. High values often identify situations in which wealth is concentrated on a small number of people. In the same spirit, we used the Gini coefficient to evaluate how the importance index was distributed across cells in our recordings. **a**) To illustrate the dependence of the Gini value from the underlying distribution, we first describe how different distributions result in different Gini indices. We considered three different distributions, from high inequality to low inequality. Here we show the distributions (left) and corresponding Lorenz curves (right) for simulated data: a uniform distribution (green), a distribution with low inequality (orange), as in the case of a distributed code, and one with strong inequality (blue) as in the case of a specialized code with few important cells and many unimportant cells. The first one is the Pareto distribution, a power-low distribution commonly used to describe social phenomena, including distribution of wealth. In this distribution, importance is concentrated on a small fraction of cells whereas the vast majority of cells have a low importance, as in a specialized code (left panel, in blue). We also considered a uniform distribution of importance whereby a cell importance is determined by chance (left panel, in green). Lastly, we considered a Gaussian distribution centered around a middle value. This case results corresponds to a highly distributed code since the vast majority of cells shares a similar amount of importance (left panel, in orange). The first step for computing the Gini index consists in evaluating the Lorenz curve, which represents what proportion of cells shares a certain proportion of importance. The area under the Lorenz curve is then compared to an ideal situation whereby all cells have the exact same amount of importance, resulting in a straight Lorenz curve, and hence in an area under the curve of 1/2 (dashed black line). Finally the Gini index is computed as the proportion of the area defined by this ideal distribution that is not covered by the Lorenz curve of the data. In the right panel we plot the Lorenz curves for the three distributions we considered and their respective Gini index values. As expected from intuition, the most equal distribution is the Gaussian distribution (right). This distribution results in an extremely low value of Gini (colored inset values on the right). **b**) The Gini index and the distribution of the importance indexes across neurons. We computed the Gini index on the values of importance indices for both DG and CA1 and found low values of Gini (between 0.11 and 0.23), indicating that the code is highly distributed.

Histograms of importance indices from the data across all DG (top) and CA1 animals (bottom). Gini index numbers are reported in each plot.

Supplemental Material S14: Procedure to destroy correlations and controls; Related to Fig. 8

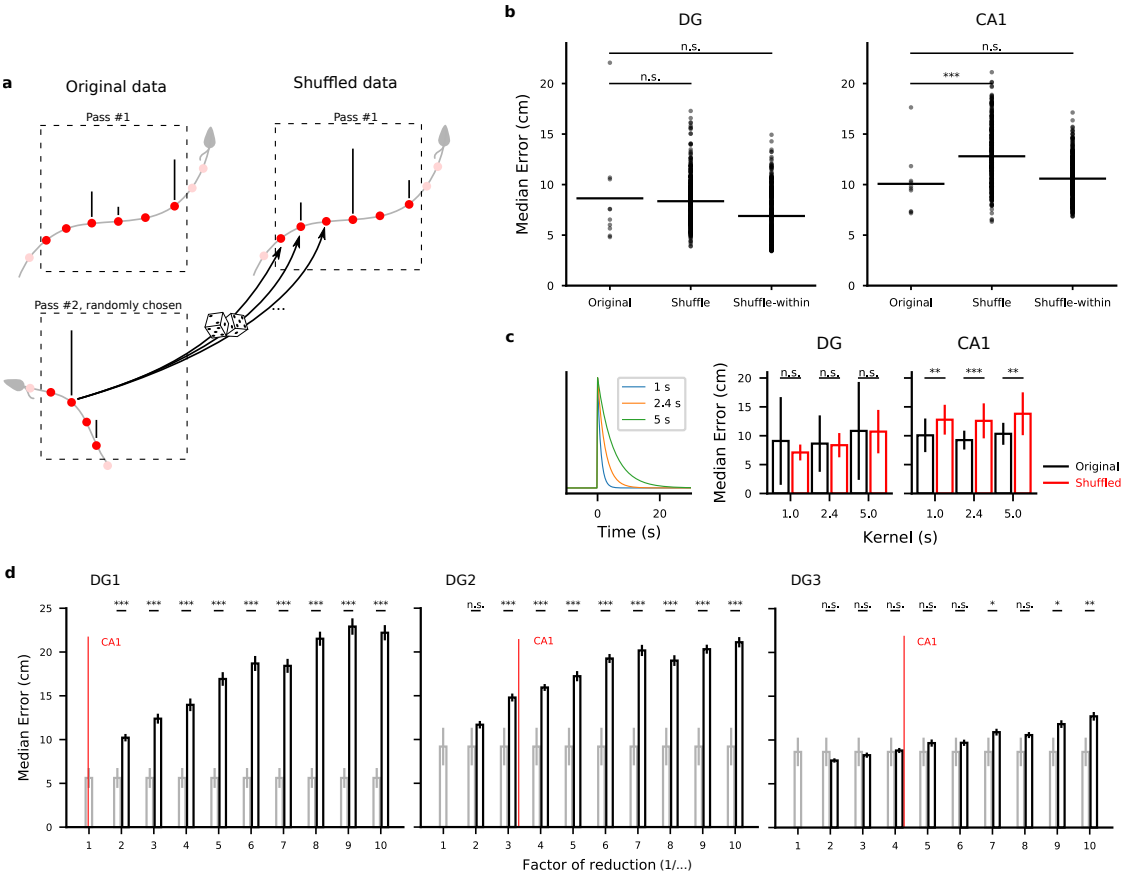

Figure S14: Procedure to destroy correlations and controls. **a**) Schematic description of the procedure for destroying correlations in one discrete location in the arena. The red dots correspond to the time bins of the recorded calcium events along the mouse trajectory (gray line) through the area of interest (dashed line). Time bins in which an event was recorded are identified with a tick mark on top of the red mark. Light red marks correspond to datapoints recorded before the animal enters and after it exits the area and are not considered. For each location, we focused on all the passes of the animal through that location and identified the entrance and exit times of each one. Then for each pass, we substituted the recorded calcium events in that pass with the ones of an other randomly chosen pass. To account for the different number of time bins in the different passes, we randomly sampled from activity from the second pass. The procedure showed an effect in CA1 data but not in DG (see **b**). This suggests different coding properties as discussed in the text. **b**) Decoding error for original data, shuffled data as in a and shuffled data when the new data was generated from the same pass to control for artifacts due to sampling. Left, DG. Right, CA1 (Mann-Whitney U test, ***p<0.001). To verify that random sampling did not introduce artifacts that would affect decoding performance, we verified that when we sampled from data within the same pass the decoding performance was not significantly affected ('shuffle-within'). After this procedure, we were confident that any effect we would observe was only due to having destroyed correlations among neurons. When we shuffled data within the same pass, we saw no effects in either CA1 nor DG data. **c**) The effects of destroying correlations do not depend on the choice of the kernel time scale that was used to convolve the sequences of events. Left: shape of the kernel function used to convolve the calcium

event data. In the main text we used 2.4 s. Right: decoding error for the original and the shuffled data in representative DG (left) and CA1 (right) mice for different choices of kernels, as in Fig.8 in the main text (Mann-Whitney U test, **p<0.005, ***p<0.001). **d**) The effects of destroying correlations depend on sparsity. Sparsity was induced by randomly selecting a given proportion of detected calcium events. We report the performance on shuffle data (black bars) compared to the original decoding performance (grey bars) for all the DG animals at different levels of sparsity. The red vertical line corresponds to the equivalent factor to obtain similar activity levels to those of the CA1 animals, for reference (Mann-Whitney U test, *p<0.05, **p<0.005, ***p<0.001). When we artificially reduced the number of events in our DG mice we observed a significant decrease of performance for increasing sparsity, however in two of the three analyzed animals the reduction was not significant at a level of sparsity that was comparable to CA1.

Supplemental Material S15: The effects of dimensionality on neural correlations and decoding performance; Related to Fig. 8

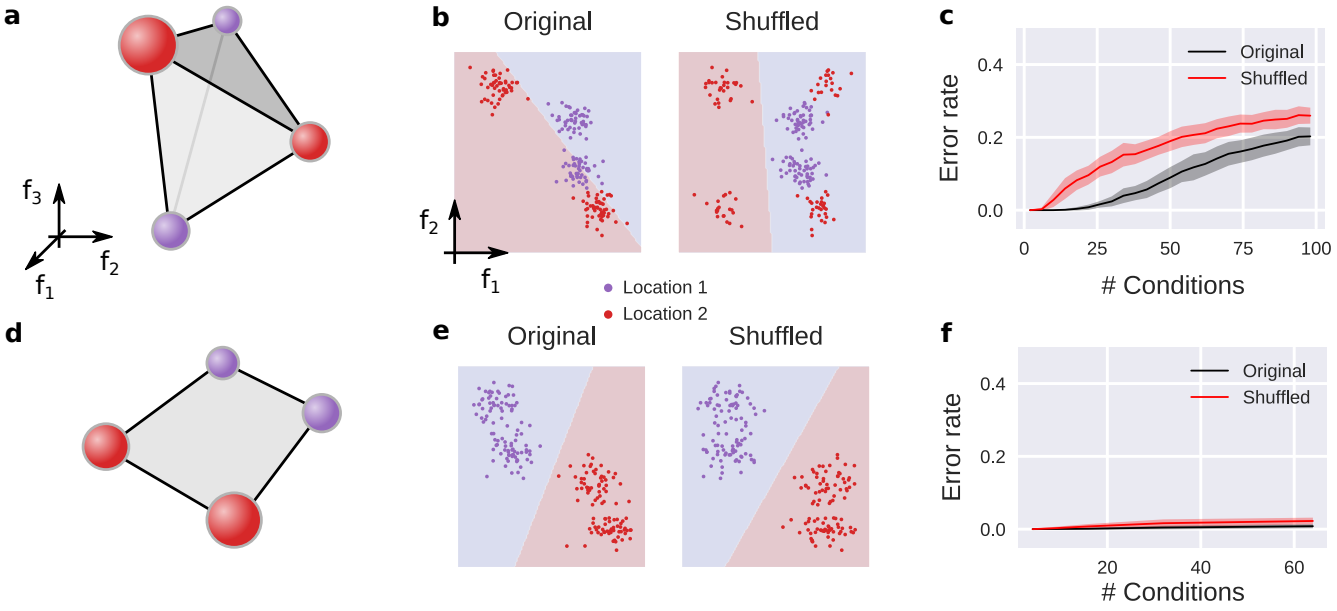

Figure S15: The effects of dimensionality on neural correlations and decoding performance. In our work we suggested that, at least in CA1, the neural correlations should not be considered as noise but rather as a reliable signal. We showed that destroying the neural correlations had an effect on the ability of our decoder to predict the animal position from CA1 activities. The effect was smaller in DG. In general, the effect of destroying correlations on decoding performance depends on the geometry of the neural representations (Abbott and Dayan, 1999), i.e., how the patterns corresponding to the different combinations of behavioral variables are distributed in the space of population activities. How do the effects of destroying correlations depend on the geometry of the neural representations? Here we studied in simulations a few cases to understand under what assumptions we should expect a disruption of the performance when the correlations are destroyed. We started by considering neural representations in the space of population activities. These are determined by different combinations of discrete behavioral variables, each of which corresponds to an experimental condition. In the experiments, some of the behavioral variables would be under direct control, such as the discretized animal's position or movement direction, while others are not. For each condition we generated a prototype pattern of population activity and a cloud of points around it by adding isotropic Gaussian noise of zero mean and unit variance (see S15). This cloud of points thus represents a set of recordings for that condition, similar to the ones generated by multiple passes of the animal through a given location in the environment as in Fig. 3 in the main text. Following our hypothesis, we generated many such clouds in the space of neural activities for the various combinations of hypothetical behavioral variables the neural activity is subject to and assigned half of them to one class, corresponding for example to one location, and half of them to another class, corresponding to a second location. The fact that the representations of each location consists of multiple conditions, naturally induces correlations among neurons in the population. This is because the activity of one neuron is tightly coupled to the activity of other neurons as prescribed by the patterns corresponding to that location. To study the effects of destroying correlations between, we first considered the cross-validated ability of a linear decoder to correctly classify the two hypothetical locations. Then, we shuffled the data in a way similar to what we described in the Methods to destroy the correlations among

neurons while keeping the information on location. Briefly, for each location, we chose one observed level of activation in that location for each neuron independently. This effectively violates the coupling imposed by the patterns of the conditions representing one location and therefore destroys correlations among neurons. Importantly, while this manipulation effectively destroys the correlations among neurons, it does not affect the statistics of the activity for each neuron with respect to each location, i.e., the spatial tuning of the neurons is not affected. We repeated this procedure multiple times to generate a new dataset of data without neural correlations. After this manipulation we trained again our decoder and compared the decoding performance with the one computed on original data, as we did for the neural recordings in the main text. To being our study of how the geometry of the neural correlations affects decoding, we considered two common scenarios observed in neural data (Rigotti et al., 2013; Bernardi et al., 2018), one in which the neural representations are 'unstructured' and one in which they are 'structured'. The first situation can be generated by randomly distributing the different conditions in the space of neural activities, as in **a**. For the second situation we considered abstract variables, i.e., the conditions were positioned at the corners of a low-dimensional hypercube randomly rotated in the space of neural activities as in **d** (Bernardi et al., 2018). To visually describe the situation, in **b** and **e** we show these two scenarios in the simple case of two neurons. We further restrict our description to the case of two conditions per location, which could correspond for instance to two directions of motion in the data. The colored regions separated by a straight line represent the decision function of a linear classifier trained on cross-validated data, hence the red points lying on the red regions are correctly classified as are the purple points lying on the blue region. Vice versa, the points lying on a region of different color are incorrectly classified. In the case of unstructured data, the four clouds of points appear randomly distributed (**b**) and, in this example, they are linearly separable. The multiple conditions impose correlations between the two neurons, in a way similar to what we discussed in Fig. 3 in the main text. Indeed, after destroying the correlations, the two location are no longer linearly separable and therefore the performance of the decoder can only worsen. The situation in the high-dimensional space of the population activities doesn't have such a straight forward graphical interpretation and so it is not straight forward to predict what would happen in the high-dimensional case. Thus, we performed simulations in the high-dimensional case (50 neurons) and compared the decoder performance before and after destroying the correlations for increasing number of conditions (**c**). We found that the effect of the correlations increases with the number of conditions (**c**), it reaches a maximum and then it decreases again for higher numbers of conditions. Moreover, when the number of conditions becomes comparable to the maximum number of neurons (50 in our simulations), the two locations may be non-linearly separable also before destroying correlations, depending on the specific random arrangement, and so the error rate increases also for the original data. In the case of structured neural representations, the situation is different (**d**, **e**, **f**). For this scenario, we generated the patterns in a way that follows abstraction as defined in (Bernardi et al., 2018) (**d**). As in the previous case, we first looked at a low-dimensional case of two neurons and two conditions per location (**e**). In this case the effect of the correlations is much weaker, as it is shown by the fact that also after destroying correlations the representations of the two locations are still linearly separable. To numerically explore the high-dimensional case, we performed similar simulations as before and varied the number of conditions while preserving the structure of abstraction (**e**). We observed that even for a large number of encoded variables the effect of destroying correlations was negligible in this case. Taken together, these results show that the geometry of the neural representations imposes neural correlations can be beneficial for decoding performance. In some of the cases we examined we also found that the effect can be negligible. It is important to note that there can be other scenarios that we did not consider here but even with the few simple cases we described we found several conditions that seem compatible with our experimental data. Furthermore, it is interesting to note that the type of neural correlations that we considered in these

models are not induced by a non-negligible covariance in the direction of the noise of neural activities. Rather they are induced by the presence of relatively small regions of the space of population activities that correspond to reliable representations to experimental conditions. It will the subject of further studies to investigate what classes of models can lead to such sparse representations. The impact of destroying neural correlations depends on the dimensionality of the neural representations and affects high dimensional representations more than low-dimensional ones. **b**) Two-dimensional projection of the activities recorded ad different times in each of the four conditions in **a**. The line separating the red and purple regions of the space corresponds to a linear separation of the red and purple locations. Left: original data. Right: datapoints obtained after destroying correlations as explained in the Methods. **c**) Decoding position, i.e., decoding performance for discriminating red from purple conditions, as a function of the number of conditions for original and shuffled data. Half of the conditions are randomly assigned to each class. **d-f**) Same as in **a-c** but the conditions are arranged on the vertices of an hypercube as in **d**.

Supplemental Material S16: The effects of shuffling on pairwise correlations; Related to Fig. 8

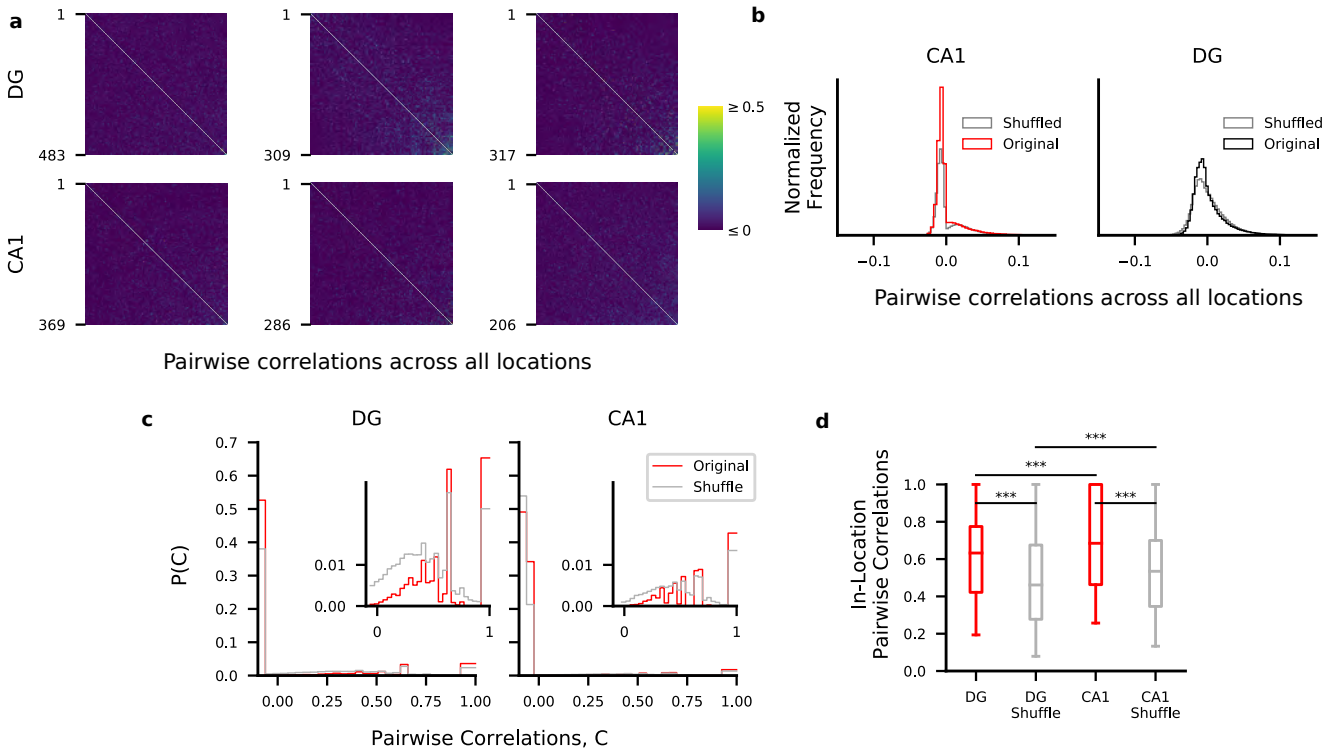

Figure S16: The effects of shuffling on pairwise correlations. Our procedure for destroying correlations shows a significant effect in CA1. To further characterize these effects, we studied the pairwise correlations before and after the shuffling. Although pairwise correlations only partially reflect the structure of correlations that is disrupted by our procedure for destroying correlations, because they include those induced by the similarity of tuning among cells, one may wonder if they would be sufficient to describe the situation. However, it is important to notice that pairwise correlations include those imposed by the tuning profiles of the cells (often referred to as signal correlations). In contrast, our procedure was specifically introduced to study the effect of correlations that appear on single passes through one location (often called noise correlations). When we destroy correlations, therefore, we disrupt those induced by the influence of variables other than position, i.e., a combination of noise and other potential signals that influence the cell activity such as movement direction. We show in Fig. 8 of the main text that our procedure does not disrupt spatial tuning of the cells. It could be therefore misleading to draw conclusions from the pairwise correlations alone about how they influence the conjunctive representations of other variables. Rather, we would like to make the point that by looking at the effects of destroying correlations while keeping the spatial tuning of the cells we can select hypothesis on the relations between the encoded variables, as we show in Fig. S14. Our analysis suggests that the structure of the pairwise correlations is only partially informative and that is important to use more sophisticated quantities to investigate the role of neural correlations. **a**) Here we show the pairwise correlation matrices for DG (top) and CA1 (bottom) animals computed using the calcium event time series across the entire session (three FOVs for each area). The values of the correlations are overall low in both regions, with medians around 0.1-0.2 but long positive tails (see the distribution in **b**, absolute values). **b**) Histograms of pair-wise correlations across all cells, computed across the entire session, pooled across mice (3 in DG and 3 in CA1) before

(red and black) and after applying the shuffling procedure to destroy the correlations (gray histograms; see Methods). We did notice a difference in the overall distribution of pairwise correlation values in the two different areas. In particular, CA1 distributions seem to show a stronger peak for negative values and tend to have a smaller frequency of values around 0. The distributions didn't qualitatively change after destroying the correlations through shuffling in either region. **c** To investigate correlations with finer detail, we studied their marginal distributions within single locations in the arena (histograms for all DG and CA1 cells combined across animals). Insets show a magnification of the y-axis. The peak of the distribution of the correlations across pairs of neurons was negative and small in magnitude since in the vast majority of passes only a few cells are active at any point in time. However, there are several pairs of neurons with a relatively strong correlation, as shown in the insets. As expected, the peak of the distribution was concentrated within a short interval of small negative values since in the vast majority of passes only a few cells would be active at any point in time. We then focused our attention at the positive side of the distribution (inset), since it is the largest contribution to correlations (in absolute terms). In line with our main results, we found higher positive correlations in CA1 than in DG (Mann-Whitney U, ***$p<0.001$). **d**) In both regions pairwise correlations are affected by our procedure. Here we show the distribution of the positive noise (in-location) correlations. In line with our main results, we found higher positive correlations in CA1 than in DG. Whisker plots show median, quart-percentile boxes, 5th and 95th percentile whiskers and Mann-Whitney U statistical tests for significantly different distributions (***$p<0.001$). Altogether, these results further confirm that although some correlations exist within DG, pairwise correlations only partially explain the effects on decoding that are caused by our shuffling procedure.

Supplemental Material S17: Importance and decoding performance for position; Related to Fig. 4

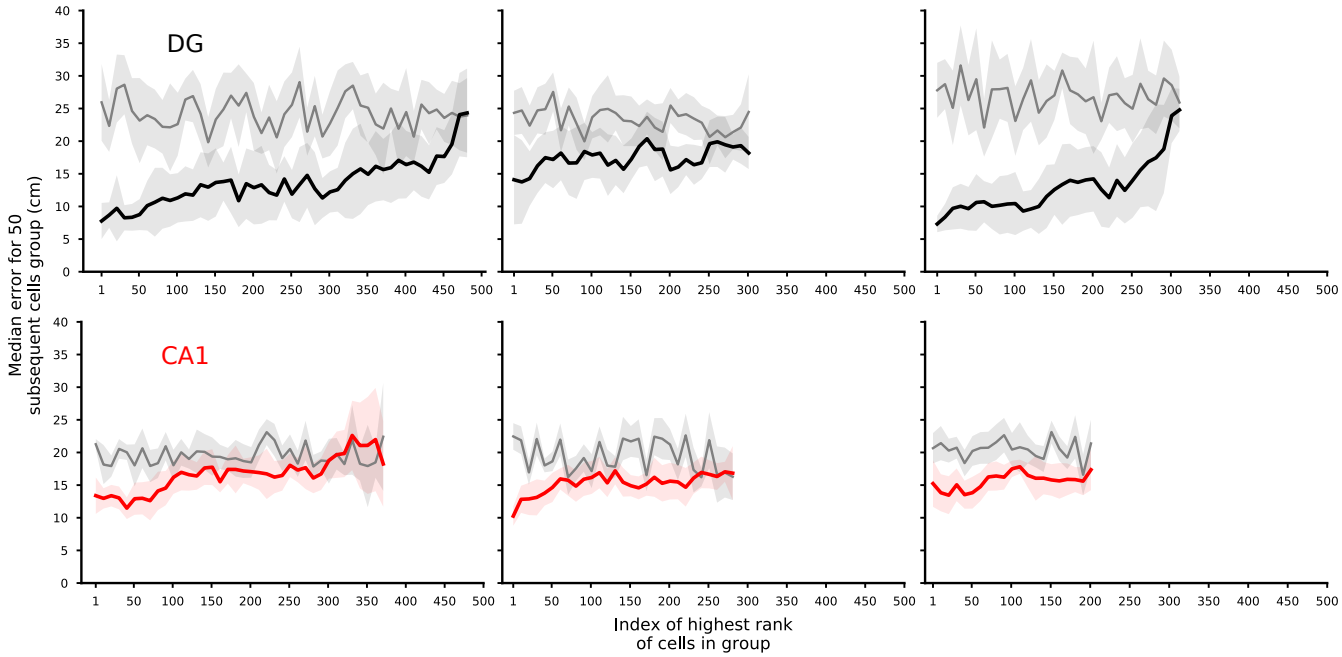

Figure S17: Importance and decoding performance for position. Decoding performance for cell ranking. Top: DG animals. Bottom: CA1. One animal per panel. See Fig. 4a in main text.

Supplemental Material S18: Importance and decoding performance for direction; Related to Fig. 6

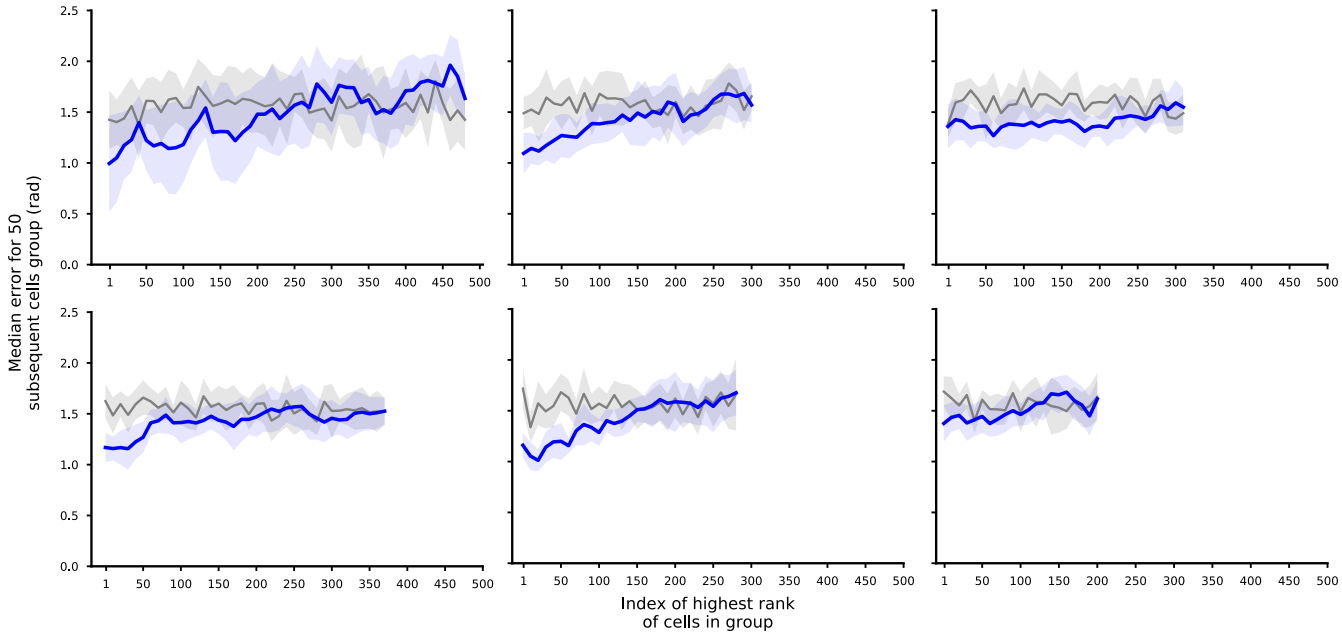

Figure S18: Importance and decoding performance for direction. Decoding direction performance for cell ranking. Top: DG animals. Bottom: CA1. One animal per panel. See Fig. 6a in main text.

Supplemental Material S19: Importance and spatial information; Related to Fig. 5

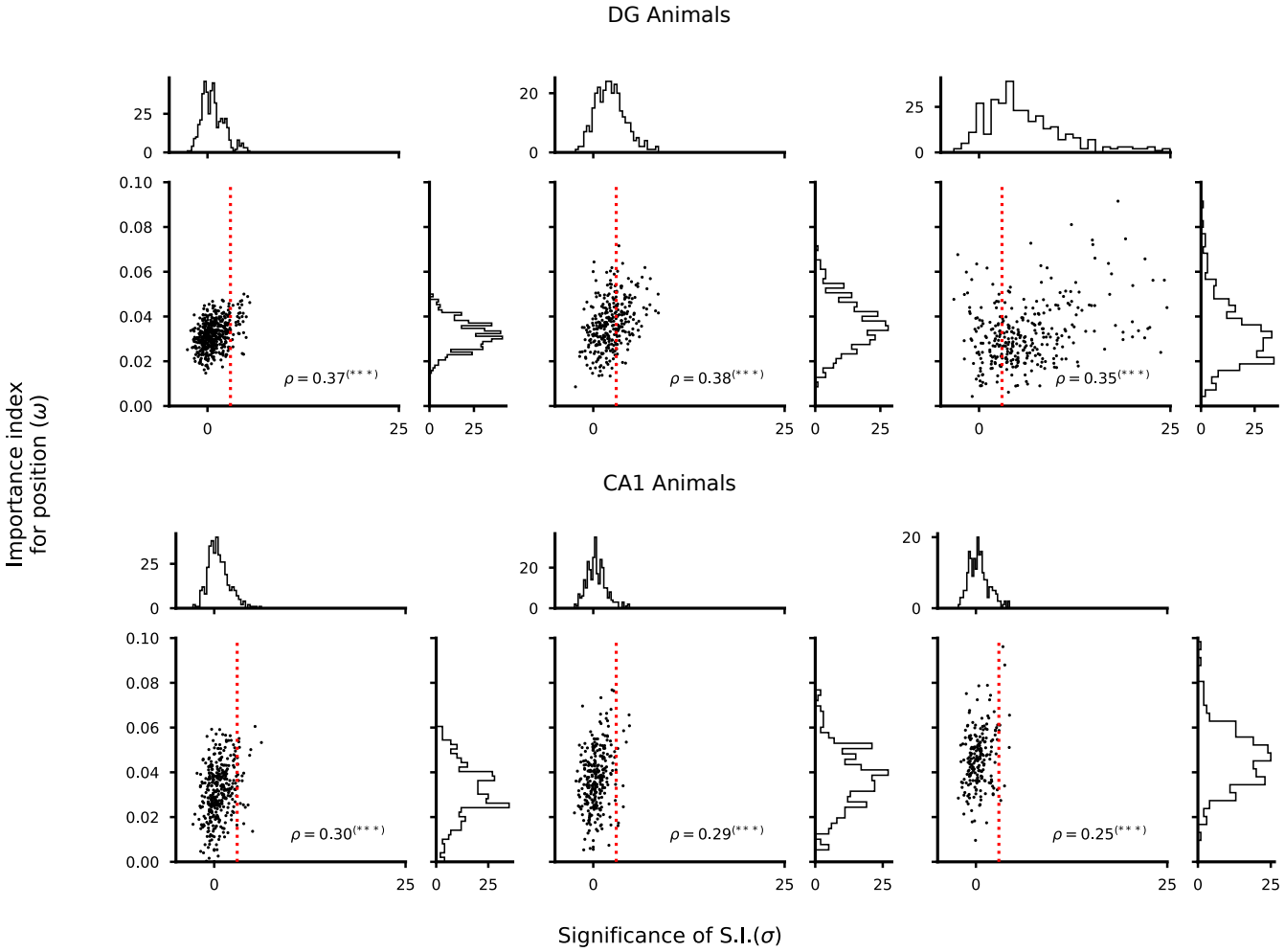

Figure S19: Importance and spatial information. Relation between importance index and significance of spatial information. Top: DG animals. Bottom: CA1. One animal per panel. See Fig. 5 in main text.

Supplemental Material S20: Importance for direction and direction information; Related to Fig. 6

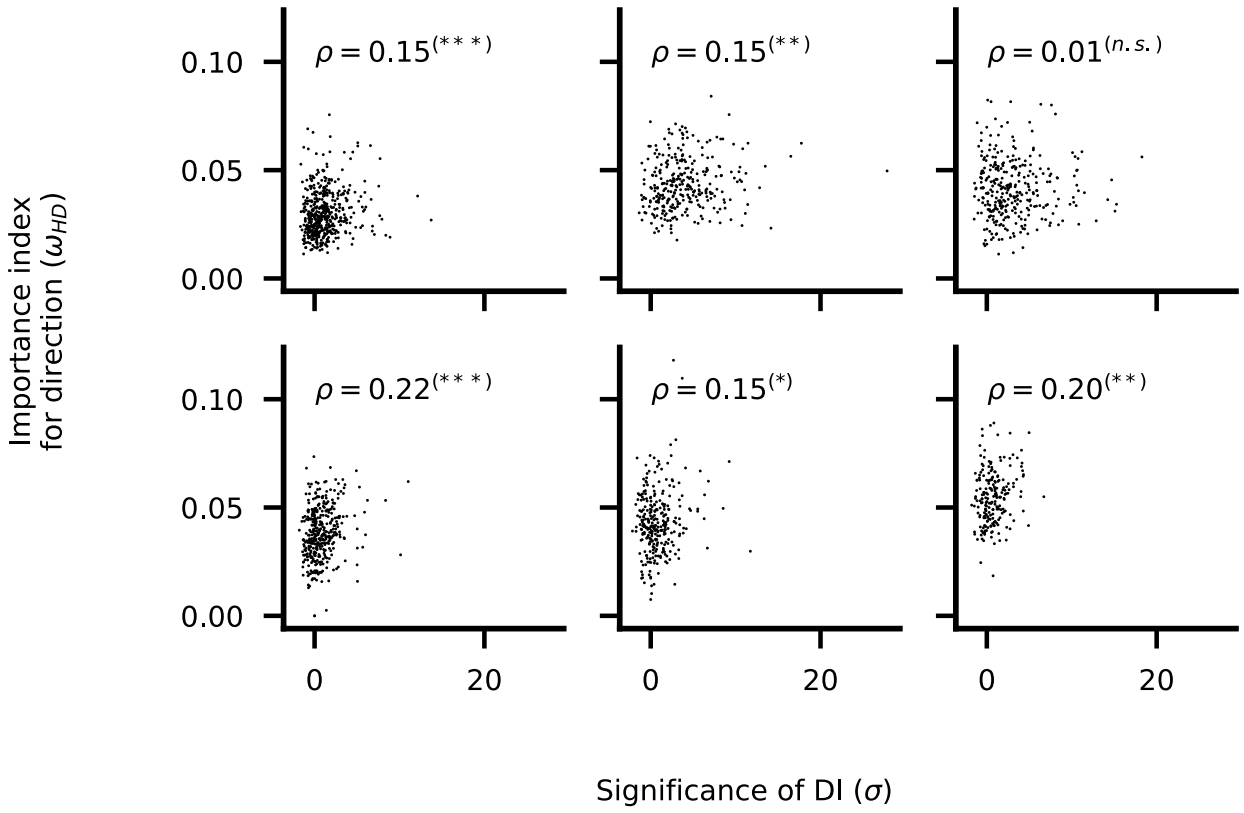

Figure S20: Importance for direction and direction information. Relation between importance index for direction of motion and significance of direction information. Top: DG animals. Bottom: CA1. One animal per panel. See Fig. 6 in main text.

Supplemental Material S21: Importance for position and importance for direction; Related to Fig. 7

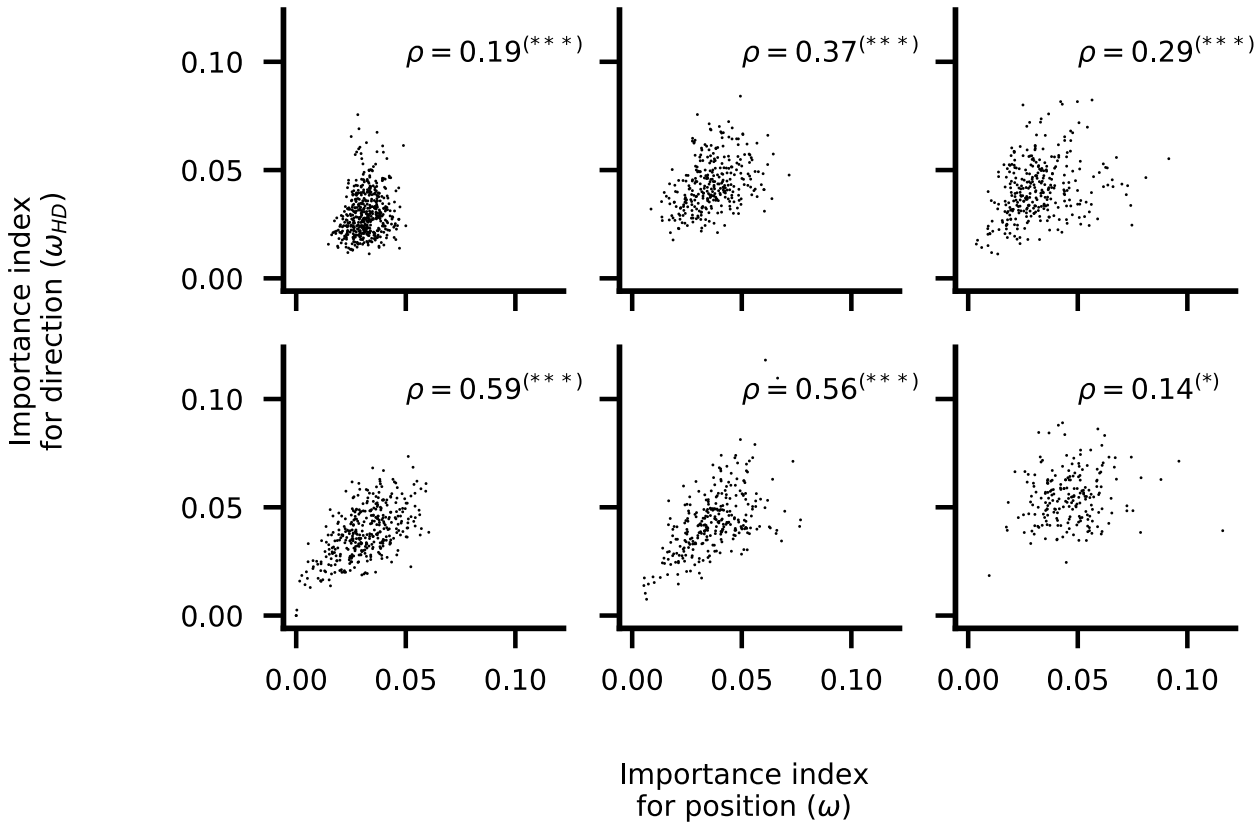

Figure S21: Importance for position and importance for direction. Relation between importance index for position and importance index for head direction. Top: DG animals. Bottom: CA1. One animal per panel. See Fig. 7 in main text.

Supplemental Material S22: Histology and lens placement

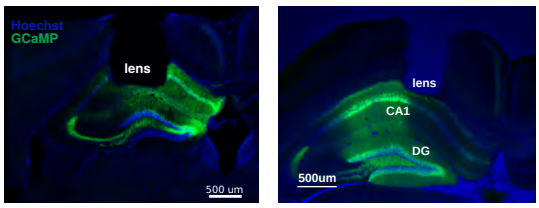

Figure S22: Histology and lens placement. Confocal image for DG (left) and CA1 (right). For all the mice presented in this report the histology confirmed the adequate placement of the lens.

Supplemental Material S23: Relation between correlations and importance index; Related to Fig.4,8

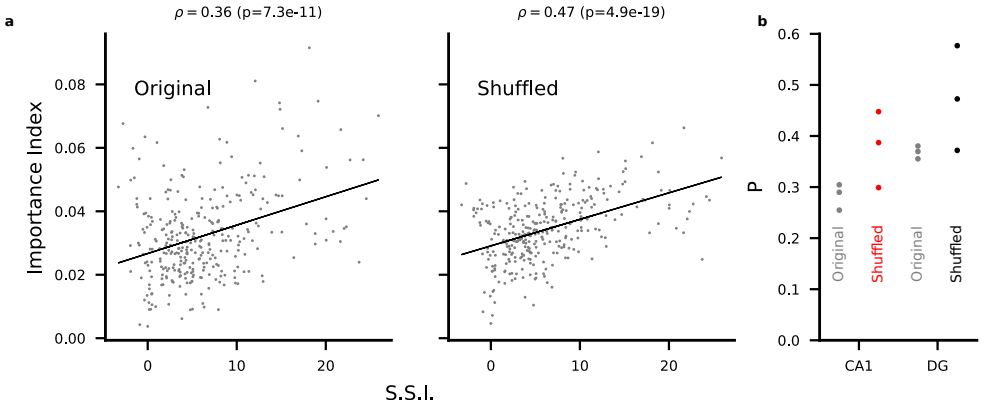

Figure S23: Relation between correlations and importance index. To study the relation between correlation and importance index, we looked at the importance index of single cells before and after destroying correlations among cells. In this way we could measure the independent contribution of single cells to the encoding of position since correlations were destroyed. On the one hand, after our shuffling procedure we would expect that the contribution of a cell to position encoding should be more tied to single cell properties such as spatial tuning. On the other hand, the shuffling procedure does not destroy the correlations that are induced by the relative tuning of cells to a given spatial location, often referred to as signal correlations, and therefore we would still expect some relation between importance index and spatial tuning. For instance, if multiple cells encode the same location, we would expect their importance index to be generally lower than that of cells tuned to location where no other cells are tuned to. Indeed, in our data we found that in most of our animals the correlation between importance index and significance of the spatial information becomes stronger after destroying correlations, suggesting that single cells properties become more relevant in the case of independent coding. However, the effect is marginal so that also in after destroying correlations both place and non-place cells contribute to accurate position decoding. Therefore, also after destroying correlations single cell properties don't fully explain importance index. Finally, we'd like to stress that in Fig. 4c we decoded position from non-place cells alone, defined as cells that do not pass a threshold for location-specific tuning. These results show that our findings are not an artifact of the choice of decoding position but are rather a manifestation of the distributed nature of the spatial code. **a**) Scatter plot of importance index in relation to the significance of spatial information (SSI) for all cells of a representative animal DG3. Left: original data. Right: same as in the left panel but with data in which correlations were destroyed through shuffling. P-values for Pearson-r correlations between the two quantities are reported on top of each panel. **b**) Values of Pearson-r correlations between importance index and SSI in original data (grey dots) and shuffled data (red, CA1; black, DG) for all the 6 mice analyzed. For the shuffled data, mean values of importance index across 20 shufflings was used.

Supplemental Material S24: Control for poorly segmented cells; Related to Fig.4,8

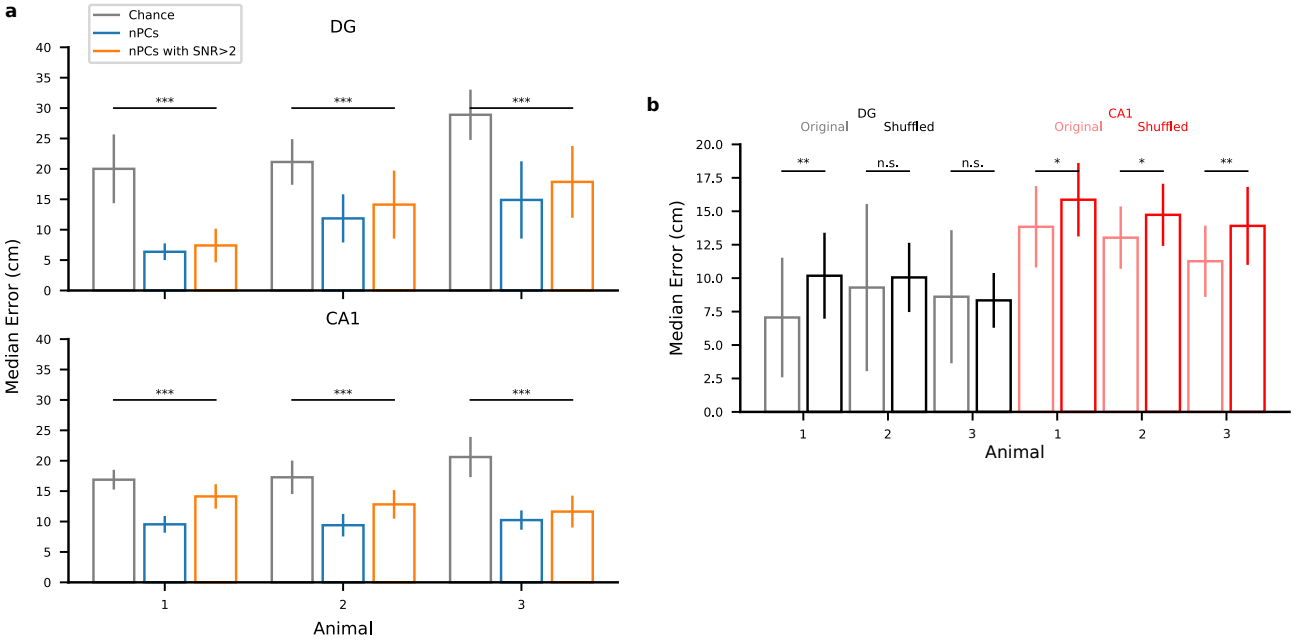

Figure S24: Control for poorly segmented cells. One may wonder if the main conclusions are due to the fact that most of the weakly spatially tuned cells may result from poor unit discrimination in the FOV. We rejected this hypothesis be removing the poorly discriminated cells and verifying that position could be decoded from non-place cells and that destroying correlation has still an effect in CA1.

To isolate the cells resulting from poor discrimination we used the signal-to-noise-ratio between the raw florescence and the deconvolved calcium trace:

$$SNR = \frac{||C||^2}{||C - C_{raw}||^2}$$

where C is the denoised calcium trace and C_{raw} is the raw calcium signal after background removal. This quantity has been introduced by the authors of the CNMF-e algorithm to assess the quality of the extracted signals (Zhou et al., 2018). We then looked at the relation between the place code and the quality of the segmentation. To verify that our main conclusions were not due to poorly identified cells classified as non-place cells, we introduced a criterion of a minimal signal to noise ratio of SNR=2. Our main results that position could be decoded from non-place cells and that destroying correlations had an effect on CA1 remain valid when we selected only neurons with a signal to noise ratio larger than 2. **a**) Position could be decoded from non-place cells also when the analysis is restricted to non-place cells whose SNR is higher than 2. Here we show the decoding position error for all DG (top) and CA1 (bottom) animals for non-place cells (blue) and also for non-place cells for which the signal-to-noise ratio (SNR) after source extraction from the video was higher than 2 (Mann-Whitney U test for significance, ***p<0.001). **b**) Decoding position error for original (light colors) and shuffled data (bright colors) after destroying correlations. Only cells with an SNR>2 were used in the decoding (28%, 91%, 98% in the 3 DG mice; 21%, 22%, 82% in the 3 CA1 mice). We therefore conclude that although untuned cells in our data are overall more likely to have a low SNR as result of poorer signal extraction than tuned cells, the main conclusions are not due to this type of artifacts in the imaging technique.

Table S1

	Authors	Notes	Mean rate
CA1	(Frank et al., 2001)	R, Ep, T-maze	0-5Hz
	(Lee et al., 2004)	R, Ep, enriched circular track	mean 1.1Hz, max 8.7Hz
	(Kelemen and Fenton, 2010)	R, Ep, shock av.	0-2Hz
	(Mizuseki et al., 2012)	R, Ep, multiple taks	(0.88±1.23)Hz
	(Bittner et al., 2017)	R, Ep, 1D	25Hz (peak in field)
	(Ziv et al., 2013)	M, Im, 1D	0.1-0.5 ev/s
	(Xia et al., 2017)	M, Im, two-chamber	∼0-0.2 ev/s
	(Allegra et al., 2019)	M, Im (2p), 1D, virtual,	0.1-0.05 ev/s
	Our work	M, Im, open field	0.06 ± 0.04 ev/s
DG	(Jung et al., 2093)	R, Ep, 8-arm	0-0.2Hz
	(Nitz and McNaughton, 2004)	R, Ep, open field,	15-30Hz (10-20Hz in CA1)
	(Leutgeb et al., 2007)	R, Ep, open field, lots of mossy	0-10 Hz (mean in-field) 5-25 Hz (peak)
	(Neunuebel and Knierim, 2012)	R, Ep, square	5-20 Hz (peak)
	(Pilz et al., 2016)	M, Im (2p), wheel,	0.01 ev/s
	(Danielson et al., 2016)	M, Im (2p), wheel, threshold-based inference	0-0.003 ev/s
	(Allegra et al., 2019)	M, Im (2p), 1D, virtual, higher selectivity in DG	0.1-0.05 ev/s
	Our work	M, Im, open field	0.15 ± 0.12 ev/s

Table S1: Firing rates in CA1 and DG. Related to Fig.4 Mean rates on CA1 and DG cells in spatial tasks (R=rats, M=mice, Ep=e-phys, Im=calcium imaging). One of the observations of our study is that the average rate of events in DG is larger than in CA1. Here we review the literature to show that our observation is compatible with the values that are reported, especially if one considers the high variability across different studies. Differences in behavioral protocols may strongly affect overall activity. This is especially true for a region like the hippocampus that is highly involved in a host of cognitive processes integrating information from virtually all the sensory areas and having a big role in memory. Most of the studies on the hippocampus in spatial tasks focus on 1-dimensional (1D) tracks (wheels or treadmills) where rats or mice are trained to collect rewards at the ends of a treadmill (e.g., experiments from the Buzsaki, Ziv, Schnitzer, Golshani, Dombeck, Losonczy labs among others) or specific mazes with several explicit cues (Knierim, Frank labs). These protocols are specifically designed to study particular properties of the activity of hippocampal cells and it may be inadequate to directly compare them to our study. Studies in freely moving rats include shock-zone avoidance (Kelemen and Fenton, 2010) and tasks combining the location of stimuli with other properties of the task such as identity and memory of previous trials (Wood et al., 1999). A relatively smaller number of studies has considered rats foraging in open fields (e.g., experiments from the Leutgeb, Moser, McNaughton labs) with a relatively small number of explicit or controlled cues. The reported quantities are also defined in various ways, for instance firing rates may be reported as mean rates across sessions (often of the order of 1 Hz) or as peaks within field (often of the order of 5-20Hz – in addition while fields are defined more or less in similar ways, cells with

significant fields are defined in inconsistent ways across studies). Furthermore, they may refer to different behavioural states (e.g., locomotion or immobility) or averaged across states. Thus, the firing properties of CA1 and dentate gyrus neurons can vary widely depending on the parameters of the experimental protocols. Differences in recording techniques are also important and they are probably one of the main sources of variability in estimates of neural activity. The rate of events reported in calcium imaging studies is always significantly smaller than the rate of spikes recorded using electrophysiological techniques. This is not surprising given that the techniques are very different (calcium is only a proxy of neuronal spiking) and often the neurons recorded using calcium imaging are of different types (see for example (Harvey et al., 2009; Dombeck et al., 2010; Ziv et al., 2013)). In electrophysiology, cells are isolated through clustering and selected using properties such as rates and shape of the action potential. A relatively large number of spikes is needed to isolate cells, introducing a bias toward most active neurons. Some labs include recordings during sleep to increase their statistical power but regions such as the dentate gyrus may express different activity profiles during sleep (Neunuebel and Knierim, 2012). In calcium imaging, cells are selected by viral expression and morphology. This allows to visualize the isolated cells, but spiking activity can only be inferred from the calcium signals. Through the relevant literature, we observe that electrophysiology studies typically report firing rates that are one order of magnitude higher than "calcium event rates" in imaging studies. Due to the relatively low sampling rate (5 Hz in our study, up to 30 Hz in other studies) and the slow temporal profile of the florescence profiles, spikes may go undetected in calcium imaging. Hence, it is highly likely that calcium events are in fact due to multiple spikes collected in a short time frame. Also, several spikes may simply be lost throughout the session due to low signal-to-noise. Lower event rates for calcium imaging with respect to electrophysiology are therefore expected. Important differences exist also among calcium imaging techniques, more specifically between one (or single) and two-photon imaging. One crucial step in the processing of calcium imaging videos is that of segmentation, i.e., the identification of the regions of interest (ROIs) in the field of view from which to compute the fluorescence signals and hence each putative cell's activity. Cell segmentation is very different in one and two-photon imaging. One-photon imaging relies on algorithms such as CNMF-e, which are sensitive to the statistics of cellular activities. Higher activity generally leads to improved source separation because of the resulting higher signal to noise ratio (SNR) in fluorescence signal, hence one-photon imaging may have a bias towards more active cells. Instead, in two-photon imaging cells are more easily identified, even in the case they are silent, and so manual segmentation or automated techniques only based on spatial profiles (e.g., apparent shape of the ROI) can be used. The bias of one-photon imaging is partially reduced by the fact that CNMF-e uses both spatial and temporal dynamics profiles to separate signal sources, hence typically even cells with low signal to-noise ratio and low activity are identified as separate sources, but a bias may still exist compared to two-photon imaging. The CNMF-e method has quickly become the *de facto* standard in the mini-endoscope imaging field, and now it is also been used in two-photon imaging studies together with other tools for automated segmentation (Pachitariu et al., 2017). Although CNMF-e data is now available for both techniques, it is still difficult to compare one and two-photon imaging because two-photon imaging is used in head-fixed preparations (e.g., mouse on a treadmill) rather than in freely-moving animals as one-photon imaging allows for. Still, it is possible that one-photon techniques underestimate the fraction of inactive cells, with a consequent inflation of the fraction of active cells and place cells (see also (Danielson et al., 2016)). An additional note should be made about methods of inference of cellular activity from fluorescence signals. This operation typically consists of many steps and has been rapidly evolving in the last few years thanks to the introduction of automated methods necessary for large scale imaging (from hundreds to tens of thousands of cells). While earlier studies applied thresholding to the fluorescence signal (typically measured as changes over an estimated baseline, $\Delta F/F$) to estimate cellular activity in time, more recent methods use more sophisticated approaches based

on models of calcium dynamics to get a more precise estimation of cellular activity (Dombeck et al., 2007; Mukamel et al., 2009; Pnevmatikakis et al., 2016; Pachitariu et al., 2017; Zhou et al., 2018). Thresholding may obviously lead to underestimating cellular activity since one calcium event may be caused by multiple spikes closely occurring in time like in bursts. Ziv and colleagues used a different approach whereby a positive trend in fluorescence signal was associated to spiking activity (Ziv et al., 2013). These different preprocessing choices constitute a further element of variability in the reported data, especially since calcium dynamics is not always known with high precision, depending on cell types and calcium indicators used, among other factors. More research will be needed to evaluate calcium imaging techniques against a ground truth of intra- or juxta-cellular recordings. This is a very active field of research and the data collected with these techniques may still reflect their methodological variability. With all these premises in mind, it is anyway useful to review the literature in which the authors report the activity of neurons in the dentate gyrus and in CA1. The firing properties of dentate gyrus granule cells are still relatively poorly understood. Most of the works that report sparsity arise from immediate gene expression studies and from rat studies by the Moser, Leutgeb, McNaughton and Knierim labs (among others). Notice that sparsity is often used with different meanings. Here we refer to the fraction of cells that are active, and not to the average activity of each neuron. Recent work from Goodsmith and Knierim (GoodSmith et al., 2017) has highlighted the difficulty in identifying granule cells using tetrode recordings. Based on cell morphology, histology and viral expression, we are highly confident that our signals correspond to the activity of granule cells of the dentate gyrus. We therefore believe that our study is the first to show dentate gyrus granule cells activity in freely moving foraging task in an open field. This makes any comparison with previous literature more difficult. Among the literature, we highlight a few recent works worth of note. A recent work by Allegra and colleagues (Allegra et al., 2019) finds very similar calcium activity profiles for CA1 and DG cells, though with a much lower spatial selectivity for CA1 cells. One of the most careful assessment of DG GCs activities (GoodSmith et al., 2017) finds rates for DG similar to CA1 with a 20-30% ratio of place cells in DG. Nitz and McNaughton (Nitz and McNaughton, 2004) report higher rates for DG cells while rats explored a novel rather than a familiar environment. This we report all the most relevant results from the literature. Overall, our observations are compatible with existing literature within the variability across studies. Notice that the differences between DG and CA1 in terms of overall activity are much smaller than the variation across studies. We therefore believe that the question of which area is more active in spatial tasks and what the ratios of place cells are in these regions remain open. We believe that imaging studies like ours targeting such comparison may strongly contribute to strengthening our understanding of the different computational properties of CA1 and DG (see for example (Allegra et al., 2019)). In conclusion, although differences between DG and CA1 place cell number may depend on several other factors such exposure times during habituation, our analysis suggests that the main drive for place cell counts and differences between these two areas comes from the different activity levels. In particular, a non-negligible proportion of CA1 cells are more sparsely active than previously reported and therefore CA1 place cells may have been over represented in place cells studies. We encourage researchers to adopt a more systematic assessment of the significance of the spatial information as we suggest in our work and to always report at least the proportion of place and non-place cells that are found.