Task-dependent mixed selectivity in the subiculum

Graphical abstract

Highlights

- CA1 and subiculum neurons respond conjunctively to position, head direction, and speed

- The degree of conjunctive coding ("mixed selectivity") is stronger in the subiculum

- Mixed selectivity is stronger during goal-directed navigation than in free foraging

- Decoding of each navigational covariate is more accurate with mixed selectivity
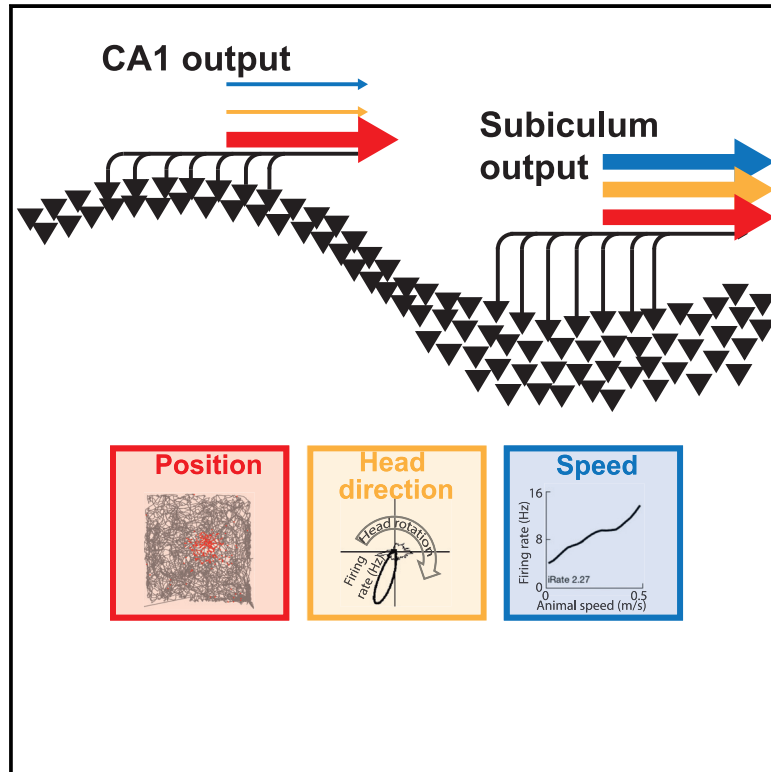
Authors

Debora Ledergerber, Claudia Battistin, Jan Sigurd Blackstad, ..., May-Britt Moser, Yasser Roudi, Edvard I. Moser

Correspondence

debora.lederberber@kliniklengg.ch (D.L.),
yasser.roudi@ntnu.no (Y.R.),
edvard.moser@ntnu.no (E.I.M.)

In brief

Ledergerber et al. show that pyramidal cells in the subiculum combine navigational variables—position, head direction, and speed—more extensively than their counterparts in hippocampal CA1 so that such mixed selectivity is stronger during goal-directed navigation than free foraging and that navigational information is decoded more efficiently from a mixed-selective code.

CellPress

Article

Task-dependent mixed selectivity in the subiculum

Debora Ledergerber,[1,2,*] Claudia Battistin,[1] Jan Sigurd Blackstad,[1] Richard J. Gardner,[1] Menno P. Witter,[1] May-Britt Moser,[1] Yasser Roudi,[1,*] and Edvard I. Moser[1,3,*]

[1]Kavli Institute for Systems Neuroscience and Centre for Neural Computation, Norwegian University of Science and Technology, Olav Kyrre s gate 9, MTFS, 7489 Trondheim, Norway
[2]Present address: Swiss Epilepsy Center, Klinik Lengg AG, Bleulerstrasse 60, 8008 Zurich, Switzerland
[3]Lead contact
*Correspondence: debora.lederberber@kliniklengg.ch (D.L.), yasser.roudi@ntnu.no (Y.R.), edvard.moser@ntnu.no (E.I.M.)
https://doi.org/10.1016/j.celrep.2021.109175

SUMMARY

CA1 and subiculum (SUB) connect the hippocampus to numerous output regions. Cells in both areas have place-specific firing fields, although they are more dispersed in SUB. Weak responses to head direction and running speed have been reported in both regions. However, how such information is encoded in CA1 and SUB and the resulting impact on downstream targets are poorly understood. Here, we estimate the tuning of simultaneously recorded CA1 and SUB cells to position, head direction, and speed. Individual neurons respond conjunctively to these covariates in both regions, but the degree of mixed representation is stronger in SUB, and more so during goal-directed spatial navigation than free foraging. Each navigational variable could be decoded with higher precision, from a similar number of neurons, in SUB than CA1. The findings point to a possible contribution of mixed-selective coding in SUB to efficient transmission of hippocampal representations to widespread brain regions.

INTRODUCTION

The hippocampus has a well-established role in mnemonic and navigational functions of the brain (Hasselmo, 2012; Morris, 2007; O'Keefe and Nadel, 1978; Scoville and Milner, 1957; Squire, 1992). Its interplay with other cortical regions is thought to be indispensable for the formation and retrieval of episodic and positional memories (Buzsáki, 1989; McClelland et al., 1995; Squire, 1992; Squire et al., 2015; Winocur and Moscovitch, 2011). The subiculum (SUB) has an important anatomical position as an interface between the hippocampus and other brain areas (Cappaert et al., 2015; O'Mara, 2006). It is the major long-range projection area of the hippocampus, the origin of a substantial part of the fornix, and the source of large parts of non-fornical output reaching a range of cortical and subcortical downstream areas. Along with the CA1 area, it sends outputs to infralimbic, prefrontal, orbitofrontal, and medial and lateral entorhinal cortices (MECs and LECs) and to subcortical structures, including the septal complex, the mammillary nucleus, the hypothalamus, the thalamus, and the amygdala (Cappaert et al., 2015; Cembrowski et al., 2018a; Ishizuka, 2001; O'Mara et al., 2001). However, even though CA1 and SUB have many overlapping target areas, their patterns of connectivity are very different; the majority of CA1 neurons each send collateral projections to at least two targets elsewhere in the brain, whereas in the case of the SUB, such branching is observed only in a minority of neurons (Bienkowski et al., 2018; Cembrowski et al., 2018b; Naber and Witter, 1998; Wit-

ter, 2006). This difference raises the possibility that the hippocampus uses CA1 and SUB outputs differentially to distribute information to downstream brain regions.

Although neuronal representations in CA1 have been known for half a century to show strong spatial selectivity, as expressed in place cells (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978), data from SUB are scarce and it has remained elusive what SUB adds to the hippocampal computation. Subpopulations of SUB neurons have broad spatial firing fields (Sharp, 1997, 2006; Sharp and Green, 1994), of which some, known as boundary-vector cells, are oriented in parallel to elongated geometric boundaries (Lever et al., 2009; Stewart et al., 2013). Vector trace cells fire at given distances and directions from discrete objects or boundaries, with firing fields that may outlast the presence of the object or boundary (Poulter et al., 2020). Yet, other SUB cells encode the animal's axis of movement (heading direction) when animals navigate on elevated multidirectional tracks (Olson et al., 2017). In all tasks in which SUB neurons have been recorded, they appear to be more broadly tuned to features of behavior or environment than neurons in other regions of the hippocampal formation. SUB is likely to receive navigational input from narrowly tuned place cells in CA1 (O'Keefe, 1976; O'Keefe and Dostrovsky, 1971); from grid cells in the MEC and pre- and parasubiculum (Boccara et al., 2010; Fyhn et al., 2004; Hafting et al., 2005); from head direction cells (Taube and Burton, 1995; Taube et al., 1990), border cells, and object vector cells (Høydal et al., 2019; Solstad et al., 2008) in the same regions; and from speed

cells in the hippocampus and MEC (Kropff et al., 2015). The specificity of these putative inputs brings up the question of what the broader representations in SUB add to the output of the hippocampal formation.

Navigation-tuned cells in the hippocampal formation express similar types of information in multiple environments. This is in contrast to neuronal representations in many other brain areas, where spatial selectivity is apparent only under task conditions relevant to the brain region, such as in prefrontal cortex (Jung et al., 1998; Padilla-Coreano et al., 2019; Pratt and Mizumori, 2001), posterior parietal cortex (Nitz, 2006; Whitlock et al., 2012), primary visual cortex (Goltstein et al., 2018; Saleem et al., 2018), amygdala (Peck et al., 2014), and nucleus accumbens (Lansink et al., 2012; Mulder et al., 2005). Navigational information in these regions may be derived from representations in the hippocampus (Remondes and Wilson, 2013; Spellman et al., 2015), but it remains unclear whether, and how, outputs from hippocampus would be modified in a task-specific manner before reaching these diverse regions. Given the potential role of SUB in distributing hippocampal output to widespread regions of the brain (Cappaert et al., 2015; Gigg, 2006; O'Mara, 2006), we hypothesized that, rather than generating *de novo* representations, SUB modifies representations from upstream neural populations in CA1, presubiculum, and entorhinal cortex to facilitate decoding by downstream regions during hippocampal-dependent behaviors.

It has been suggested that networks consisting of neurons that encode multiple stimulus features simultaneously or conjunctively, using a "mixed selectivity" (MS) code, have several computational advantages over networks in which neurons respond predominantly to single features (Miller et al., 1996; Rigotti et al., 2013). Besides a high representational capacity, networks with high-dimensional coding in individual neurons have the advantage that a wide span of task-relevant aspects is accessible to linear classifiers, as the number of classifications that can be performed by a linear readout grows exponentially with the dimensionality of the information carried by the neurons (Fusi et al., 2016). Increasing the level of MS in SUB might therefore be a mechanism by which the hippocampal formation makes relevant output more accessible to downstream target regions.

With these advantages of MS in mind, we asked if SUB modifies output from the hippocampus by combining, in individual neurons, multiple features of the navigation experience in ways that depend on current task goals. Considering that much of the output from CA1 is also passed on to SUB, we performed simultaneous *in vivo* electrophysiological recordings in these regions in rats performing either random foraging or a spatial navigation task. We compared representations in SUB and CA1 for three navigational covariates, namely, position (P), head direction (H), and speed (S). We report that individual neurons in SUB combine these behavioral covariates more extensively than their counterparts in CA1 and more strongly during goal-directed spatial navigation than during free foraging. This coding scheme was paralleled by more accurate decoding of the navigational covariates from SUB than from CA1, providing regions downstream of SUB with broad spectra of information even from limited numbers of SUB output cells.

RESULTS

Anatomical location of recording electrodes

In order to understand what the SUB adds to the navigational output of the hippocampus, we performed extracellular recordings in CA1 and SUB (Figure 1A). All tetrodes were placed in the dorsal one-third of each region (see sectioning plane in Figure 1A). In Nissl-stained coronal sections, CA1 was identified as the narrow, densely packed layer of small pyramidal cells that extends from CA2 (with a less compact cell layer and larger neurons) to SUB (defined as the thicker, more diffuse layer of medium-sized neurons located dorsomedially to CA1). Moving away from the septal pole, SUB gradually widens, extending all the way medially until an additional granule layer is added ventrally, almost at the midsagittal side of the hemisphere. This additional layer belongs to the retrosplenial cortex (RSC; area 29) in the septal part of the hippocampal formation (RSC in Figure S1) and to the dorsal presubiculum in more temporal parts of the structure (preSUB in Figures 1B and S1A). At the septal pole, CA1 continues medially into the fasciola cinereum (FC in Figures 1B and S1; Boccara et al., 2015). Recordings from FC were not included in our study. Recordings from nearby SUB were included if the tetrodes were more than 50 um away from FC (SUB* in Figure 1B and Figure S1; see "Histology and reconstruction of tetrode placement" in STAR Methods).

We isolated 760 putative principal neurons from recordings in 8 rats (see STAR Methods; Table S1). Only cells with clear cluster separation from other background and other neurons were accepted (Figures S2A–S2C). Among the accepted neurons, 421 were located in SUB and 325 in CA1 (346 neurons in 3 rats, of which 96 in CA1 and 249 in SUB, were recorded in sessions in which CA1 and SUB were recorded in parallel). In SUB, the recording electrodes were distributed quite evenly along the proximo-distal axis (from CA1 to FC, RSC, or preSUB), whereas in CA1, there was a bias for the electrodes to be positioned in mid to distal parts of the subfield, i.e., nearer the SUB boundary (only 2 out of 18 tetrodes were located near the proximal end of CA1, near CA2) (tetrodes 1 and 11 in animal 24101 in Figure S1). However, the number of neurons recorded was distributed more evenly along the proximo-distal axis (approximate neuron numbers in the respective regions: 35 in proximal CA1, 223 in mid CA1, 49 in distal CA1, 292 in proximal , and 136 in distal SUB; arbitrary boundaries dividing subfields in three equal bands; Table S1).

Position coding in CA1 and SUB

With the aim of investigating the neurophysiological properties of SUB during active foraging and navigation, rats were trained in 2 tasks. First, 5 of the 8 rats were trained in an open field (OF) foraging task in which rats searched for randomly scattered cookie crumbs in a dimly-lit 1.5-m-wide square enclosure on a floor consisting of a black even mat without incisions. A cue card was attached on one of the walls. Second, all 8 rats were trained in a spatial task (ST) adapted from Pfeiffer and Foster (2013) in which the rats alternated between free search and goal-directed memory-based navigation. The same enclosure was used as for the OF but the floor mat was removed,
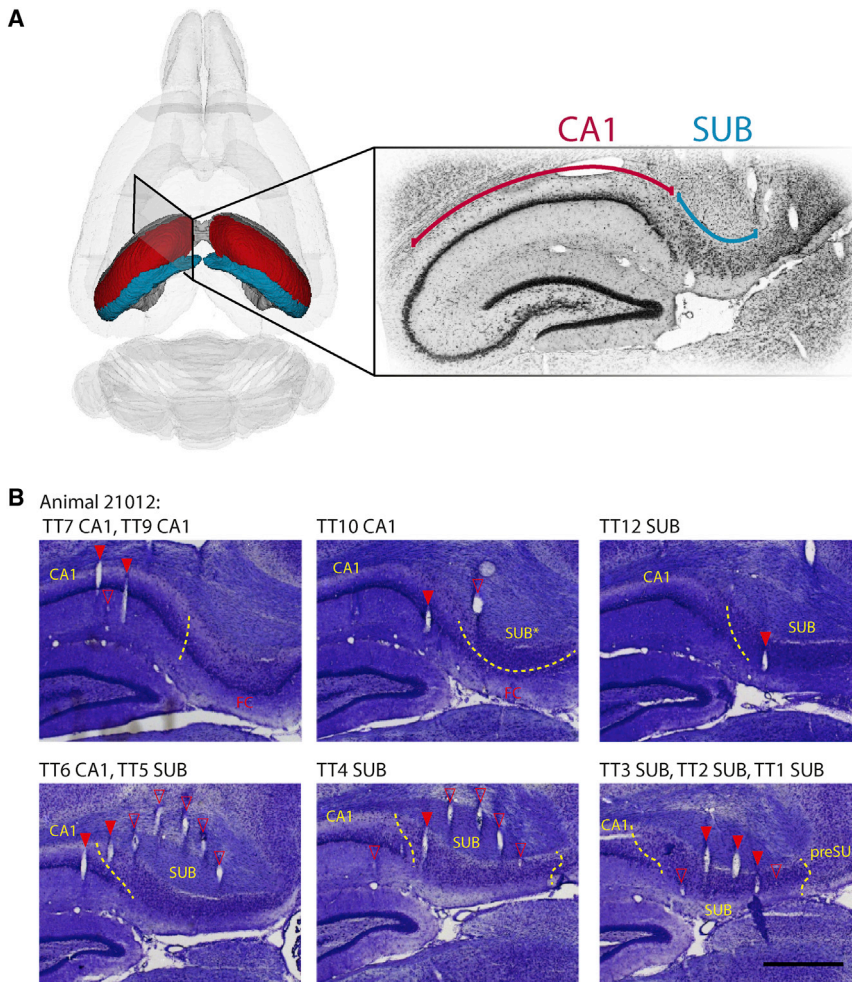
A

B

uncovering a square grid of 1-cm-diameter holes in the floor. The walls of the arena and their location in the room with respect to all distal landmarks remained constant between OF and ST (Figures S2D and S2E). During the ST, chocolate oat milk was provided alternately at a fixed "home well" and in a randomly selected well ("random well") of the arena (Pfeiffer and Foster, 2013). The animals memorized the home well location and learned to navigate back to it straight after consuming the randomly placed reward (Figure S2D), as reflected in their behavioral latencies (Figure S2F). Training in OF and SF was conducted on alternating sessions.

We first compared the spatial tuning properties of CA1 and SUB neurons. These analyses were performed on combined data from ST and OF because no major differences in spatial tuning were detected in separate analyses (Figures S3A–S3D). In agreement with previous studies (Kim et al., 2012; Sharp and Green, 1994), the average firing rate of principal neurons was significantly higher in SUB (4.44 ± 0.14 Hz) than in CA1 (1.7 ± 0.12 Hz, p = 4.6e-53, Welch's test; Figures 2A and S3A). Similarly, as in previous studies (Sharp, 1997, 2006; Sharp and Green, 1994), CA1 cells had sharply defined firing fields (Figure 2A; 95% of spikes in CA1 neurons fell into 25.7% ± 1.1%

of the spatial bins), whereas SUB neurons tended to fire continuously across the environment (Figure 2A; 95% of the neuron spikes fell into 56.3% ± 0.7% of the spatial bins), often expressing some degree of boundary vector-like bias in activity, which is in line with earlier findings (Lever et al., 2009; Stewart et al., 2013) (Figures 2A, S3A, and S5B). The information rate of the neurons, or their amount of position information per time interval (Skaggs et al., 1993), was similar in CA1 and SUB (Figures 2B and S5C; medians and median absolute deviation [MAD] are 0.47 ± 0.27 bits/s in CA1 and 0.42 ± 0.26 bits/s in SUB; p = 0.02, Mann-Whitney U-test), whereas information content, or the information a neuron's spikes provide about P (Skaggs et al., 1993), was markedly smaller in SUB (Figure 2C and S5D; median and MAD for CA1: 0.84 ± 0.57 bits/spike, for SUB: 0.11 ± 0.16 bits/spike; p = 1.1e−70, Mann-Whitney U-test).

The difference in information content but not rate between CA1 and SUB raises the question of what benefits are conferred by increasing the mean firing rate and therewith sacrificing information per spike during the processing step from CA1 to SUB. One possibility is that SUB, by integrating positional information over longer timescales, might combine input from CA1 with information from other sources (including inputs encoding variables besides P). With this possibility in mind, we investigated whether information about other behavioral variables is expressed in spike trains of SUB neurons.
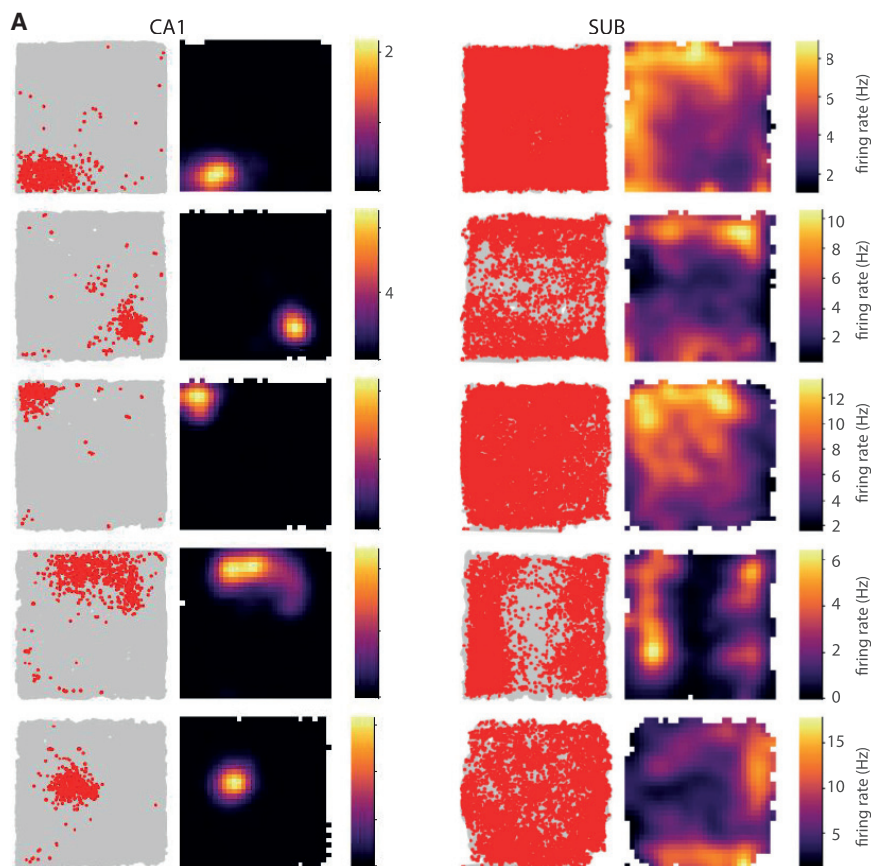
Figure 2. Difference in position coding in CA1 and SUB

(A) Example path plots (left) and rate maps (right) from neurons in CA1 (left column) and SUB. Path of the animal during the entire session is shown in gray with emitted spikes overlaid in red. Color bars to the right indicate firing rate.

(B and C) Frequency distributions showing P information rate (B) and information content (C) for all neurons recorded in CA1 (red) and SUB (blue) in all animals. Dashed line, median of each distribution.
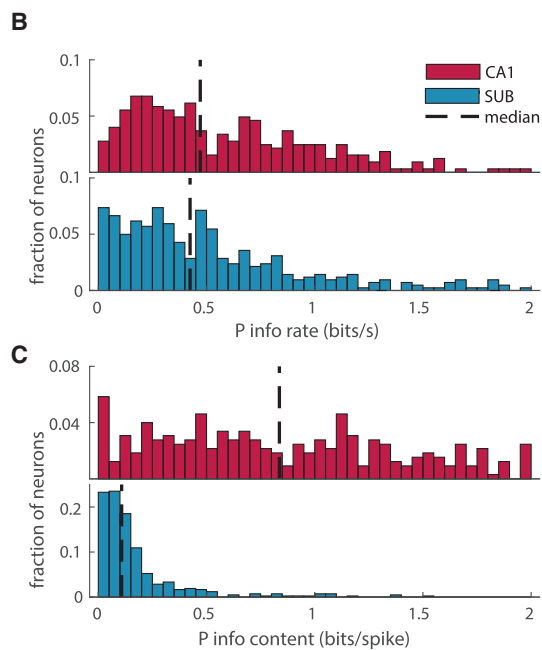
Representation of multiple covariates including H and S in SUB and CA1

Previous studies found that subsets of CA1 and SUB neurons show some degree of modulation by navigational variables like head direction (H) modulation and running speed (S). In CA1, place cells may be tuned to H inside their place fields (Acharya et al., 2016; Langston et al., 2010; Leutgeb et al., 2000). Weak H tuning has also been reported for spatially modulated neurons in SUB (Sharp and Green, 1994). Similarly, place cells in CA1 respond to some degree to running speed, whereas speed tuning has not been reported in SUB to our knowledge (Czurkó et al., 1999; Kropff et al., 2015; McNaughton et al., 1983).

Based on these observations, we set out to compare quantitatively, in the same recordings, the extent to which cells in CA1 and SUB represent navigational covariates besides P, like H and S. The H tuning curves of the recorded neurons revealed various degrees of modulation by H, in both CA1 (Figure 3A, 2nd column) and SUB (Figure 3B, 2nd column). Consistent with previous reports in CA1 (McNaughton et al., 1983; Skaggs et al., 1993), the information rate in CA1 neurons was lower for H than for P (Figures 3C and 3D; median and MAD of the CA1 population was 0.47 ± 0.27 bits for P and 0.04 ± 0.04 bits for H; $p = 6.87e-67$ Mann-Whitney U-test). In contrast, we found that in SUB neurons, information rates for H and P were in the same order of magnitude, albeit they were significantly different (median and MAD for H: 0.22 ± 0.15 bits/s and for P: 0.42 ± 0.26 bits/s; $p = 4.4e-13$ Mann-Whitney U-test). CA1 and SUB neurons showed similar levels of H information content (medians and MAD 0.07 ± 0.06 bits/spike for CA1 and 0.06 ± 0.04 bits/spike for SUB; $p = 0.98$ Mann-Whitney U-test).

A similar pattern could be seen for speed information rates and contents (Figure 3E and 3F). CA1 and SUB neurons had similar information content for speed (0.07 ± 0.07 bits/spike in CA1 versus 0.04 ± 0.04 bits/spike in SUB), whereas information rates for speed were higher in SUB than CA1 neurons (0.04 ± 0.04 bits/s in CA1 and 0.10 ± 0.10 bits/s in SUB; $p = 6.8e-9$ Mann-Whitney U-test).

Modeling spike rate with a GLM reveals MS in SUB

Measuring information rate separately for each different covariate (P, H, and S) entails some important limitations. These three behavioral covariates are strongly interdependent, meaning that an effect that is selectively tuned to one covariate might manifest spuriously when measured in relation to a different covariate (Acharya et al., 2016). This undesirable effect may result in both false-positive and false-negative tuning outcomes. One solution to this challenge is to use multiple-variable statistical methods, such as the generalized linear model (GLM). The GLM considers multiple covariates simultaneously and thus allows the influences on a cell's activity to be ascribed in a principled manner to the covariates that provide the strongest prediction.

Therefore, we extended the analysis by using a Poisson GLM framework (Hardcastle et al., 2017; McCullagh and Nelder, 1989) to investigate tuning of the neurons to P, H, and S, while accounting for any deceptive correlations between the covariates induced by sampling (ST and OF still combined; Figures S3E–S3G). Alongside the behavioral covariates (P, H, and S), we

also included two covariates representing basic influences on neural activity, namely, ensemble activity (E), defined as the Z-scored spike count summed over all other neurons on the same tetrode as the cell in question; and the theta phase (T) of the filtered local field potential (see STAR Methods). These two covariates were treated as "internal covariates," as their influence is exerted only in the local neuronal network and they are themselves modulated by the animal's behavior. They were important, however, for accurately fitting the model to the spike counts, as they remove a possible bias due to correlation between P, H, and S and the internal covariates. For analyzing the functional properties of the two brain regions, we focused our attention on the behavioral covariates P, H, and S (see STAR Methods). All the covariates (P, H, S, E, and T) were binned along their respective dimension (30×30 bins for P; and 10 bins for each H, S, E, and T), and continuity in the predicted tuning curves was enforced by a smoothness prior (STAR Methods). All models resulting from specific subsets of the covariates (ranging from single-covariate models, e.g., only P, to the most complex model containing all five covariates [PHSET]) were trained and tested by 10-fold cross-validation for every cell. Bin size and smoothness were treated as hyperparameters and selected by optimizing the cross-validated log-likelihood (LLH) for each single covariate model (Figure S4). For each neuron, the full-model GLM (PHSET) yielded tuning curves for P, H, and S (Figures 4A–4C), which qualitatively reproduced the tuning curve for the data (Figures 4A–4C).

Model selection was performed in a forward stepwise fashion, starting from single-covariate models and adding one covariate at a time by using a non-parametric test and a 5% significance level (see STAR Methods). Except in Figure S3G, models containing internal covariates E and/or T were pooled with corresponding models without this covariate (e.g., neurons best modeled by the PHT, PHE, and PHTE were counted to the group of the PH model). For every neuron, we determined which model provided the best fit to the neuron's firing properties in the data (Figure 4D), which in turn allowed us to assess model performance at the population level (Figure 4E). The proportion of neurons for which the most complex model (PHS) performed significantly better than any simpler model was higher for SUB than CA1 (53.7% in SUB versus 46.5% in CA1, percentages of each region; Figure 4E; difference, >99.9 percentile of a distribution of shuffled data in which cells were randomly assigned to anatomical region). Conversely, the proportion of neurons from one region best fitted by the PS model was significantly lower in SUB than in CA1 (25.9% in CA1 and 8.8% in SUB; difference, >99.9 percentile of a shuffled distribution; Figure 4E). These findings may suggest that both CA1 and SUB cells express high levels of MS, but the degree of mixing is stronger in SUB. To confirm that our GLM framework correctly picked up neurons with combined tuning properties, we simulated synthetic neurons by using inhomogeneous Poisson spike trains that were either purely tuned to P, H, or S or expressed known mixed combinations of the three. Classifying these synthetic cells with our GLM showed that the model selection procedure is not prone to overfitting by including spurious covariates (Figure S4F).

One caveat of the model selection analysis is that it only allows us to conclude that a number of CA1 and SUB neurons were
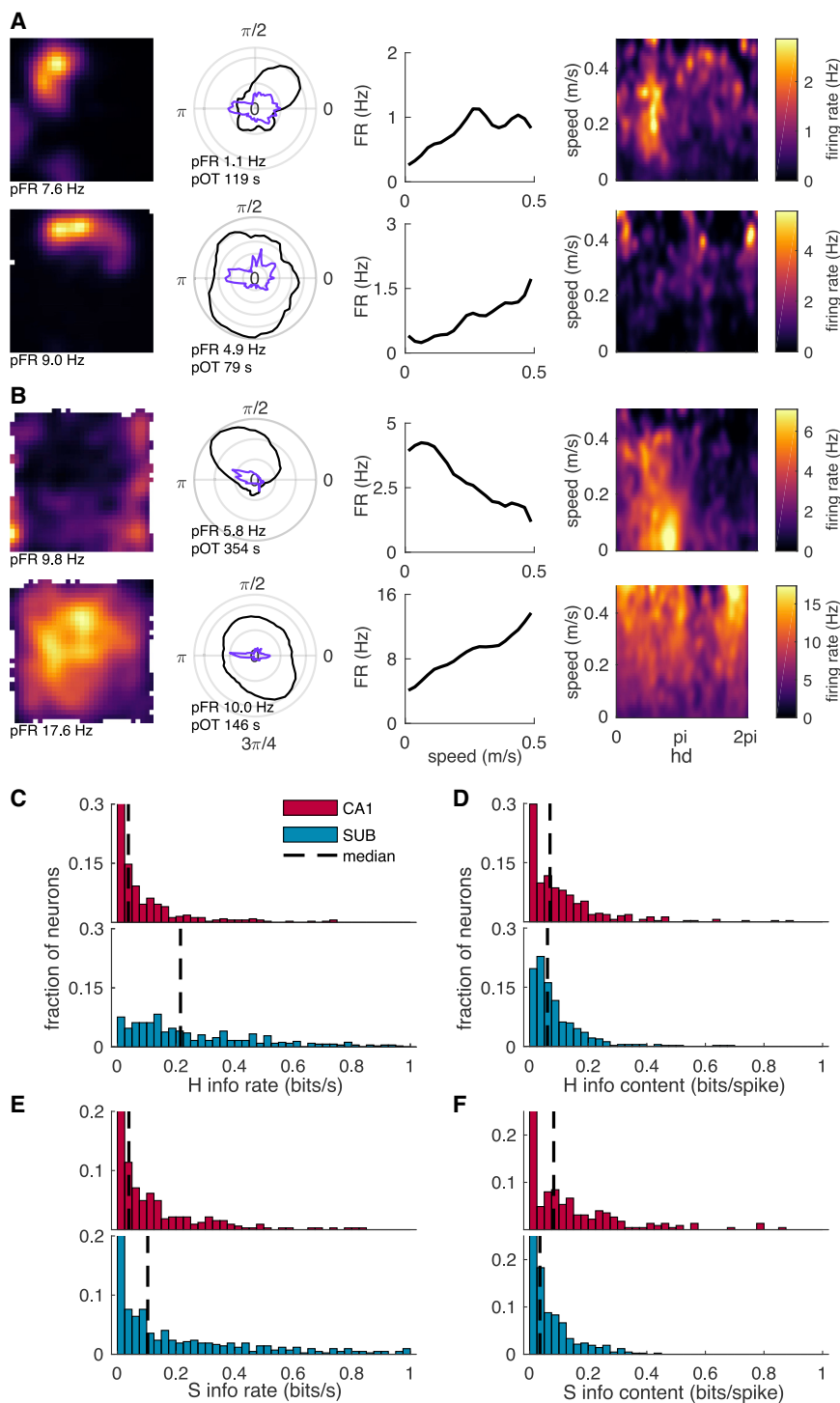
A

pFR 7.6 Hz

pFR 9.0 Hz

B

pFR 9.8 Hz

pFR 17.6 Hz

C

D

E

F

(rSCC) by taking the difference between the LLH of the selected model and the LLH of the model with the respective covariate removed. This term was then divided by the square root of the sum of the squares of all differences. The contributions of covariates not used in the selected model were set to zero (see equation in STAR Methods section "Relative single covariate contributions"). rSCC was in many cases correlated with information rate for the respective covariate (Figure 4F), although occasionally there were large differences. Although the measure of information rate is biased due to correlations between covariates, rSCC is derived from the GLM framework, allowing us to simultaneously quantify the contribution of multiple covariates. This leads to a more differentiated estimate of the contribution of individual covariates to the cell's firing rate (Figure 4F, see rSCC versus information rate in the two SUB neurons).

At the population level, plotting the rSCC for CA1 and SUB neurons separately (Figure 4G) revealed that CA1 neurons were largely distributed along the axis for P, showing very little contribution of H and S to their firing pattern, despite the fairly large number of cells that selected the PHS model in the above analyses. In contrast, SUB neurons were scattered throughout the entire space, indicating the differential combinations of the three covariates determining the firing of the neurons.

modulated by P, H, and S in a statistically significant manner. It did not inform us about how strong this modulation was. As a next step, we therefore investigated how strongly each neuron's firing rate was determined by either an individual covariate (P, H, or S) or a combination of them. Using the GLM framework, we quantified the relative contribution of every single covariate
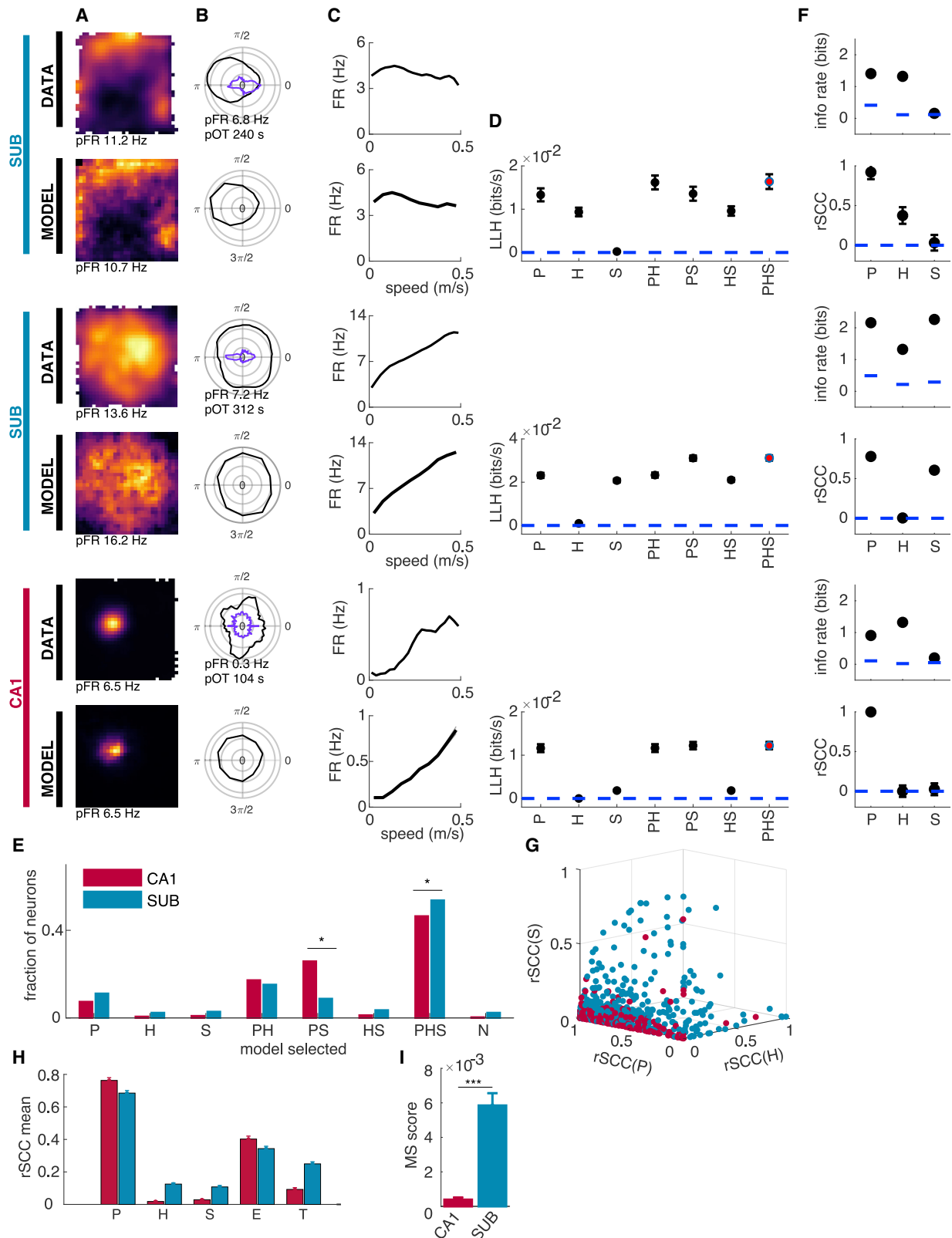
Similarly, wide scatter was seen in both proximal and distal SUB (Figure S5G). These observations suggest that, despite showing a high level of MS in terms of model selection, CA1 neurons singled out one covariate—P—that dominated their tuning properties, whereas SUB neurons more strongly combined multiple covariates, expressing a conjunctive code. The internal covariates E and T were an important factor for determining the firing rates of neurons in both regions. E was represented similarly in both populations, whereas T was the more important covariate in spike trains of the SUB than of the CA1 population (Figure 4H). In order to quantify this coding difference between CA1 and SUB, we introduced a MS score. The MS score is the product of all rSCCs in each neuron (see STAR Methods section "Mixed selectivity score (MS-score)"). This results in a maximal score for neurons that have equal contributions of P, H, and S to their firing rate. The average MS score for SUB neurons was 5.8e−3 ± 7.1e−4, which is one order of magnitude larger than that for the CA1 neurons, which had an MS score of 3.8e−4 ± 1.2e−4 (Figure 4I; Wilcoxon rank-sum test, p = 6.86e−9). MS scores were comparable between proximal and distal parts of SUB (Figure S5H). Thus, in contrast to CA1 cells, SUB neurons integrate information regarding P, H, and S through a highly mixed-selective code in which every neuron expresses differential contributions of the three covariates.

MS in SUB is task modulated

To better understand the function of MS during navigation, we next compared the neural representations of behavioral covariates between the two tasks (ST and OF). Although tuning to P, H, and S (Figures S3A and S3B), as well as information rate and information content (Figures S3C and S3D), Figures S3 failed to show much difference between OF and ST, we hypothesized that the level of MS in the network is more strongly dependent on task factors. Using the previously introduced GLM, we found that, for CA1, the distribution of selected models was similar between OF and ST (Figure 5A; no changes in model selection were higher than the 99.9th percentile of a shuffled distribution). However, the proportion of SUB neurons for which the most complex model (PHS) performed significantly better than the others increased from OF to ST (Figure 5A). Other changes in model selection between OF and ST were not significant; however, the in-

crease from 39.4% to 55.9% of neurons best fitted with the PHS model was significantly beyond the 99.9th percentile of a distribution of shuffled data (Figure 5A). A total of 2.3% and 1.4% of the neurons in SUB and CA1, respectively, were tuned to P only during ST and not in OF. In SUB, 6.9% and 15.4% more principal neurons were tuned to H and S, respectively, in ST but not in OF. In CA1, 5.8% of the neurons lost their tuning for S after transitions from OF to ST (Figure 5B). When considering tuning to E and T (which were always fitted in parallel to the other three covariates), it was apparent that the activity in the largest proportion of SUB neurons (35.6% in OF and 51.6% in ST) was best fitted by all five covariates (Figure S3E), whereas in CA1, the largest proportions of neurons (42.5% in OF and 41.5% in ST) were best fitted by a four-covariate model (Figure S3E), in most cases PHET or PSET (see Figure S3G). The task change did not affect the overall likelihood of the best-performing model (Figure S3F), indicating that the fitting procedure performed similarly in both behavioral situations.

The increase in MS in SUB was reflected in the relative contribution of each single covariate to the firing rate of the neurons. When we plotted the rSCC separately for OF and ST, it appeared that during the ST condition the SUB datapoints are more dispersed in the 3D plot, indicating a more equal contribution of the three covariates to the firing pattern of the cells (Figure 5C). To quantify this phenomenon, we calculated the MS score introduced before. Taking the product between all rSCC s of the three covariates showed that, after transitions from OF to ST, there was a significant increase in MS in the SUB (from 1.3e−3 ± 0.58e−3 in OF to 2.8e−3 ± 0.42e−3 in ST, p = 0.00018, Wilcoxon rank-sum test). In CA1, the MS score was low under both task conditions (0.02e−3 ± 0.02e−3 in OF to 0.07e−3 ± 0.01e−3 in ST, p = 0.47, Wilcoxon rank-sum test, Figure 5D). In summary, even if there is no obvious remapping between OF and ST (neither in CA1 nor in SUB), the population decoding improved in the ST condition.

High decoding accuracy for P, H, and S in SUB

The previous analyses established that a statistical model requires the mix of P, H, and S covariates to best explain the data and more so in SUB than CA1 and in ST than OF. However, it is not straightforward to directly translate these results into

Figure 4. Modeling the data with a generalized linear model (GLM) shows contribution of P, H, and S to firing rates of individual neurons

(A–C) Spatial rate maps and tuning curves for H and S for three example neurons generated from data (top panels labeled with DATA) or from spike trains reproduced with the GLM (bottom panels labeled with MODEL). Symbols as in Figure 3. Note minimal tuning to H in the CA1 neuron (pFR, ~0.3 Hz).

(D) Log-likelihoods (black circles) resulting from fitting the spike train of each neuron in (A)–(C) with models encoding one covariate—P, H, or S—or with combined models containing two or more covariates. Models that performed significantly better than all less-complex models are indicated in red. Error bars indicate SEM across 10 cross-validation folds. Log-likelihoods from models containing ensemble activity (E) or theta (T) as additional covariates (Figure S3) were collapsed with simpler models containing P, H, and/or S (for example PE is collapsed with P).

(E) Bar graphs showing fractions of neurons from CA1 (red bars) and SUB (blue bars) selecting the different models. With the exception of the PHS and PS models, models were not selected at significantly different frequencies by the two populations. PS, however, was selected by 25.9% CA1 compared to 8.8% of SUB neurons and PHS was selected by 46.5% of the neurons in CA1 as compared to 53.7% in SUB (difference, >99.9 percentile of a model-selection shuffled distribution).

(F) Information rate (info rate; top panels) and relative single-covariate contribution (rSCC; bottom panels) for the neurons shown to the left in (A)–(D). Note that even if all three neurons in (D) select the most complex model (PHS), the rSCCs indicate that the firing of the CA1 neuron is predominantly determined by the animal's position, whereas the firing of the two SUB neurons is determined by P and H or P and S, respectively.

(G) 3D plot of the rSCCs of P, H, and S. Each neuron is represented by one datapoint in red for CA1 and in blue for SUB. Note abundance of SUB neurons with high values for more than one rSCC.

(H) rSCC mean and SEM across the populations of neurons in CA1 (red) and SUB (blue) for all covariates, namely, P, H, S, E, and T.

(I) Mixed selectivity score (MS score) defined as the product between rSCC(P), rSCC(H), and rSCC(S). Error bars indicate SEM.
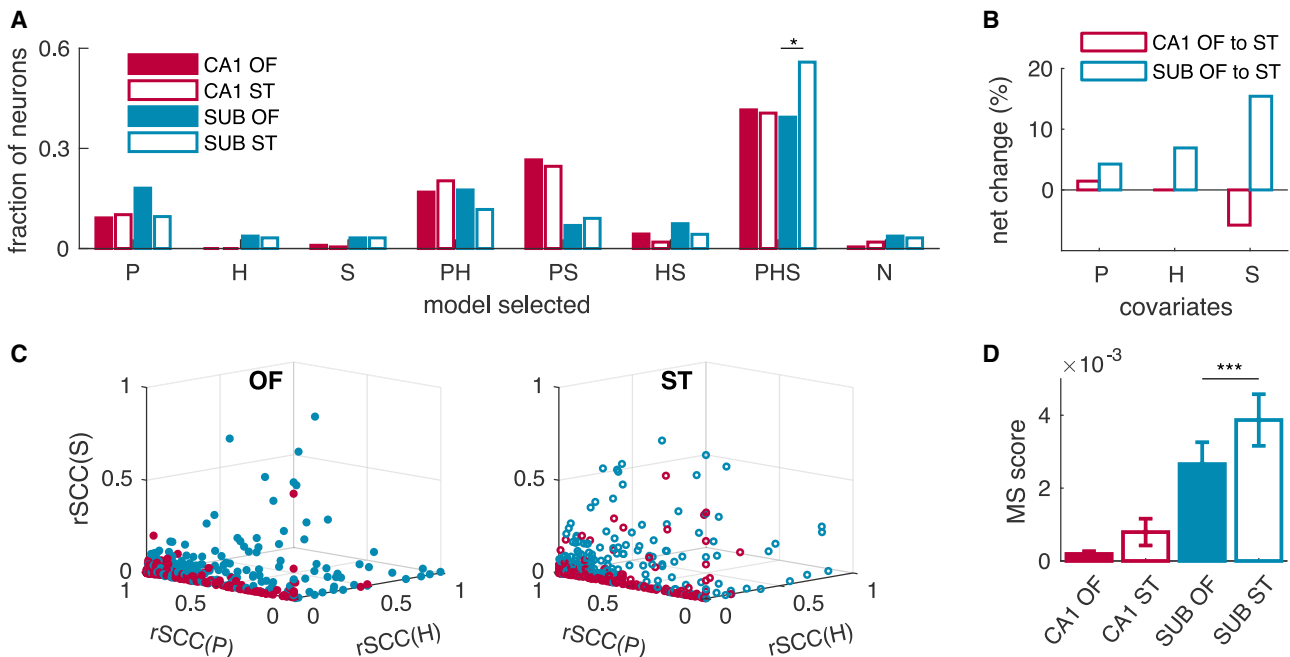
Figure 5. SUB neurons exhibit more mixed selectivity in the spatial task (ST) than during random foraging in the open field (OF)
(A) Model selection in CA1 (red bars) and SUB (blue bars) for recordings in OF (filled bars) and ST (empty bars). In SUB but not CA1, the population of neurons best fitted by the PHS model is significantly increased, in response to the task change from OF to ST.
(B) Percentage of neurons that add or remove the covariates P, H, and S after the task change.
(C) rSCC for P, H, and S for each neuron in the OF (left graph) and the ST (right graph) for the CA1 (red datapoints) and the SUB (blue datapoints).
(D) MS score in CA1 (red bars) and SUB (blue bars) in the OF (full bars) and the ST (empty bars). Error bars indicate SEM.

what can actually be read out from these codes in upstream cell populations. In order to understand what the large amount of MS in SUB might add to the hippocampal output, we therefore tried to decode P, H, and S simultaneously from cell populations in either CA1 or SUB.

Methodological constrains limited the number of simultaneously recorded neurons in SUB (the recording range was three to five neurons per SUB tetrode). In order to study decoding performance as a function of population size and extrapolate the decoding power of the two regions across tasks for a large number of neurons, we exploited the GLM framework to resample all the recorded neurons over two concatenated sessions for ST and OF, respectively (see STAR Methods section "Resampling of neurons"). To attain population sizes large enough for decoding, we randomly assembled populations of neurons that had not been simultaneously recorded in the first place. This allowed decoding of P, H, and S at the same time. Decoding of all three covariates was more accurate when using data from SUB than CA1 (Figures 6A and 6B). To assess the significance of this difference, we shuffled the neural populations with respect to their anatomical origin and computed the differences between two sets of randomized selections of neurons (ensuring equal probability to select neurons from the CA1 or the SUB cell population; Figure S6). This shuffling analysis showed that the difference in decoding accuracy deviated significantly from a distribution of shuffled differences (p < 10e−3) for decoding P, H, and S in the OF and the ST. Although smaller decoding errors for SUB than for CA1 were found across all population sizes, these differ-

ences were only significantly different from the differences in a shuffled distribution for population sizes larger than 100 neurons for P, larger than 20 neurons for H, and larger than 50 neurons for S (for population sizes of 5, 20, 50, 100, and 150, respectively, significance values were, for P: p = 0.66, 0.61, 0.32, 0.005, and 0; for H: p = 0.02, 0, 0, 0, and 0 and for S: p = 0.595, 0.240, 0.042, 0.033, and 0.009).

To better understand the difference in coding in SUB versus CA1, we also estimated the relative amount of perfectly decoded bins (PDBs). The PDB is the number of time bins in the time series for which the decoder settled on the exact bin of P, H, or S where the animal was in during that time bin. For all three covariates, namely, P, H, and S, the amount of PDB was higher when using SUB as opposed to CA1 neurons (Figure 6C; these differences were significantly different from randomized distributions; Figure S6A). Figure S7A shows furthermore that in same-length time windows, the decoder reaches bin size decoding of P, H, and S in SUB whereas in CA1 it does not. This finding indicates that, for all three covariates, the neural code in SUB allowed more accurate decoding than the one from CA1, suggesting that the code of the SUB is an effective way of transmitting signals to downstream target regions.

Comparing decoding accuracies across the two tasks showed that the decoding error for P, H, and S was significantly reduced in the ST versus OF (Figures 6A and 6B; p < 10e−3 compared to a shuffled distribution; Figure S6B). Decoding accuracy was higher in the ST than the OF also when counting the number of PDBs (Figure 6C). Although this was true for both CA1 and
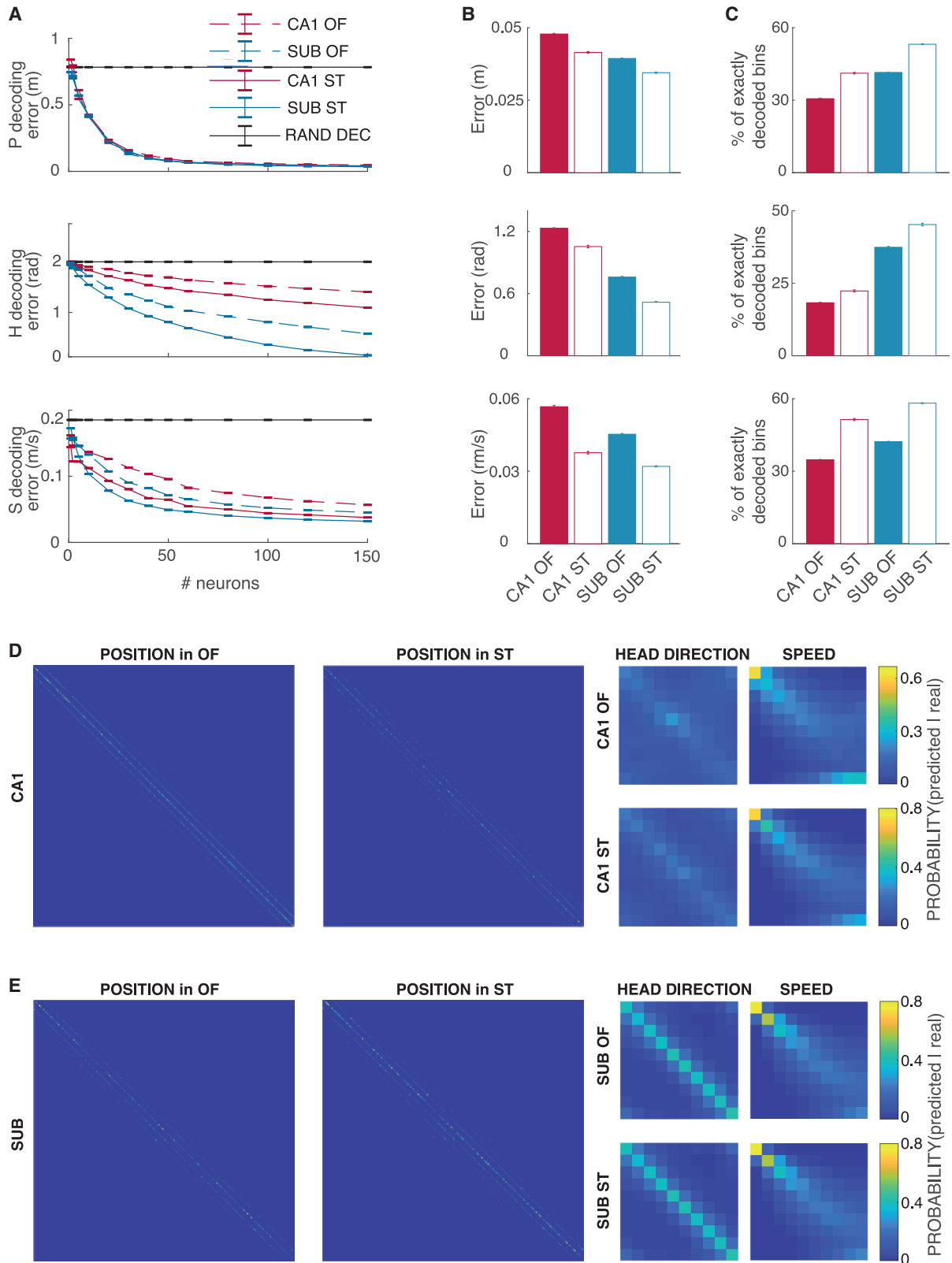
(legend on next page)

SUB, overall, the SUB reached a significantly lower number of decoding errors as well as higher proportions of PDB for P, H, and S (Figures 6A and 6B; p < 10e−3 in shuffling test of significance; Figure S6B). In order to ensure that the improved decoding in the ST did not result from a behavioral bias, we plotted confusion matrices displaying decoding accuracies for the different bin identities (Figures 6D and 6E). For H and P, the confusion matrices did not show any bias toward disproportionally better decoding in certain bin clusters of the data than others. For speed decoding, however, we observed a bias toward better decoding in the lowest speed bins (0–0.1 m/s). This bias was present for decoding from CA1 and SUB and for the OF and ST and might be a result of the disproportionally high sampling in these bins. Furthermore, even though the average speed was similar between OF and ST, the distribution was skewed toward lower and higher values (meaning that in the ST, the animal would be more often entirely immobile in order to consume the reward and then run faster to the next well). We therefore tested if decoding was affected by these differential speed distributions. By sampling the speed equally for decoding from OF and ST, we could show that decoding of all three covariates—P, H, and S—was still improved in the ST over the OF (Figures S7B–S7D).

Decoding was more accurate in the ST for both SUB and the CA1 data (see Figure 6A). Even if decoding accuracy from SUB outperforms the one from CA1 data for every behavioral covariate and session type, the improved decoding cannot be solely attributed to higher levels of MS of SUB neurons. As correlations between behavioral covariates (independent of the neural spiking) might affect decoding results, we estimated the mutual information (MI) between P, H, and S. This analysis shows that MI between behavioral covariates tends to be higher in ST than in OF (MI in ST versus OF between P and H: 1.25 bit versus 0.58 bit; between P and S: 0.76 bit versus 0.51 bit; between H and S: 0.01 bit versus 0.02 bit). Because our decoder maximizes its likelihoods for all three behavioral covariates simultaneously, the higher MI between behavioral covariates likely affects decoding accuracy independently of MS coding in SUB (even if the conjoint covariate activity is nowhere part of the model). Higher decoding accuracies will therefore be a combination of higher levels of MS in SUB neurons and a better readout due to stronger correlated behavioral covariates in the ST.

One might argue that the superiority of decoding in the SUB data might be affected by our choice of GLM as a modeling framework, as the goodness-of-fit of a model always depends on the choice of the measure used. However, in our data, neural activity in CA1 and SUB is equally well modeled by the GLM when the goodness-of-fit is measured by the difference in the likelihood of the best model and the average firing rate model (Figure S7E). Furthermore, when the goodness-of-fit is measured with explained deviance, the model fit is even better in CA1 than in SUB neurons (Figure S7F). Additionally, the task choice does not appear to affect the model performance. We therefore consider it unlikely that the decoding results should be merely a consequence of our choice of the GLM as a modeling framework.

The higher decoding accuracy in SUB is somewhat surprising given that the tuning curves of SUB neurons are broader and have a more slowly changing visual appearance than tuning curves from CA1 neurons (see tuning curves for P, H, and S in Figures 2 and S5). A possible explanation for this would be that at the population level, neural activity in SUB is fluctuating at a shorter spatial scale than in CA1. The visual appearance of the spatial autocorrelations for all neurons recorded in SUB and CA1 (Figure S7G) seems to suggest that the two population vectors change at different scales. Plotting the average population vector autocorrelation as a function of distance between spatial bins (Figure S7H) shows that the peak in SUB is narrower than that in CA1. The more accurate P decoding from SUB data might therefore reflect a faster change of the population code between nearby spatial locations in SUB than that in CA1. This therefore points to an important role of SUB at integrating multiple information streams from the hippocampal formation and providing an accurate navigational code to downstream regions.

DISCUSSION

Downstream regions of the hippocampal formation rely on output from the CA1 and SUB subfields during behaviors when navigational and mnemonic information is relevant. Although spatial coding in CA1 has been investigated extensively in place cells, it has remained elusive how SUB transforms the positional code. Here, we record neural activity during two different spatial behaviors and show that SUB provides a mixed selective code, from which three correlates of navigation—P, H, and S—can each be decoded at higher accuracies, from a similar number of neurons, than from the CA1 output. The presence of MS is consistent with an early report of place and H tuning in the same SUB cells (Sharp and Green, 1994) but takes the finding further by quantifying the mixing of three navigational variables and by showing that decodability is enhanced compared to that in CA1. Our experiments further demonstrate that MS in SUB increases with navigational task demands, suggesting a

---

Figure 6. More accurate decoding of P, H, and S from a fixed number of cells in SUB compared to CA1

(A) Decoding of P (top panel), H (middle panel), and S (bottom panel) for different numbers of resampled neurons. For all three covariates, decoding errors are smaller in SUB (blue) than CA1 (red) and smaller in the ST (full lines) than in the OF (dashed lines).

(B and C) Decoding error (B) and percentage of perfectly decoded bins (the number of time bins in which the decoded was the exact same as where the animal roamed at this instance; C) in a population of 150 neurons, shown separately for P (top panels), H (middle panels), and S (bottom panels). Note smaller decoding error for position in SUB than CA1. Error bars correspond to 1 SEM over 20 samples of the identities of the resampled cells.

(D and E) Confusion matrices for decoding in CA1 (D) and SUB (E). Columns show decoded covariates, namely, P (left), S (middle), and H (right); rows show tasks, namely, OF (top) and ST (bottom). Within each confusion matrix C, rows identify decoded covariate bins, wehreas columns identify bins for the covariate in the data, such that position confusion matrices are 900 × 900 and H or S matrices are 10 × 10. The color-coded element C_{ij} of the confusion matrix represents the fraction of time bins for which the decoded covariate bin is i given that the actual covariate bin was j. Bright-yellow colors indicate a maximum overlap, whereas dark-blue colors indicate no overlap between decoded and actual covariate.

behaviorally dependent information transfer to downstream regions. The activity of individual SUB neurons, which are informative on a broader spatial scale (and consequently, a longer temporal scale) than CA1 cells, may represent an integration of inputs from CA1 with information from other areas. In the present work, the combination of covariates into a mixed representation is demonstrated for three navigational parameters, namely, P, H, and S. However, it has been shown that hippocampal neurons respond also to other covariates like odors (Dusek and Eichenbaum, 1997), texture (Wood et al., 2000), and time (Eichenbaum, 2017; Fortin et al., 2002; Hampson and Deadwyler, 2003; Kesner et al., 2002). This raises the possibility that SUB neurons combine a wide range of covariates, resulting potentially in a behavior-dependent mixed code for retrieval of a broad spectrum of experiences in downstream regions.

In an information theoretical analysis, Kim et al. (2012) showed that SUB is well suited to transmit information about the animal's position by using fewer neurons than the CA1 region. We show here, with experimental data, that we can indeed decode P at a higher accuracy from SUB data than from CA1 data, when the number of cells is the same. A similar enhancement was seen in SUB for decoding H and S. For P, decoding can also be extremely accurate in CA1 when a high number of neurons is available (Pfeiffer and Foster, 2013; Wilson and McNaughton, 1993), but it is unclear if the density of connectivity to distant downstream regions is high enough to integrate over such large arrays of inputs. As SUB is the major recipient of CA1 output (Cappaert et al., 2015), projection density to this region might allow SUB to integrate over a large array of neurons in order to translate the sparse code from CA1 into a code that can be decoded accurately from a lower number of neurons at the next stage. Neurons in CA1 have been shown to combine different behavioral covariates in tasks for which multiple factors, such as olfactory, tactile, and positional cues, are salient to the animal (Komorowski et al., 2009). This type of MS has high similarities with later defined non-linear MS in other systems (Fusi et al., 2016; Rigotti et al., 2013). In contrast, the SUB neurons we show in the present study are highly mixed-selective even in situations in which the cognitive map is primarily concerned with representing spatial parameters. The GLM seems to be picking up both non-linear and linear MS. Although the model selection process appears to be sensitive to even small contributions of different covariates to the firing rate of the neurons (non-linear mixing), the MS score is sensitive to the degree at which a given covariate modulates the firing rate (linear mixing). SUB neurons appear to combine the different covariates by extensive linear mixing, allowing it to broadcast the hippocampal output in a manner that is complementary to CA1 and may facilitate efficient readout in distant brain areas.

Although P, H, and S are represented in both CA1 and SUB neurons, the combinatorial expression of those variables was a lot more common in SUB neurons. MS neurons have previously been found in a number of cortical and subcortical brain areas (Asaad et al., 1998; Freedman and Assad, 2009; Hardcastle et al., 2017; Meister et al., 2013; Rigotti et al., 2013; Rishel et al., 2013), including the hippocampal areas CA1 and CA3 (Acharya et al., 2016; McKenzie et al., 2014; Wood et al., 2000). Here, we provide additional evidence that CA1 neurons encode multiple variables simultaneously, but we show further that representations in this subfield are dominated by P, whereas other navigational covariates have a less predominant influence on the firing rates. So even if from the perspective of the model selection process, MS exists in both CA1 and SUB, SUB displays a more evenly mixed code, as expressed in the MS score.

Our recordings in the ST further show that neurons in SUB change their level of MS depending on the task in which the animal is involved. Although it has been observed that SUB neurons change their strength of tuning to movement direction depending on the animal's behavior (Olson et al., 2017), we show here that task-dependent changes in tuning strength, in environments with high visual similarity, are not limited to direction (H in our case) but also encompass S and P. Higher cortical areas like the prefrontal and the posterior parietal cortex have previously been shown to adapt their mixed selective neural response to different task demands (Mante et al., 2013; Parthasarathy et al., 2017; Raposo et al., 2014; Stokes et al., 2013). The present findings suggest that, in the hippocampal formation, navigational output is modified by increasing the level of MS through a processing step in SUB. This is accompanied with improved decoding accuracies in both the OF and the ST. As the decoding accuracies in the ST also increase when compared to OF for CA1 data, these improvements cannot be explained solely at the level of neuronal representation but likely result from a combination of increased MS and a higher correlation between behavioral covariates. In concrete terms, this means that in SUB, in which H and S are highly expressed in the firing patterns of the cells, these covariates might help improve decoding accuracy for P and vice versa. This likely serves a mechanism for downstream regions to obtain access to a broad spectrum of hippocampal information when it is behaviorally relevant. However, although mixed-selective coding of P, H, and S proved to be effective at predicting variation on these variables, we cannot claim that these variables constitute the only neural code of SUB. The models compare only the covariates that are put into them. As SUB has been shown to be important for learning and memory (Cembrowski et al., 2018a; Morris et al., 1990; Roy et al., 2017), it is likely that its activity relates also to behavioral or cognitive variables that were absent from our analysis of navigational behavior.

Previous theoretical work has shown that mixed-selectivity coding schemes offer certain computational advantages over single-variable selectivity. When multiple covariates are combined in various quantities in the population of neurons projecting to a downstream target region, a downstream neuron can linearly combine an arbitrary subset of the inputs in order to reconstruct the value of a single covariate (Fusi et al., 2016; Ganguli and Sompolinsky, 2012). This means that, with no need for precise prepatterned connections, all the information can be transmitted, allowing downstream regions to integrate different covariates without having to repattern connections. Keeping in mind that neurons from SUB may consist of genetically compartmentalized subpopulations with each projecting to only a few selected target areas (Bienkowski et al., 2018; Cembrowski et al., 2018b; Naber and Witter, 1998; Witter, 2006), the mixed code from SUB may be specifically tailored to ensure that a wide range of covariates computed in the

Cell Reports
Article

hippocampal formation is accessible to distant projection areas, despite the limited number of SUB cells that may project there. Although the widespread projections from CA1 provide position information to a large array of downstream areas, the SUB output ensures the inclusion of a broader spectrum of task-relevant information in the hippocampal output, which is encoded in ways that can be integrated efficiently by downstream target regions.

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
- METHOD DETAILS
  - Electrode implantation and surgery
  - Recording procedures
  - Behavioral procedures
  - Histology and reconstruction of tetrode placement
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Spike sorting and single-unit selection
  - Tracking, rate maps and tuning curves
  - Spatial correlation across session type
  - Calculating information rate and information content
  - Preparation of the data for the Poisson GLM
  - Poisson GLM
  - Poisson spiking
  - Testing performance of GLM framework with synthetic cells
  - Learning
  - Hyperparameters optimization
  - Model selection
  - Splitting between external and brain internal variables
  - Relative single covariate contributions (rSCC)
  - Mixed selectivity score (MS-score)
  - Tuning curves predicted by the model
  - Goodness of fit
  - Testing for significance in model selection between CA1 or SUB and between OF or ST
  - Decoding
  - Resampling of neurons
  - Randomization of decoding
  - Statistical testing

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109175.

REFERENCES

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.

Asaad, W.F., Rainer, G., and Miller, E.K. (1998). Neural activity in the primate prefrontal cortex during associative learning. Neuron 21, 1399–1407.

Bienkowski, M.S., Bowman, I., Song, M.Y., Gou, L., Ard, T., Cotter, K., Zhu, M., Benavidez, N.L., Yamashita, S., Abu-Jaber, J., et al. (2018). Integration of gene expression and brain-wide connectivity reveals the multiscale organization of mouse hippocampal networks. Nat. Neurosci. 21, 1628–1643.

Boccara, C.N., Sargolini, F., Thoresen, V.H., Solstad, T., Witter, M.P., Moser, E.I., and Moser, M.-B. (2010). Grid cells in pre- and parasubiculum. Nat. Neurosci. 13, 987–994.

Boccara, C.N., Kjonigsen, L.J., Hammer, I.M., Bjaalie, J.G., Leergaard, T.B., and Witter, M.P. (2015). A three-plane architectonic atlas of the rat hippocampal region. Hippocampus 25, 838–857.

Brown, E.N., Barbieri, R., Ventura, V., Kass, R.E., and Frank, L.M. (2002). The time-rescaling theorem and its application to neural spike train data analysis. Neural Comput. 14, 325–346.

Buzsáki, G. (1989). Two-stage model of memory trace formation: a role for "noisy" brain states. Neuroscience 31, 551–570.

Cappaert, N., Strien, N.M., and Witter, M.P. (2015). Hippocampal Formation. In The Rat Nervous System (Academic Press).

Cembrowski, M.S., Phillips, M.G., DiLisio, S.F., Shields, B.C., Winnubst, J., Chandrashekar, J., Bas, E., and Spruston, N. (2018a). Dissociable Structural and Functional Hippocampal Outputs via Distinct Subiculum Cell Classes. Cell 173, 1280–1292.e18.

Cembrowski, M.S., Wang, L., Lemire, A.L., Copeland, M., DiLisio, S.F., Clements, J., and Spruston, N. (2018b). The subiculum is a patchwork of discrete subregions. eLife 7, e37701.

Czurkó, A., Hirase, H., Csicsvari, J., and Buzsáki, G. (1999). Sustained activation of hippocampal pyramidal cells by 'space clamping' in a running wheel. Eur. J. Neurosci. 11, 344–352.

Dusek, J.A., and Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. Proc. Natl. Acad. Sci. USA 94, 7109–7114.

Eden, U.T., and Kramer, M.A. (2010). Drawing inferences from Fano factor calculations. J. Neurosci. Methods 190, 149–152.

Eichenbaum, H. (2017). On the Integration of Space, Time, and Memory. Neuron 95, 1007–1018.

Fortin, N.J., Agster, K.L., and Eichenbaum, H.B. (2002). Critical role of the hippocampus in memory for sequences of events. Nat. Neurosci. 5, 458–462.

Freedman, D.J., and Assad, J.A. (2009). Distinct encoding of spatial and nonspatial visual information in parietal cortex. J. Neurosci. 29, 5671–5680.

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

Fyhn, M., Molden, S., Witter, M.P., Moser, E.I., and Moser, M.-B. (2004). Spatial representation in the entorhinal cortex. Science 305, 1258–1264.

Ganguli, S., and Sompolinsky, H. (2012). Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. Annu. Rev. Neurosci. 35, 485–508.

Gigg, J. (2006). Constraints on hippocampal processing imposed by the connectivity between CA1, subiculum and subicular targets. Behav. Brain Res. 174, 265–271.

Goltstein, P.M., Meijer, G.T., and Pennartz, C.M. (2018). Conditioning sharpens the spatial representation of rewarded stimuli in mouse primary visual cortex. eLife 7, e37683.

Guisan, A., and Zimmermann, N.E. (2000). Predictive habitat distribution models in ecology. Ecol. Modell. 135, 147–186.

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

Hampson, R.E., and Deadwyler, S.A. (2003). Temporal firing characteristics and the strategic role of subicular neurons in short-term memory. Hippocampus 13, 529–541.

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.

Hasselmo, M.E. (2012). How We Remember: Brain Mechanisms of Episodic Memory (MIT Press).

Høydal, Ø.A., Skytøen, E.R., Andersson, S.O., Moser, M.-B., and Moser, E.I. (2019). Object-vector coding in the medial entorhinal cortex. Nature 568, 400–404.

Ishizuka, N. (2001). Laminar organization of the pyramidal cell layer of the subiculum in the rat. J. Comp. Neurol. 435, 89–110.

Jung, M.W., Qin, Y., McNaughton, B.L., and Barnes, C.A. (1998). Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. Cereb. Cortex 8, 437–450.

Kesner, R.P., Gilbert, P.E., and Barua, L.A. (2002). The role of the hippocampus in memory for the temporal order of a sequence of odors. Behav. Neurosci. 116, 286–290.

Kim, S.M., Ganguli, S., and Frank, L.M. (2012). Spatial information outflow from the hippocampal circuit: distributed spatial coding and phase precession in the subiculum. J. Neurosci. 32, 11539–11558.

Komorowski, R.W., Manns, J.R., and Eichenbaum, H. (2009). Robust conjunctive item-place coding by hippocampal neurons parallels learning what happens where. J. Neurosci. 29, 9918–9929.

Kraus, B.J., Brandon, M.P., Robinson, R.J., 2nd, Connerney, M.A., Hasselmo, M.E., and Eichenbaum, H. (2015). During Running in Place, Grid Cells Integrate Elapsed Time and Distance Run. Neuron 88, 578–589.

Kropff, E., Carmichael, J.E., Moser, M.-B., and Moser, E.I. (2015). Speed cells in the medial entorhinal cortex. Nature 523, 419–424.

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

Lansink, C.S., Jackson, J.C., Lankelma, J.V., Ito, R., Robbins, T.W., Everitt, B.J., and Pennartz, C.M.A. (2012). Reward cues in space: commonalities and differences in neural coding by hippocampal and ventral striatal ensembles. J. Neurosci. 32, 12444–12459.

Leutgeb, S., Ragozzino, K.E., and Mizumori, S.J.Y. (2000). Convergence of head direction and place information in the CA1 region of hippocampus. Neuroscience 100, 11–19.

Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., and Burgess, N. (2009). Boundary vector cells in the subiculum of the hippocampal formation. J. Neurosci. 29, 9771–9777.

Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503, 78–84.

McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. 102, 419–457.

McCullagh, P., and Nelder, J.A. (1989). Generalized Linear Models (Chapman and Hall/CRC).

McKenzie, S., Frank, A.J., Kinsky, N.R., Porter, B., Rivière, P.D., and Eichenbaum, H. (2014). Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. Neuron 83, 202–215.

McNaughton, B.L., Barnes, C.A., and O'Keefe, J. (1983). The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. Exp. Brain Res. 52, 41–49.

Meister, M.L.R., Hennig, J.A., and Huk, A.C. (2013). Signal multiplexing and single-neuron computations in lateral intraparietal area during decision-making. J. Neurosci. 33, 2254–2267.

Miller, E.K., Erickson, C.A., and Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J. Neurosci. 16, 5154–5167.

Morris, R. (2007). Theories of Hippocampal Function. In The Hippocampus Book (Oxford University Press), pp. 581–714.

Morris, R.G., Schenk, F., Tweedie, F., and Jarrard, L.E. (1990). Ibotenate Lesions of Hippocampus and/or Subiculum: Dissociating Components of Allocentric Spatial Learning. Eur. J. Neurosci. 2, 1016–1028.

Mulder, A.B., Shibata, R., Trullier, O., and Wiener, S.I. (2005). Spatially selective reward site responses in tonically active neurons of the nucleus accumbens in behaving rats. Exp. Brain Res. 163, 32–43.

Naber, P.A., and Witter, M.P. (1998). Subicular efferents are organized mostly as parallel projections: a double-labeling, retrograde-tracing study in the rat. J. Comp. Neurol. 393, 284–297.

Nitz, D.A. (2006). Tracking route progression in the posterior parietal cortex. Neuron 49, 747–756.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. Exp. Neurol. 51, 78–109.

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

O'Keefe, J., and Nadel, L. (1978). The hippocampus as a cognitive map (Oxford University Press).

O'Mara, S. (2006). Controlling hippocampal output: the central role of subiculum in hippocampal information processing. Behav. Brain Res. 174, 304–312.

O'Mara, S.M., Commins, S., Anderson, M., and Gigg, J. (2001). The subiculum: a review of form, physiology and function. Prog. Neurobiol. 64, 129–155.

Olson, J.M., Tongprasearth, K., and Nitz, D.A. (2017). Subiculum neurons map the current axis of travel. Nat. Neurosci. 20, 170–172.

Padilla-Coreano, N., Canetta, S., Mikofsky, R.M., Alway, E., Passecker, J., Myroshnychenko, M.V., Garcia-Garcia, A.L., Warren, R., Teboul, E., Blackman, D.R., et al. (2019). Hippocampal-Prefrontal Theta Transmission Regulates Avoidance Behavior. Neuron 104, 601–610.e4.

Parthasarathy, A., Herikstad, R., Bong, J.H., Medina, F.S., Libedinsky, C., and Yen, S.-C. (2017). Mixed selectivity morphs population codes in prefrontal cortex. Nat. Neurosci. 20, 1770–1779.

Peck, E.L., Peck, C.J., and Salzman, C.D. (2014). Task-dependent spatial selectivity in the primate amygdala. J. Neurosci. 34, 16220–16233.

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. (2009). Time-rescaling methods for the estimation and assessment of non-Poisson neural encoding models. In Advances in Neural Information Processing Systems 22, Y. Bengio, D. Schuurmans, J.D. Lafferty, C.K.I. Williams, and A. Culotta, eds. (Curran Associates, Inc.), pp. 1473–1481.

Poulter, S., Lee, S.A., Dachtler, J., Wills, T.J., and Lever, C. (2020). Vector trace cells in the subiculum of the hippocampal formation. Nat. Neurosci. *24*, 266–275.

Pratt, W.E., and Mizumori, S.J.Y. (2001). Neurons in rat medial prefrontal cortex show anticipatory rate changes to predictable differential rewards in a spatial memory task. Behav. Brain Res. *123*, 165–183.

Raposo, D., Kaufman, M.T., and Churchland, A.K. (2014). A category-free neural population supports evolving demands during decision-making. Nat. Neurosci. *17*, 1784–1792.

Remondes, M., and Wilson, M.A. (2013). Cingulate-hippocampus coherence and trajectory coding in a sequential choice task. Neuron *80*, 1277–1289.

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.

Rishel, C.A., Huang, G., and Freedman, D.J. (2013). Independent category and spatial encoding in parietal cortex. Neuron *77*, 969–979.

Roy, D.S., Kitamura, T., Okuyama, T., Ogawa, S.K., Sun, C., Obata, Y., Yoshiki, A., and Tonegawa, S. (2017). Distinct Neural Circuits for the Formation and Retrieval of Episodic Memories. Cell *170*, 1000–1012.e19.

Saleem, A.B., Diamanti, E.M., Fournier, J., Harris, K.D., and Carandini, M. (2018). Coherent encoding of subjective spatial position in visual cortex and hippocampus. Nature *562*, 124–127.

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

Sharp, P.E. (1997). Subicular cells generate similar spatial firing patterns in two geometrically and visually distinctive environments: comparison with hippocampal place cells. Behav. Brain Res. *85*, 71–92.

Sharp, P.E. (2006). Subicular place cells generate the same "map" for different environments: comparison with hippocampal cells. Behav. Brain Res. *174*, 206–214.

Sharp, P.E., and Green, C. (1994). Spatial correlates of firing patterns of single cells in the subiculum of the freely moving rat. J. Neurosci. *14*, 2339–2356.

Skaggs, W.E., McNaughton, B.L., and Gothard, K.M. (1993). An Information-Theoretic Approach to Deciphering the Hippocampal Code. In Advances in Neural Information Processing Systems 5, [NIPS Conference] (Morgan Kaufmann Publishers Inc.), pp. 1030–1037.

Solstad, T., Boccara, C.N., Kropff, E., Moser, M.-B., and Moser, E.I. (2008). Representation of geometric borders in the entorhinal cortex. Science *322*, 1865–1868.

Spellman, T., Rigotti, M., Ahmari, S.E., Fusi, S., Gogos, J.A., and Gordon, J.A. (2015). Hippocampal-prefrontal input supports spatial encoding in working memory. Nature *522*, 309–314.

Squire, L.R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. Psychol. Rev. *99*, 195–231.

Squire, L.R., Genzel, L., Wixted, J.T., and Morris, R.G. (2015). Memory consolidation. Cold Spring Harb. Perspect. Biol. *7*, a021766.

Stewart, S., Jeewajee, A., Wills, T.J., Burgess, N., and Lever, C. (2013). Boundary coding in the rat subiculum. Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, 20120514.

Stokes, M.G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. Neuron *78*, 364–375.

Taube, J.S., and Burton, H.L. (1995). Head direction cell activity monitored in a novel environment and during a cue conflict situation. J. Neurophysiol. *74*, 1953–1971.

Taube, J.S., Muller, R.U., and Ranck, J.B., Jr. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. J. Neurosci. *10*, 420–435.

Whitlock, J.R., Pfuhl, G., Dagslott, N., Moser, M.-B., and Moser, E.I. (2012). Functional split between parietal and entorhinal cortices in the rat. Neuron *73*, 789–802.

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

Winocur, G., and Moscovitch, M. (2011). Memory transformation and systems consolidation. J. Int. Neuropsychol. Soc. *17*, 766–780.

Witter, M.P. (2006). Connections of the subiculum of the rat: topography in relation to columnar and laminar organization. Behav. Brain Res. *174*, 251–264.

Wood, E.R., Dudchenko, P.A., Robitsek, R.J., and Eichenbaum, H. (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. Neuron *27*, 623–633.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Cresyl Violet for Nissl stain	Sigma Alderich	1.05235
Deposited data		
Spike-time, position, head direction and theta phase data; histology data	This paper	Provided upon request
Experimental models: Organisms/strains		
Rat: Long-Evans	Charles River, in-house	https://www.criver.com/products-services/find-model/long-evans-rat?region=3616
Software and algorithms		
MATLAB R2019b	Mathworks	https://se.mathworks.com/
MClust 4.4	A.D. Redish	http://redishlab.neuroscience.umn.edu/MClust/MClust.html
Custom-made analysis code	This paper	Provided upon request
Other		
Polyimide-coated platinum-iridium (90-10%) wires	California Fine wire	N/A
Hyperdrive	Custom made	N/A
Microdrive	Axona LTD	N/A
Axio Scan.Z1	Zeiss	N/A

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Edvard I. Moser (edvard.moser@ntnu.no).

Materials availability
This study did not generate new unique materials.

Data and code availability
Datasets and code supporting the current study will be provided upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Data were collected from eight male Long Evans rats, which were experimentally naive and 3–5 months old (350–600 g) at the time of implantation. The rats were group housed with 3–8 of their male littermates prior to surgery and were singly housed in large Plexiglas cages (45 × 44 × 30 cm) thereafter. The rats were kept on a 12 h light/12 h dark schedule, and humidity and temperature were strictly controlled. The experiments were performed in accordance with the Norwegian Animal Welfare Act and the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes. All experiments were approved by the Norwegian Food Safety Authority.

METHOD DETAILS

Electrode implantation and surgery
Tetrodes were constructed from four twisted 17-μm polyimide-coated platinum-iridium (90%/10%) wires (California Fine Wire). The electrode tips were plated with platinum to reduce electrode impedances to between 120–300 kΩ at 1 kHz.

Anesthesia was induced by placing the animal in a closed Plexiglas box filled with 5% isoflurane vapor. Subsequently, the animal received a subcutaneous injection of buprenorphine (0.03 mg kg^{-1}), atropine (0.05 mg kg^{-1}) and meloxicam (1.0 mg kg^{-1}) and was mounted on a stereotactic frame. The animal's body rested on a heat blanket to maintain its core body temperature during the surgical procedure. Anesthesia was maintained with isoflurane, with air flow at 1.0 liters/min and isoflurane concentration 0.75%–3%, as determined according to breathing patterns and reflex responses.

The scalp midline was subcutaneously injected with the local anesthetic lidocaine (0.5%) prior to incision. After removal of the periost, holes were drilled vertically in the skull, into which screws (M1.4) were inserted. Two screws positioned over the cerebellum were used as the electrical ground. Craniotomies were drilled anterior to the transverse sinus. Subsequently, the animal was implanted with either a hyperdrive containing 14 independently moveable tetrodes (seven animals), or a microdrive, containing a single bundle of eight tetrodes. Hyperdrive implants were always on the left side. Hyperdrive tetrodes were implanted perpendicular to the long axis of the HPC (at a 45 deg angle) between −5.5 to −5.7 mm AP and 1.2 to 1.8 ML. Each tetrode was immediately advanced by 940 μm. Microdrive tetrodes were inserted in the right HPC at −6.9 AP (from right sinus) and 2.0 ML. The tetrodes were immediately inserted to a depth of 1.5 mm. Implants were secured with dental cement (Meliodent). 8–12 h after the beginning of the surgery, the animal was treated with an additional dose of buprenorphine (0.03 mg kg^{-1}).

Recording procedures

Over the course of 1–3 weeks, tetrodes were lowered in steps of 320 μm or less, until high-amplitude theta-modulated activity appeared in the local field potential at a depth of approximately 2.0 mm. In hyperdrive experiments, at least one of the tetrodes was used to record a reference signal from white matter areas. The drive was connected to a multichannel, impedance matching, unity gain headstage. The output of the headstage was conducted via a lightweight multiwire tether cable and through a slip-ring commutator to a Neuralynx data acquisition system (Neuralynx, Tucson, AZ; Neuralynx Digital Lynx SX, for all hyperdrive-implanted animals) or via a counterbalanced lightweight multiwire cable to an Axona acquisition system (Axona Ltd., Herts, UK, for the one microdrive-implanted animal). Both cables allowed the animal to move freely within the available space. Unit activity was amplified by a factor of 3,000–5,000 and bandpass filtered 600–6,000 Hz (Neuralynx) or 800–6,700 Hz (Axona). Spike waveforms above a threshold set by the experimenter (50–80 μV) were timestamped and digitized at 32 kHz (Neuralynx) or 48 kHz (Axona) for 1 ms. In some Neuralynx recordings, the raw signals were also recorded (32 kHz). Local field potential (LFP) signals were recorded from one per tetrode for the hyperdrives and one in total per Axona microdrive. LFP signals were amplified by a factor of 250–1,000, low-pass filtered at 300–475 Hz and sampled at 1,800–2,500 Hz. The LFP channels were recorded referenced to the ground screw positioned above the animal's cerebellum (Neuralynx and Neuropixels) or against an electrode from one microdrive tetrode (Axona). For Neuralynx and Axona recordings, LEDs on the headstage were used to track the animal's movements at a sampling rate of 25 Hz (Neuralynx) or 50 Hz (Axona).

Behavioral procedures

The rats were food restricted, maintaining their weight at a minimum of 90% of their free-feeding body weight, and were food deprived 12–16 h before each training or recording session. During the 3–6 weeks prior to surgery and testing, the animals were trained to find the wells filled with chocolate oat milk and alternate between targeted search and direct run toward the home well (ST). When animals achieved a success rate higher than 90% in the ST they were also introduced to random foraging in the open field (OF) (approximately 1 to 1.5 weeks prior to implantation). The arena was a 150 × 150 cm square box with a black floor mat during the OF and a rubber-spray covered plastic plate with 37 small hemisphere incisions of 1 cm diameter arranged in a regular lattice. A cue card (a white A4 paper) on one of the walls indicated orientation. The two environments were located in the exact same place with all distal cues constant and surrounded by the same 50-cm-high black walls. As a reward, vanilla or chocolate biscuit crumbs were randomly scattered in the OF condition and in the ST, chocolate oat milk was provided via the wells in the floor. Curtains were not used, and abundant visual cues were available to the foraging rat. Between sessions in the OF and ST, the rat was placed next to the arena on an elevated flowerpot lined with towels.

Each recording day consisted of two OF and two ST sessions of each roughly 30 min. The ST was run such that all 36 random wells had to be visited at least once, resulting in 36 to 38 home run trials. Care was given that all the arena floors and walls were clean prior to beginning each recording. Data from multiple sessions of the same type was concatenated for analysis purposes. Recordings were generally performed during the dark phase of the 12 h/12 h light cycle.

Histology and reconstruction of tetrode placement

Rats were anesthetized with isoflurane (5%) and then received an overdose of sodium pentobarbital. They were subsequently perfused intracardially with saline followed by 4% formaldehyde. The brains were extracted, stored in 4% formaldehyde to be later frozen and cut in coronal sections (animals 20360, 22295, 23783, 24101 and 24116) and para-coronal sections (animal 20382, 21012 and 22098; sections were 45 degrees offset of coronal and sagittal sections in order to align with the angle of drive implantation). Sections of 30 μm were subsequently stained with cresyl violet (Nissl) and the relevant parts of CA1 and SUB were collected for analysis. For hyperdrive implants, all tetrodes from the 14-tetrode bundle were individually identified from digital photomicrographs by comparing tetrode traces from successive sections.

The depth of the recording sites was determined post hoc by comparing the deepest visible electrode trace in the tissue with the distance the electrodes were moved between the recording day and the day the animal was perfused. E.g. in Figure 1B tetrode 5 is visible in two images: the leftmost and the middle image in the second row. The track in the middle image, indicated with an empty red triangle (second triangle from the left), is the lowest point in the brain sections where the electrode track could be observed. Between the day the animal was perfused and the last day a neuron had been recorded on tetrode no. 5 the electrode had been moved 270 μm. We therefore assume that the recording sites for this electrode had been approximately 270 μm above this point. This is located in the proximal SUB which at this proximo-distal level extends above the CA1 cell layer. For tetrode no. 12 in animal 20382 the electrode was 100 μm above the recording end point and in animal 20360 the last recordings of tetrodes 8, 9 and 11 have been 100, 210 and 300 μm above the location of the electrode-tip on the day of perfusion. For four-tetrode microdrives, the tetrode bundle as a whole was localized with a similar method.

In order to approximate the anatomical distribution of the recorded cells along the proximo-distal axis, the respective regions were subdivided into proximal, middle and distal CA1 and proximal or distal SUB, respectively. The boundary between the respective subregions was approximated by dividing the region in every section into three or two (for CA1 or SUB, respectively) equally large subregions (see black dashed lines in sections of Figure S1).

QUANTIFICATION AND STATISTICAL ANALYSIS

Spike sorting and single-unit selection
Spike sorting was performed offline using manual cluster cutting methods with MClust (A. D. Redish, http://redishlab.neuroscience.umn.edu/MClust/MClust.html). Spike rate autocorrelation and cross-correlation were used as additional tools for separating or merging spike clusters. Single units were discarded if more than 0.5% of their inter-spike interval distribution was comprised of intervals less than 2 ms or if they had a mean spike rate less than 0.5 Hz. Additionally, interneurons were separated and excluded from the dataset by removing clusters with narrow waveforms and high firing rates during manual spike sorting.

Tracking, rate maps and tuning curves
Animal position was estimated by tracking the LEDs or reflective markers mounted on the implant. Only time epochs in which the animal was moving at a speed above 2 m s^{-1} were used for spatial analyses.

To generate 2D rate maps for the open-field arena, position estimates were binned into a 5 × 5 cm square grid. The spike rate in each position bin was calculated as the number of spikes recorded in the bin, divided by the time the animal spent in the bin. The resultant 2D rate map was smoothed with a Gaussian kernel with σ = 1.5 bins.

The animal's head direction was determined from the relative positions of LEDs or reflective markers on the implant. Head direction tuning curves were calculated by binning the head direction estimates into 6° bins. The spike rate in each angular bin was calculated as the number of spikes recorded in the bin divided by the time the animal spent in the bin. The resultant tuning curve was smoothed with a Gaussian kernel with σ = 1 bin, with the ends of the tuning curve wrapped together.

Spatial correlation across session type
For the spatial correlations across session types (i.e., spatial memory task (ST) versus open field (OF) the data from open field and foster maze were concatenated, respectively. Half of the session was chosen by taking 5 minutes intervals of the data which then were shifted forward n times with 30 s intervals. For each shifted sub-portion of the data, the rate map was calculated and correlated with the other half of the data. The number of shifts needed to get a stable value was determined in a saturation process where for every cell increasing number of shifts were used until three consecutive values of the correlation were below 0.5 standard deviations.

Calculating information rate and information content
Information rate (I_{rate}) and information content ($I_{content}$) were calculated as described in Skaggs et al. (1993). Briefly, every covariate space was binned (900 bins of 5 X 5 cm for position, 10 bins for head direction and 10 bins for speed) and the neurons spikes were allocated to the respective bin the animal was occupying while they were emitted. The total number of spikes in every bin was then divided by the total session time to receive $\lambda(x)$, the average firing rate in bin x. Using λ, the mean firing rate of the neuron across the entire session and $P(x)$, the probability of the animal to occupy bin x, information rate for each neuron across the entire session was calculated by:

$$I_{rate} = \int_x \lambda(x) log_2 \frac{\lambda(x)}{\lambda} P(x) dx$$

To calculate information content per spike, I_{rate} was divided by the average firing rate of the neuron across the entire session λ.

$$I_{content} = \frac{I_{rate}}{\lambda}$$

In order to ensure that the information values are minimally affected by any biases, we subtracted from all information measures the mean of the value obtained for the same measure for 100 shuffled spike distributions. The values reported in the text are with these biases removed, although our results do not qualitatively change if we do not remove the bias.

Preparation of the data for the Poisson GLM

Spike time stamps obtained from the spike sorting procedure were binned using bins of length dt = 20 ms to get a spike count train vector $\mathbf{n} = \{n(t)\}_{t=1}^{T}$ per each cell, where $n(t)$ is the spike count in time bin t, while T is the total number of time bins. The behavioral covariates taken into account are:

 P xy-position of the rat on the horizontal plane;
 H head direction of the rat on the horizontal plane;
 S running speed of the rat;
 while the internal covariates consist of:
 T theta phase, i.e., the phase of the band-pass filtered (5 to 12 Hz) and Hilbert-transformed local field potential.
 E ensemble spike signal, defined as the z-scored sum of the spike counts of all neurons simultaneously recorded at the same tetrodes.

All covariates were interpolated at the centers of the time bins to achieve the same temporal resolution of the spike counts. By means of an optimization procedure over the entire population of neurons (see section Hyperparameters optimization) we chose to bin each covariate C in to N_C bins. Finally for each covariate C in each time bin t we built a binary state vector $\mathbf{X}^C(t)$ of length N_C, with entries $X_i^C(t) = 1$ if the animal behavioral state at time t fell in the i-th bin for the covariate C, while $X_i^C(t) = 0$ otherwise. Unless explicitly stated the OF and ST sessions were concatenated.

Poisson GLM

We adopted a Poisson GLM to explicitly model the stochastic response of each neuron to the covariates C. The choice of a Poisson random component as used in Hardcastle et al. (2017) is further motivated in subsection '*Poissonian spiking*'. The models $\mathcal{M}$ we consider, are combinations of behavioral and internal covariates $C = P, H, S, T, E$. For a given model $\mathcal{M}$ the probability of recording k spikes in time bin t of length dt is Poissonian:

$$Prob\big(k\big|\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}}\{\beta^C\}_{C\in\mathcal{M}}\big) = exp\big(-FR(\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}})dt\big)\frac{(FR(\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}})dt)^k}{k!} \qquad \text{(Equation 1)}$$

where $FR(\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}})$ is the expected firing rate in time bin t, and

$$FR(\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}}) = exp\left(\sum_{C\in\mathcal{M}}\sum_{i}\beta_i^C X^C(t)_i\right)\Big/ dt \qquad \text{(Equation 2)}$$

depends on the behavioral/internal state vectors $\mathbf{X}^C$ in time bin t and the vector of predictors β^C for all covariates $C\in\mathcal{M}$. The predictors are estimated by the learning procedure explained in section "Learning."

Poisson spiking

We selected the Poisson distribution to fit the stochasticity of the firing process after explorative analysis of our data and rigorous testing. First, we observed that the spiking data are not binary for time bin lengths $dt \geq 1$ ms. This indicated that the Bernoulli distribution may not be a suitable choice. Second, we verified that the inter-spike-interval (ISI) distribution is well fitted by an exponential up to 60 ms (the deviations at 1-2 ms can be attributed to the refractory period). The ISI distribution expected for Poisson firing is indeed exponential.

In addition, we looked at the Fano Factors of the spike count N for each cell, defined as:

$$PFano\ Factor(N) = \frac{variance(N)}{mean(N)} \qquad \text{(Equation 3)}$$

This analysis indicated that the *variance*(N) versus *mean*(N) plots are well fitted by the line that bisects the first quadrant of the coordinate system for time bin lengths $dt \geq 40$ ms. The coefficient of determination of the fit $R - square$ is an indication of how much the spiking process deviates from a homogeneous Poisson process, for which a Fano Factor of 1 is expected in the long run.

Notice that the Poisson GLM in Equation 1 is not a homogeneous Poisson process (the firing rate of Equation 1 depends on the covariates which vary in time), it becomes homogeneous Poisson once the covariates in Equation 1 are fixed. We used the asymptotic distribution of the Fano Factor for the homogeneous Poisson (Eden and Kramer, 2010) to test how poissonian the spike counts are at fixed covariates. We run the test after conditioning the spike counts on the covariate position in a 30 times 30 grid superimposed to the recording box. The number of position bins was optimized for this test and bins with occupancy shorter than 250 ms were excluded by this analysis. The test with 5% significance level resulted in an average fraction of position bins rejected by cell of 0.0589 ± 0.0009. Finally, instead of conditioning on position, we conditioned on time, by testing the Fano Factor of the spike count in 500 ms time windows. The conditioning in time is motivated by slow varying covariates which enforce to the firing rate a different

timescale from the spiking timescale. This approach resulted into an even smaller average rejection rate, suggesting that position may not be the only covariate modulating the activity of this cells. The average fraction of 500 ms time windows in which the Poisson hypothesis has to be rejected with a 5% confidence is 0.0152 ± 0.0002.

Testing performance of GLM framework with synthetic cells

We simulated pure position, head-direction and speed tuned inhomogeneous Poisson spike trains and mixed combinations of the three to validate the performance our GLM framework.

Learning

In order to determine the selectivity of a neuron to the covariates $C = P, H, S, T, E$, given the recorded spike count train vector $\mathbf{k}$ and the vectors of covariates $\mathbf{X}^C(t)$ in each time bin t, we optimized the predictors β_i^C of the Poisson GLM in Equation 1 for each model $\mathcal{M}$ to minimize the cost function:

$$L\left[\{\beta^C\}_{C\in\mathcal{M}}\big|\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}}, \gamma, \mathbf{k}\right] = -\sum_t \ln\text{Prob}\left(k(t)\big|\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}}, \{\beta^C\}_{C\in\mathcal{M}}\right) + \frac{1}{2}\sum_{C\in\mathcal{M}}\gamma^C\sum_i\left(\beta_i^C - \beta_{i-1}^C\right)^2 \qquad \text{(Equation 4)}$$

such that the learned parameters are $\{\beta^C\}_{C\in\mathcal{M}} = \text{argmin}_{\{\beta\}_{C\in\mathcal{M}}}L[\{\beta\}_{C\in\mathcal{M}}|\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}}, \gamma, \mathbf{k}]$. The first term in the loss function of Equation 4 is the negative log-likelihood of the spike count train vector $\mathbf{k}$, while the second term is a penalty on large differences in parameters between nearest neighboring covariate bins and therefore enforces smoothness in the model predicted tuning curves. The smoothness hyperparameter γ^C controls the strength of the smoothness penalty for the covariate C and was optimized *a priori* on the entire population of neurons, as explained in section Hyperparameters optimization. The minimization of the function $L[\{\beta\}_{C\in\mathcal{M}}|\{\mathbf{X}^C(t)\}_{C\in\mathcal{M}}, \gamma, \mathbf{k}]$ in the parameter space was performed using the MATLAB *minunc* function. The learned parameters β were used for estimating model performance, constructing model predicted tuning curves (see section Tuning curves predicted by the model) and decoding.

In addition to the models containing the behavioral/internal covariates we also fitted a model with only the bias term ($FR(\beta_0) = \exp(\beta_0)/dt$), which we term the average firing rate model. Model performance for each cell was assessed in a 10-fold cross-validation setup: of the whole spike count train one tenth was held out as test set, while the rest constituted the training dataset, on which parameters were learned. The difference in log-likelihood of the test spike count train between the model $\mathcal{M}$ and the average firing rate Poisson model (also learned on the training data) was taken and then divided by the number of time bins in the test set. The cross-validation procedure was repeated for 10 non-overlapping test sets, resulting in 10 values of the log-likelihood increase over the average firing rate model per time bin $\{LLH_i(\mathcal{M})\}_{i=1}^{10}$. Their average across cross-validation folds $LLH(\mathcal{M}) = \frac{1}{10}\sum_i LLH_i(\mathcal{M})$ was regarded as the main indicator of model $\mathcal{M}$ performance.

Hyperparameters optimization

For optimizing the hyperparameters (Figure S4), number of bins N_C and smoothness hyperparameter γ^C for each covariate $C = P, H, S, T, E$, we computed the model performance LLH(C) on a grid in the hyperparameters space ($N_C = \{2,5,10,20,30,40,50,60,70\}$, $\gamma^C = \{0.08, 0.8, 0, 80, 800\}$), for each of the single covariate model $\mathcal{M} = C$ separately and for each cell independently. The values of LLH(C) at fixed hyperparameters were then averaged over all cells in the population, and the values of the hyperparameters at its maximum on the grid were taken as candidate optimal hyperparameters. To compare model performance at the maximum to its neighboring vertices on the grid in the hyperparameters space, we employed the one tailed Wilcoxon signed-rank test between the population vector of LLH(C) at the maximum and the same vector in its nearest neighbors on the grid. The values of the hyperparameters including the smallest number of covariate bins N_C, whose corresponding vector of average normalized log-likelihoods was not significantly different from the maximum (p > 0.05 in the Wilcoxon signed-rank test), were chosen as new candidate optimal hyperparameters. This procedure was then iterated until all neighboring models with smaller numbers of covariate bins were significantly different in model performance according to our test. The hyperparameters were then held fixed during learning on the entire population of neurons and for all models.

The optimal number of bins and smoothness hyperparameters are respectively $N^P = 30$ (along each edge of the squared enclosure, 900 in total) and $\gamma^P = 8$ for position, $N_H = 10$ and $\gamma^H = 800$ for head direction, $N_S = 10$ and $\gamma^S = 800$ for speed, $N_T = 10$ and $\gamma^T = 800$ for theta, $N_E = 20$ and $\gamma^E = 80$ for ensemble activity.

Model selection

Learning was performed for all 31 models $\mathcal{M}$, i.e., the single covariate models P, H, S, E, T, the two covariate models $PH, PS, HS, PE, PT...$, all the three, the four covariate models up to the five covariates model $PHSET$. The learning method is described in section "Learning." We adopted a forward model selection procedure (Hardcastle et al., 2017) aiming at selecting the model with the highest performance and the smallest number of covariates. As a starting point we used the average firing rate model (only bias). So, we considered the model with the largest performance LLH($\mathcal{M}$) among the single covariate models, e.g., $\mathcal{M}^1 = S$. The vector of log-likelihood increases over the average firing rate model $\{LLH\}_{i=1}^{10}$ across cross-validation folds for the selected single covariate model $\mathcal{M}^1$ was compared to the same vector for the average firing rate model (a vector of zeros), by means of the one-tailed Wilcoxon

signed rank test. If significantly different ($p < \alpha$ in the one tailed Wilcoxon signed rank test), then the single covariate model became the new candidate selected model, otherwise the average firing rate model was selected and the search in the model space interrupted. In the former case, the model with the largest LLH($\mathcal{M}$) among the two covariate models, including the covariate selected in the former step, e.g., $\mathcal{M} = PS, HS, TS, SE$, was considered, e.g., $\mathcal{M}^2 = PS$. In full analogy with the previous selection step, the vector of log-likelihood increases over the average firing rate model $\{LLH\}_{i=1}^{10}$ across cross-validation folds for the single covariate model $\mathcal{M}^1$ was compared to the same vector for the two covariate model $\mathcal{M}^2$, by means of the Wilcoxon signed-rank test. If the mean of the vector for $\mathcal{M}^2$, LLH($\mathcal{M}^2$), was significantly larger than the mean of the vector for $\mathcal{M}^1$, LLH($\mathcal{M}^1$), ($p < \alpha$ in the one tailed Wilcoxon signed-rank test) the model $\mathcal{M}^2$ was identified as candidate model. If not significant the search in the model space was interrupted and the single covariate model $\mathcal{M}^1$ became the selected model. The forward search was carried on in this fashion by including one covariate at each step to the candidate selected model, up to the selected model resulting in a not significantly different outcome in the Wilcoxon signed-rank test. We checked the sensitivity of the model selection procedure to the significance level α.

Splitting between external and brain internal variables

In the current study we were interested in how CA1 and SUB neurons compare in their mapping of navigational covariates P, H and S. The covariates E and T (ensemble firing and theta) were treated as brain internal states which helped fitting the model but did not directly explain the behavior of the animal. We therefore used them as auxiliary variables which helped to get the best possible fit, by on the one hand 'explaining away' spikes that otherwise would have been attributed to position, head direction or speed and on the other reducing the level of noise in the spike trains such that other spikes could be better attributed to the correct navigational covariate. For example, it might be conceivable that a neuron has a higher firing rate always when the the other neurons at the tetrodes (E) have a high firing rate because the neuron is not well enough isolated from the ensemble. In a scenario in which the ensemble responds to speed, including E in our model will attribute some spikes of that neuron to ensemble firing and less to the speed response. For an extended graph with all models displayed separately, see Figure S3G.

Relative single covariate contributions (rSCC)

The model selection procedure is highly sensitive to individual covariates. Even very small influences of covariates on the firing rates of the neurons are classified as significant. In subsequent analyses we therefore investigated how strongly each neuron's firing rate is determined by either an individual covariate. We measured the contribution of a single covariate C included in the selected model $\mathcal{M}^{\hat{a}}$ in explaining the firing rate of a neuron in terms of the difference in model performance between the selected model and the model resulting from removing the covariate C from the selected model, $\mathcal{M}^{\hat{a}}, C$. We therefore defined the rSCC (C) for covariate C$\in \mathcal{M}^{\hat{a}}$ as LLH($\mathcal{M}^{\hat{a}}$) $-$ LLH($\mathcal{M}^{\hat{a}}, C$) normalized by

$$\sqrt{\sum_{C \in \mathcal{M}^{\star}} LLH(\mathcal{M}^{\hat{a}}) - LLH(\mathcal{M}^{\hat{a}}, C)^2}$$

where LLH($\mathcal{M}$) is the normalized log-likelihood of the model $\mathcal{M}$ on the held-out data averaged across cross-validation folds of PHS, C refers to the two-covariate model that does not include the covariate C. Single covariate contributions of covariates not included in the selected model were set to zero.

Mixed selectivity score (MS-score)

The MS-score was defined as the product of rSCC of position, head direction and speed:

$$MS = \prod_{C \varepsilon \{P, H, S\}} rSCC(C)$$

The product of rSCC is a measure that maximizes at the point where different covariates contribute equally to the firing rate of the cell. As a measure of mixed selectivity, it was therefore preferred over the sum of rSCC or the maximal Euclidean distance from zero as both of these options would be similar for two neurons with either a large contribution of an individual covariate and small contributions the two other, or similarly large contributions of the three covariates.

Tuning curves predicted by the model

Model predicted tuning curves were constructed on the basis of the model including all covariates, $\mathcal{M} = PHSTE$, whose parameters β were learned as explained in section "Learning." A tuning curve is the average firing rate of a neuron as a function of relevant stimulus parameters, in our case covariate bins; for the model predicted tuning curves the average is taken over the distribution of the other covariates. In case of uniform sampling and assuming independence between covariates, the expected value of the firing rate in the i-th bin of covariate C^* for Poisson neurons as defined in Equation 4 is:

$$\mathbb{E}\left[FR\left(\{\mathbf{X}^C\}_{C \neq C^*}, X_i^{C^*} = 1\right)\right]_{\{x^c\}_{C \neq C^*} | X^{C^*} = 1} = \frac{\exp\left(\beta_i^{C^*}\right)}{dt} \prod_{C \neq C^*} \frac{1}{N_C} \sum_j \exp\left(\beta_j^C\right). \qquad \text{(Equation 5)}$$

Goodness of fit

In order to assess the goodness of fit of the GLM, in addition to LLH($\mathcal{M}$), the average of the explained deviance across cross validation folds exD($\mathcal{M}$) was estimated (Brown et al., 2002; Guisan and Zimmermann, 2000; Kraus et al., 2015; Pillow, 2009). The explained deviance of model $\mathcal{M}$ in cross-validation fold i is defined as $exD_i(\mathcal{M}) = \frac{LLH_i(\mathcal{M})}{LLH_i(S)}$ where S is the so-called saturated model, a maximum-likelihood Poisson GLM optimized on the test data (in practice the firing rate FR in Equation 2 is set equal to the spike count k(t) in each time bin t). Then $exD(\mathcal{M}) = \frac{1}{10}\sum_{i=1}^{10} exD_i(\mathcal{M})$ quantifies the fraction ascribable to model $\mathcal{M}$ of log-likelihood increase with respect to the average firing rate model of the best Poisson model.

Testing for significance in model selection between CA1 or SUB and between OF or ST

For Figures 4E and 5B we used a shuffling procedure to test whether the type of model selected for a proportion of neurons was significantly affected by the anatomical location of the neurons or the task in which the neurons were recorded. For the pair of conditions being compared (either CA1 versus SUB, or OF versus ST), each neuron was randomly assigned to one of the two conditions with equal probability, and for each model (e.g., PH or PHS) the absolute difference was calculated between the proportion of neurons selecting it in the two conditions. This process was repeated 10000 times, yielding a distribution of shuffled differences for each model. If the actual difference between the two groups was outside the 99.9[th] percentile of 10000 such shuffled differences, we considered the difference for a given model (e.g., PH or PHS) to be significant.

Decoding

We exploited the Poisson GLM framework defined in section Poisson GLM for decoding simultaneously position P, head direction H and speed S. We employed the model including all behavioral covariates *P*, *H*, *S* and the internal covariate theta T. Given the learned parameters $\beta^{\boldsymbol{C}}$ with $C = P, H, S \cup I$, the vectors of spike counts $\mathbf{k}_n$, for cells $n = 1, \ldots N$ and the internal observed covariates $\mathbf{X}^I(t)$, the decoded covariates at time t, $\{\mathbf{X}^C(t)\}_{C \in PHS}$, were chosen as those maximizing the objective function:

$$\Gamma(\{\mathbf{X}^C(t)\}_{C \in PHS} | \mathbf{k}_n\,_{n=1}^N, \mathbf{X}^I(t)) = \sum_{s=-3\sigma}^{3\sigma} K(s) \sum_{n=1}^N \log \mathrm{Prob}(k_n(t+s) | \{\mathbf{X}^C(t)\}_{C \in PHS}, \mathbf{X}^I(t), \{\beta^C\}_{C \in P,H,S \cup I}). \qquad \text{(Equation 6)}$$

In Equation 6 *K*(s) is a Gaussian kernel of zero mean and standard deviation σ, while $\mathrm{Prob}(k_n(t+s) | \{\mathbf{X}^C(t)\}_{C \in PHS}, \mathbf{X}^I(t), \{\beta^C\}_{C \in P,H,S \cup I})$ is the probability of observing the spike count k_n at time $t+s$, given the covariates $\{\mathbf{X}^C(t)\}_{C \in PHS}, \mathbf{X}^I(t)$ according to the Poisson model in Equation 2 with parameters. $\{\beta^C\}_{C \in P,H,S \cup I}$

Decoding is then maximizing the weighted log-likelihood of the spike counts of all cells in a time window of size $6\sigma + 1$ centered at the time point *t*. The parameter σ has been optimized for the entire dataset (both regions and both tasks).

Maximization of the objective function Γ was performed by grid search on the covariates lattice defined by the binning adopted for learning, where the number of bins has been optimized as explained in Section "Hyperparameters optimization."

Resampling of neurons

For decoding we sampled differently sized subpopulations (from 5 to 150 neurons) and maximized the likelihood of the spike counts in the space of all three behavioral covariates. Within the selected population, resampled neurons were randomly assembled from different sessions without consideration of whether they had been recorded in parallel or not in the first place. From each neuron we used the parameters Learned using the GLM model with covariates PSTH (sec. "Encoding") and generated Poisson spikes using the model predicted firing rates for Poisson spikes. All the combinations of position, head direction, speed and theta were realized during a recorded session of one of the animals in this study. To differentiate between OF and ST we resampled neurons recorded in the OF and in the ST separately and only employed ST recorded neurons for decoding the ST session and OF recorded neurons for decoding the OF session. This allowed us to decode the three covariates P, H and S simultaneously at every time bin t (while theta was given) by optimizing the likelihood of the weighted spike counts from Gaussian filtered time window $t - \Delta t$ to $t + \Delta t$ (where $2\Delta t$ is the size of the decoding window: 400 ms). The GLM used for encoding provides an estimate of the tuning curves in the 6-dimensional covariate space (2 dimensions for P and 1 for each H, S, E and T), assuming that each covariate contributes multiplicatively to the instantaneous firing rate. In absence of this (or alternative) assumptions on how covariates combine to contribute to the firing rate of a cell, a reliable estimate of the tuning curves in this 6 dimensional behavioral space would require much higher coverage.

Randomization of decoding

The randomized decoding distributions for Figure S6, are built by taking the difference between two means from decoding 20 times from a randomly selected subpopulation of neurons, while ensuring equal probability to select CA1 or SUB neurons. The randomization runs over 1000 iterations and it is assessed whether the difference between decoding error of CA1 and SUB population is larger than the 99th percentile of the randomized differences.

Statistical testing

The experimenters were not blind to subject treatments and no statistical methods were employed to predetermine sample size. Details regarding data distributions, statistical tests and sample size are presented in the main text, figures, and figure legends. Data presented as mean ± standard error of the mean (SEM), unless otherwise specified. All data was analyzed with custom-written MATLAB scripts (https://se.mathworks.com/). Nonparametric tests were used to analyze data violating normal distribution assumptions (Mann–Whitney U test), and all statistical tests performed were two-tailed with significance level set at $p < 0.05$ while P-values where always provided as numerical values.