

The Structure of Hippocampal CA1 Interactions Optimizes Spatial Coding across Experience

Michele Nardin,^{1,4}  Jozsef Csicsvari,¹ Gašper Tkáčik,¹ and Cristina Savin^{2,3}

¹Institute of Science and Technology Austria, Klosterneuburg AT-3400, Austria, ²Center for Neural Science, New York University, New York, New York 10003, ³Center for Data Science, New York University, New York, New York 10011, and ⁴Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia 20147

Although much is known about how single neurons in the hippocampus represent an animal's position, how circuit interactions contribute to spatial coding is less well understood. Using a novel statistical estimator and theoretical modeling, both developed in the framework of maximum entropy models, we reveal highly structured CA1 cell-cell interactions in male rats during open field exploration. The statistics of these interactions depend on whether the animal is in a familiar or novel environment. In both conditions the circuit interactions optimize the encoding of spatial information, but for regimes that differ in the informativeness of their spatial inputs. This structure facilitates linear decodability, making the information easy to read out by downstream circuits. Overall, our findings suggest that the efficient coding hypothesis is not only applicable to individual neuron properties in the sensory periphery, but also to neural interactions in the central brain.

Key words: hippocampus; maximum entropy models; network topology; neural coding; noise correlations

Significance Statement

Local circuit interactions play a key role in neural computation and are dynamically shaped by experience. However, measuring and assessing their effects during behavior remains a challenge. Here, we combine techniques from statistical physics and machine learning to develop new tools for determining the effects of local network interactions on neural population activity. This approach reveals highly structured local interactions between hippocampal neurons, which make the neural code more precise and easier to read out by downstream circuits, across different levels of experience. More generally, the novel combination of theory and data analysis in the framework of maximum entropy models enables traditional neural coding questions to be asked in naturalistic settings.

Introduction

The dual role of the hippocampal formation in memory (Scoville and Milner, 1957; Eichenbaum, 2000) and spatial navigation (O'Keefe and Dostrovsky, 1971; Morris et al., 1982) is reflected in two distinct views on hippocampal coding: the place field view (Moser and Paulsen, 2001; McNaughton et al., 2006) that reduces spatial information to tuning properties of individual neurons,

and the ensemble view (Harris et al., 2003; Harris, 2005) that focuses on subsets of units that are co-activated together as the substrate for memory (Hopfield, 1982). Nonetheless, recent results blur the line between the two (Meshulam et al., 2017; Stefanini et al., 2020; Hazon et al., 2022). Furthermore, it remains unclear how experience shapes the organization of CA1 population responses and what effects such changes may have on the encoding of spatial information.

Experience affects the properties of hippocampal cells in many ways. From a single-cell perspective, reliable position-dependent spiking is detectable after a few minutes in a novel environment (Wilson and McNaughton, 1993; Leutgeb et al., 2004), with a gradual reduction in overall firing, sharpening of place fields and sparsification of responses with increasing familiarity (Karlsson and Frank, 2008). From a pairwise interactions' perspective, total correlations among pairs of place cells are stronger in novel environments (Gava et al., 2021), although less is known about their structure and computational role. CA1 inhibition, the primary mechanism for circuit interactions in this region, is weaker in novel environments as compared with familiar ones (Nitz and McNaughton, 2004; Sheffield et al., 2017; Arriaga and

Received Feb. 9, 2023; revised Sep. 11, 2023; accepted Sep. 14, 2023.

Author contributions: G.T. and C.S. designed research; M.N. and C.S. performed research; J.C. contributed unpublished reagents/analytic tools; M.N. and C.S. analyzed data; M.N. and C.S. wrote the first draft of the paper; M.N., J.C., G.T., and C.S. edited the paper; M.N., G.T., and C.S. wrote the paper.

M.N. was supported by the European Union Horizon 2020 Grant 665385. J.C. was supported by the European Research Council Consolidator Grant 281511. G.T. was supported by the Austrian Science Fund (FWF) Grant P34015. C.S. was supported by an Institute of Science and Technology fellow award and by the National Science Foundation (NSF) Award No. 1922658. We thank Peter Baracska, Karola Kaefer, and Hugo Malagon-Vina for the acquisition of the data. We also thank Federico Stella, Wiktor Mlynarski, Dori Derdikman, Colin Bredenberg, Roman Huszar, Heloisa Chiossi, Lorenzo Posani, and Mohamady El-Gaby for comments on an earlier version of the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Cristina Savin at csavin@nyu.edu.

<https://doi.org/10.1523/JNEUROSCI.0194-23.2023>

Copyright © 2023 the authors

Han, 2019). In parallel, the main afferents to CA1 are initially noisier (Cohen et al., 2017; Pedrosa and Clopath, 2020) and have weaker spatial tuning, which improves with familiarity (Leutgeb et al., 2004; Hafting et al., 2005; Barry et al., 2012). Moreover, CA1 needs both inputs for detailed spatial representation (Brun et al., 2002, 2008). Taken together, these observations suggest that the CA1 circuit is in a different dynamic regime during novelty, with distinct local circuit interactions.

Circuit interactions shape both spatial tuning and noise correlations. In general, noise correlations are considered to be an obstacle to optimal information coding, especially in sensory areas (Moreno-Bote et al., 2014; Kanitscheider et al., 2015; Rumyantsev et al., 2020). Nevertheless, there are scenarios where they can improve coding quality at the population level (Tkacik et al., 2010; da Silveira and Berry, 2014; Valente et al., 2021; Panzeri et al., 2022), which might be relevant for the hippocampus. Traditionally, the importance of noise correlation on population coding has been assessed by comparing the information content of the population responses with an alternative code where the neurons preserve their tuning but whose responses are otherwise independent (Stefanini et al., 2020; Hazon et al., 2022). While attractive as a statistical analysis, such manipulations may not represent a neural code that is biologically realizable. Since mechanistically circuit interactions shape both single-cell and pairwise statistics (Sheffield and Dombeck, 2019), they cannot be decoupled in the real brain. In this paper, we explore the consequences of such a constraint for the hippocampal neural code.

Measuring CA1 noise correlations is fraught with technical difficulties (Mathis et al., 2013; Tocker et al., 2015; Dunn et al., 2015; Posani et al., 2017). Free behavior in 2D environments renders traditional approaches for estimating noise correlations inapplicable, because of the lack of stimulus repeats. Existing approaches struggle to account for complex dependencies as they are either too data inefficient (Mathis et al., 2013) or make strong assumptions about the functional form of the interactions (Dunn et al., 2015). Moreover, oscillations (Colgin, 2013) act as global sources of co-modulation obscuring the fine structure of pairwise neural co-variability. The key challenge is to partition neural covariability into an explainable component, driven by position, oscillations, and other global signals, and unexplained, or “excess” correlations, which reflect local interactions.

Here, we develop a new statistical test for detecting excess correlations in freely moving animals, and explore their significance for the encoding of spatial information in CA1. Our method allows us to robustly detect network interactions by comparing hippocampal responses against a maximum entropy null model (Savin and Tkačik, 2017) that optimally captures the cells’ place preference and population synchrony (Engel et al., 2001). When applied to CA1 tetrode recordings from rats during open field exploration, our analysis detects structured excess correlations, more prominently during novelty. These highly structured excess correlations optimize spatial information and facilitate its downstream readout in both the familiar and novel environment, with differences reflecting the different selectivity of spatial inputs in the two environments. Taken together, our results suggest that CA1 local circuitry readjusts to changes in its inputs so as to improve population-level stimulus representation, in line with efficient coding predictions (Tkacik et al., 2010).

Materials and Methods

Experimental procedures

Datasets and subjects

We analyzed data from two previously published datasets (Kaefer et al., 2019; Stella et al., 2019). All procedures involving experimental animals were conducted in accordance with Austrian animal law (Austrian federal law for experiments with live animals) under a project license approved by the Austrian Federal Science Ministry. Four adult male Long–Evans rats (Janvier) were used for the experiments by Stella et al. (2019). We further analyzed two wild-type littermate control animals from Kaefer et al. (2019), generated by breeding two DISC1 heterozygous Sprague Dawley rats. Rats were housed individually in standard rodent cages (56X40X26 cm) in a temperature and humidity-controlled animal room. All rats were maintained on a 12/12 h light/dark cycle and all testing performed during the light phase. Food and water were available *ad libitum* before the recording procedures and bodyweight at the time of surgery was 300–375 g.

Surgery

The first four animals (Stella et al., 2019) were implanted with microdrives housing 32 (2×16) independently movable tetrodes targeting the dorsal CA1 region of the hippocampus bilaterally. Each tetrode was fabricated out of four 10- μ m tungsten wires (H-Formvar insulation with Butyral bond coat California Fine Wire Company) that were twisted and then heated to bind them into a single bundle. The tips of the tetrodes were then gold-plated to reduce the impedance to 200–400 k Ω . During surgery, the animal was under deep anesthesia using isoflurane (0.5–3% MAC), oxygen (1–2 l/min), and an initial injection of buprenorphine (0.1 mg/kg). Two rectangular craniotomies were drilled at relative to bregma (centered at AP = -3.2 ; ML = ± 1.6), the dura mater removed and the electrode bundles implanted into the superficial layers of the neocortex, after which both the exposed cortex and the electrode shanks were sealed with paraffin wax. Five to six anchoring screws were fixed on to the skull and two ground screws (M1.4) were positioned above the cerebellum. After removal of the dura, the tetrodes were initially implanted at a depth of 1–1.5 mm relative to the brain surface. Finally, the micro-drive was anchored to the skull and screws with dental cement (Refobacin Bone Cement R). Two hours before the end of the surgery the animal was given the analgesic Metacam (5 mg/kg). After a one-week recovery period, tetrodes were gradually moved into the dorsal CA1 cell layer (stratum pyramidale).

The last two animals (Kaefer et al., 2019) were implanted with microdrives housing 16 independently movable tetrodes targeting the right dorsal CA1 region of the hippocampus. Each tetrode was fabricated out of four 12- μ m tungsten wires (California Fine Wire Company) that were twisted and then heated to bind into a single bundle. The tips of the tetrodes were gold-plated to reduce the impedance to 300–450 k Ω . During surgery, the animal was under deep anesthesia using isoflurane (0.5–3%), oxygen (1–2 l/min), and an initial injection of buprenorphine (0.1 mg/kg). A rectangular craniotomy was drilled at -3.4 to -5 mm AP and -1.6 to -3.6 mm ML relative to bregma. Five to six anchoring screws were fixed onto the skull and two ground screws were positioned above the cerebellum. After removal of the dura, the tetrodes were initially implanted at a depth of 1–1.5 mm relative to the brain surface. Finally, the microdrive was anchored to the skull and screws with dental cement. Two hours before the end of surgery the analgesic Metacam (5 mg/kg) was given. After a one-week recovery period, tetrodes were gradually moved into the dorsal CA1 cell layer.

After completion of the experiments, the rats were deeply anesthetized and perfused through the heart with 0.9% saline solution followed by a 4% buffered formalin phosphate solution for the histologic verification of the electrode tracks.

Behavioral procedures

Each animal was handled and familiarized with the recording room and with the general procedures of data acquisition. For the first four animals (Stella et al., 2019), 4–5 d before the start of recording, animals were familiarized at least 30 min with a circular open-field environment (diameter = 120 cm). On the recording day, the animal

underwent a behavioral protocol in the following order: exploration of the familiar circular open-field environment (40 min), sleep/rest in rest box (diameter = 26 cm, 50 min). Directly after this rest session the animals also explored a novel environment for an additional 40 min and rested after for 50 min. The novel environment recordings were performed in the same recording room but in an enclosure of a different geometric shape but similar size (e.g., a square environment of 100-cm width). The wall of both the familiar and novel environment enclosures was 30 cm in height, which limited the ability of the animal to access distal room cues. In addition, in two animals a 50-min sleep/rest session was performed before the familiar exploration.

For the last two animals (Kaefer et al., 2019), 2–3 d before the start of recording, animals were familiarized with a circular open-field environment (diameter = 80 cm). On the recording day, the animal underwent a behavioral protocol in the following order: 10-min resting in a bin located next to the open-field environment, exploration of the familiar open-field environment (20 min), sleep/rest in the familiar open-field environment (20 min), exploration of a novel open-field environment (20 min), sleep/rest in the novel open-field environment (20 min). While the familiar environment was kept constant, the novel environment differed on every recording day. The novel open-field arenas differed in their floor and wall linings, and shapes. The recordings for the familiar and novel conditions were performed in the same recording room.

During open-field exploration sessions, food pellets (MLab rodent tablet 12 mg, TestDiet) were scattered on the floor to encourage foraging and therefore good coverage of the environment.

Data acquisition

A headstage with 64 or 128 channels (4×32 or 2×32 channels, Axona Ltd) was used to preamplify the extracellular electric signals from the tetrodes. Wide-band (0.4 Hz to 5 kHz) recordings were taken and the amplified local field potential and multiple-unit activity were continuously digitized at 24 kHz using a 128-channel (resp. 64-channels) data acquisition system (Axona Ltd). A small array of three light-emitting diode clusters mounted on the preamplifier headstage was used to track the location of the animal via an overhead video camera. The animal's location was constantly monitored throughout the daily experiment. The data were analyzed offline.

Data processing

Spike sorting

The spike detection and sorting procedures were performed as previously described (O'Neill et al., 2006). Action potentials were extracted by first computing power in the 800–9000 Hz range within a sliding window (12.8 ms). Action potentials with a power >5 SD from the baseline mean were selected and spike features were then extracted by using principal components analyses. The detected action potentials were segregated into putative multiple single units by using automatic clustering software (<http://klustakwik.sourceforge.net/>). These clusters were manually refined by a graphical cluster cutting program. Only units with clear refractory periods in their autocorrelation and well-defined cluster boundaries were used for further analysis. We further confirmed the quality of cluster separation by calculating the Mahalanobis distance between each pair of clusters (Harris et al., 2000). Afterwards, we also applied several other clustering quality measures and selected only cells which passed stringent measures. In particular we implemented: isolation distance and l-ratio (Schmitzer-Torbert et al., 2005), Inter Spike interval (ISI) violations (Hill et al., 2011) and contamination rate. We employed the code available on GitHub: <https://github.com/cortex-lab/sortingQuality>. The criteria for the cells to be considered for analysis were: (1) isolation distance >10 th percentile, (2) ISI violations <0.5 , and (3) contamination rate <90 th percentile. Periods of waking spatial exploration, immobility, and sleep were clustered together and the stability of the isolated clusters was examined by visual inspection of the extracted features of the clusters over time. Putative pyramidal cells and putative interneurons in the CA1 region were discriminated by their autocorrelations, firing rate, and waveforms, as previously described (Csicsvari et al., 1999).

Data inclusion criteria

We set a minimum average firing rate of >0.25 Hz for each cell, across both familiar and novel environments. Tetrodes were 0.4 mm apart, guaranteeing that no two tetrodes would pick up the same cell (Henze et al., 2000). Nonetheless, we monitored for that possibility by measuring the cross-correlogram of cells on different tetrodes, and found no suspicious pairs of cells that could be duplicates. The final dataset consisted of 294 putative excitatory and 128 putative inhibitory cells across 6 animals. Considering only pairs of units recorded on different tetrodes, the dataset includes a total of 9511 excitatory-excitatory (EE) pairs, 7848 excitatory-inhibitory (EI), and 1612 inhibitory-inhibitory (II) pairs.

Spiking data were binned in 25.6-ms time windows, reflecting the sampling rate for positional information. We excluded bins where (1) the animal was static (speed <3 cm/s); (2) sharp-wave ripple oscillatory activity was high, i.e., periods with power in the band 150–250 Hz in the top fifth percentile (Csicsvari et al., 2000; O'Neill et al., 2006); (3) theta oscillatory activity was particularly low, with power in the band 5–15 Hz in the lowest fifth percentile; it is known that hippocampal theta oscillations support encoding of an animal's position during spatial navigation and reduces overall synchrony of population (Buzsáki and Moser, 2013; Mizuseki and Buzsáki, 2014).

Null model of population responses and detection of excess correlations

We construct a null model for population responses ($y_1, \dots, y_N$) that takes into account the position of the animal, $\mathbf{s}$ and the population syn-

chrony, $k = \sum_i y_i$, but is otherwise maximally variable. We use this

model to generate a large ensemble of surrogate datasets, that match the data with respect to tuning but without additional noise correlations. Using these surrogates allow us to estimate an empirical distribution of (total) pairwise correlations under the null model, which we then compare to data.

Under the assumption that spike counts have mean $\lambda(\mathbf{s}, k)$ with Poisson noise, the distribution of the joint neural responses under the null model factorizes as:

$$p_{ind}(\mathbf{y} | \mathbf{s}, K) = \prod_i \text{Poisson}(y_i | \lambda_i(\mathbf{s}, k)). \quad (1)$$

The Poisson distribution has maximal entropy over a very large class of distributions on $\mathcal{N}_0$ (Harremoës, 2001; Johnson, 2007). One important caveat is that the population synchrony depends on the neural responses themselves, which introduces the additional constraint that $k = \sum_i y_i$ for each of these surrogate draws, something that we enforce

by rejection sampling (Press et al., 1992). The only remaining step is to estimate the tuning function of each cell, $\lambda_i(\mathbf{s}, k)$, which we achieve using a nonparametric approach based on Gaussian Process (GP; Rasmussen, 2003) priors.

Tuning function estimation

Here, we briefly describe the key steps of the approach and refer the reader to (Savin and Tkacik, 2016) for further details. For an introduction to Gaussian Processes inference, we refer the reader to (Rasmussen, 2003).

The data are given as T input pairs, $\mathcal{D} = \{\mathbf{x}_t, \mathbf{y}_t\}_{t=1,2,\dots,T}$, where $\mathbf{x}_t$ denotes the input variables, defined on a 3D lattice for the $2d$ -position of the animal in the environment and population synchrony, defined as $k = \sum_{n=1}^N y_t^{(n)}$; $\mathbf{y}_t$ denotes spike counts of N neurons in the t th time bin ($dt = 25.6$ ms).

Neural activity of each single neuron is modeled as an inhomogeneous Poisson process with firing rate dependent on input variables, $\lambda(\mathbf{x}_t)$. We use a Gaussian Process (GP) before specify the assumption that the neuron's tuning is a smooth function of the inputs, with an exponential link function, $f = \log \lambda$, $f \sim \mathcal{GP}(\mu, c)$, with mean function $\mu(\cdot)$ and

covariance function $c(\cdot, \cdot)$. In particular, we use a product of squared exponential (SE) kernels for the covariance function:

$$c(\mathbf{x}, \mathbf{x}') = \prod_{d=1}^3 c_d(x_d, x'_d) = \prod_{d=1}^3 \rho_d \exp(x_d - x'_d)^2 / 2\sigma_d^2, \quad (2)$$

This allows the prior covariance matrix C to be decomposed as a Kronecker product: $C = C_1 \otimes C_2 \otimes C_3$, dramatically increasing the efficiency of the fitting procedure (Flaxman et al., 2015). We denote the set of parameters that uniquely determine f as $\theta = \{\mu, \rho, \sigma\}$.

We choose input points $\mathbf{x}_* = (\mathbf{s}, k)$ that corresponded to the binned 2D location $\mathbf{s}$ of the animal (5-cm bins) and binned population synchrony k [10 equally weighted bins, each containing 10% of the data, i.e., the bin edges correspond to the (0th, 10th..., 100th) percentiles]. Our aim is to find a distribution over expected log-firing rates at each input point $\mathbf{x}_*$; we denote these by $f_* = f(\mathbf{x}_*)$.

Given a set of parameters $\hat{\theta} = \{\hat{\mu}, \hat{\rho}, \hat{\sigma}\}$, and a set of input values $\mathbf{x}_*$ (as defined above), we infer the predictive distribution over log-firing rates given data and parameters as $p(f_* | \mathcal{D}, \mathbf{x}_*, \hat{\theta})$. This distribution can be computed by marginalizing over $\mathbf{f}$, which denotes a collection of function values of f evaluated at a given set of points $X = \{\mathbf{x}_1, \dots, \mathbf{x}_I\}$, i.e., $\mathbf{f} = f(X) \sim \mathcal{N}(\hat{\mu}, \hat{C}) \in \mathbf{R}^I$, where $\hat{C}_{ij} = c(\mathbf{x}_i, \mathbf{x}_j | \hat{\rho}, \hat{\sigma})$.

$$p(f_* | \mathcal{D}, \mathbf{x}_*, \hat{\theta}) = \int p(f_* | \mathcal{D}, \mathbf{x}_*, \hat{\theta}, \mathbf{f}) p(\mathbf{f} | \mathcal{D}, \hat{\theta}) d\mathbf{f} \quad (3)$$

This distribution is intractable, but can be approximated by using a Laplace approximation for $p(\mathbf{f} | \mathcal{D}, \hat{\theta})$ so that ultimately $p(f_* | \mathcal{D}, \mathbf{x}_*, \hat{\theta}) \approx \mathcal{N}(\mu_{f_*}, \sigma_{f_*}^2)$.

The parameters θ used in the previous paragraph are fitted for each cell by maximizing the marginal likelihood of the data given parameters: $\hat{\theta} = \arg\max_{\theta} \log P_{\theta}(\{\mathbf{y}_1, \dots, \mathbf{y}_T\}, \{\mathbf{x}_1, \dots, \mathbf{x}_T\})$. This step and the previous one are repeated until the convergence of data log-likelihood. We refer the reader to refs (Flaxman et al., 2015; Savin and Tkacik, 2016) for detailed derivations and computational implementation of the learning algorithm.

Finally, thanks to the exponential link function, the inferred firing rate of an individual input point $\lambda(\mathbf{x}_*) = \exp(f_*)$ is log-normally distributed, whose mean and variance can be computed as:

$$\mathbb{E}(\lambda(\mathbf{x}_*)) = \exp(\mu_{f_*} + \sigma_{f_*}^2 / 2) \quad (4)$$

and

$$\text{Var}(\lambda(\mathbf{x}_*)) = \exp(\sigma_{f_*}^2 - 1) \exp(2\mu_{f_*} + \sigma_{f_*}^2). \quad (5)$$

This completes our probabilistic inference of single-cells tuning functions. Computing $\text{Var}(\lambda(\mathbf{x}_*))$ allows us to propagate the uncertainty of tuning function estimation throughout all our analyses, as explained in the following subsection.

Generating surrogate data

At each moment in time, given the position $\mathbf{s}$ and population synchrony k , the GP tuning estimate provides a distribution over possible firing rates for cell i , $\lambda_i(\mathbf{s}, k)$, as a log normal distribution, $\log \lambda_i \sim \mathcal{N}(\mu_{f_i}, \sigma_{f_i}^2)$. This captures uncertainty about the tuning of the cell, given the data. Our method allows us to propagate this uncertainty throughout all the subsequent analyses by sampling firing rates from this distribution (formally, by treating neural responses as arising from a doubly stochastic point model). We generate surrogate spike counts in two steps. First, we sample the mean firing from this $p(\lambda_i | \mathbf{s}, K)$ distribution. Second, for each λ_i sample, we draw the corresponding spike count from $\text{Poisson}(\lambda_i)$. Applying this procedure for all cells and all time points generates a surrogate dataset from the unconstrained null model. We enforce

the constraint $\sum_i y_i = k$ by discarding and redrawing samples that

do not satisfy it. In rare cases (<2% of data), it was not possible to replicate the desired k statistic, i.e., achieving the desired k required >500 re-samplings. Such time bins were excluded from subsequent analysis (both for real data and all surrogates). We generate a total of 1000 surrogate datasets.

Inference of excess correlations

We use the pairwise correlations between neural responses as the test statistic and compare it to the distribution of pairwise correlations expected under the null model that assumes that the firing rate of cells is only driven by the stimulus and the synchrony of the population, without further pairwise interactions.

Given the Pearson correlation coefficient between the activities of cells i and j computed on real data, c_{ij} , and c_{ij}^{γ} , the same quantity computed on a surrogate dataset $\{\mathbf{y}_{1:t}^{\gamma}\}$ for $\gamma = 1, 2, \dots, 1000$. We define the quantity we refer to as “excess correlations” as:

$$w_{ij} = \frac{c_{ij} - \langle c_{ij}^{\gamma} \rangle}{\sigma(c_{ij}^{\gamma})}, \quad (6)$$

where $\langle \cdot \rangle$ denotes the sample average and σ the sample standard deviation of c_{ij}^{γ} . Assuming that the c_{ij}^{γ} distribution is normal, this quantity is closely related to confidence bounds, and p -values (via the error function). An excess correlation is deemed significant if $|w_{ij}| > 4.5$, which corresponds to a p -value threshold of $p = 0.05$ with a Bonferroni correction for >7500 multiple comparisons. This threshold was chosen so as to counteract the effect of multiple comparisons, and make sure that our results are not influenced by spurious correlations.

Validation procedures

To validate our method, we construct an artificial dataset with known interactions, by sampling from a coupled stimulus dependent MaxEnt model. We consider $N = 50$ neurons and binary activations $\mathbf{y} = (y_1, \dots, y_N)^T$ for any given time window. The distribution of responses $\mathbf{y}$ given a location-stimulus $\mathbf{s}$ and synchrony level k is

$$p(\mathbf{y} | \mathbf{s}, k) \propto \exp \left(h \sum_i f_i(\mathbf{s}) y_i + \sum_{i>j} W_{ij} y_i y_j - h_0(k) \sum_i y_i \right), \quad (7)$$

where $\mathbf{s} \in \{\mathbf{s}_1, \dots, \mathbf{s}_K\}$ is a spatial position chosen from a set of discrete locations uniformly spaced in the environment, and k -dependent parameter $h_0(k)$ is chosen to ensure that $\sum_i y_i = k$ on average. The spatial

input tuning to each cell, $f_i = f_i(\mathbf{s})$, is modeled as a 2D. Gaussian bump with continuous boundary conditions, mean randomly drawn from a uniform on $[0, 1] \times [0, 1]$ and fixed covariance 0.1I. The parameter h is used to vary the input strength, as explained in the following subsections. For validation in Figure 1 and Extended Data Figure 1-1, we use $h = 2, 3, 4$. We try to match the general statistics of the data as closely as possible. In particular, we match the true time-dependent occupancy, s_t , observed in a 20-min exploration session, and the corresponding time-dependent synchrony observed in the same session, k_t , by sampling one population activity vector (after adequate burn-in time) at each time point $\mathbf{y}(t) \sim p(\mathbf{y} | \mathbf{s}_t, k_t)$ using Gibbs sampling (Geman and Geman, 1984).

Given this artificial dataset, we analyze it with the same processing pipeline that we use for the neural recordings and compare the estimated interactions w_{ij} with the ground truth couplings W_{ij} , which are randomly and independently drawn from $\mathcal{N}(0, 1)$. Furthermore, we generate data with the same constraints but without any interactions. We assess the ability of our statistical test to detect true interactions using the receiver operating characteristic (ROC), and estimate false positive rates for our statistical test.

Additional validation for instability in tuning

We further validated our method on synthetic data simulated using unstable tuning inputs. We model this phenomenon by spatially tuned

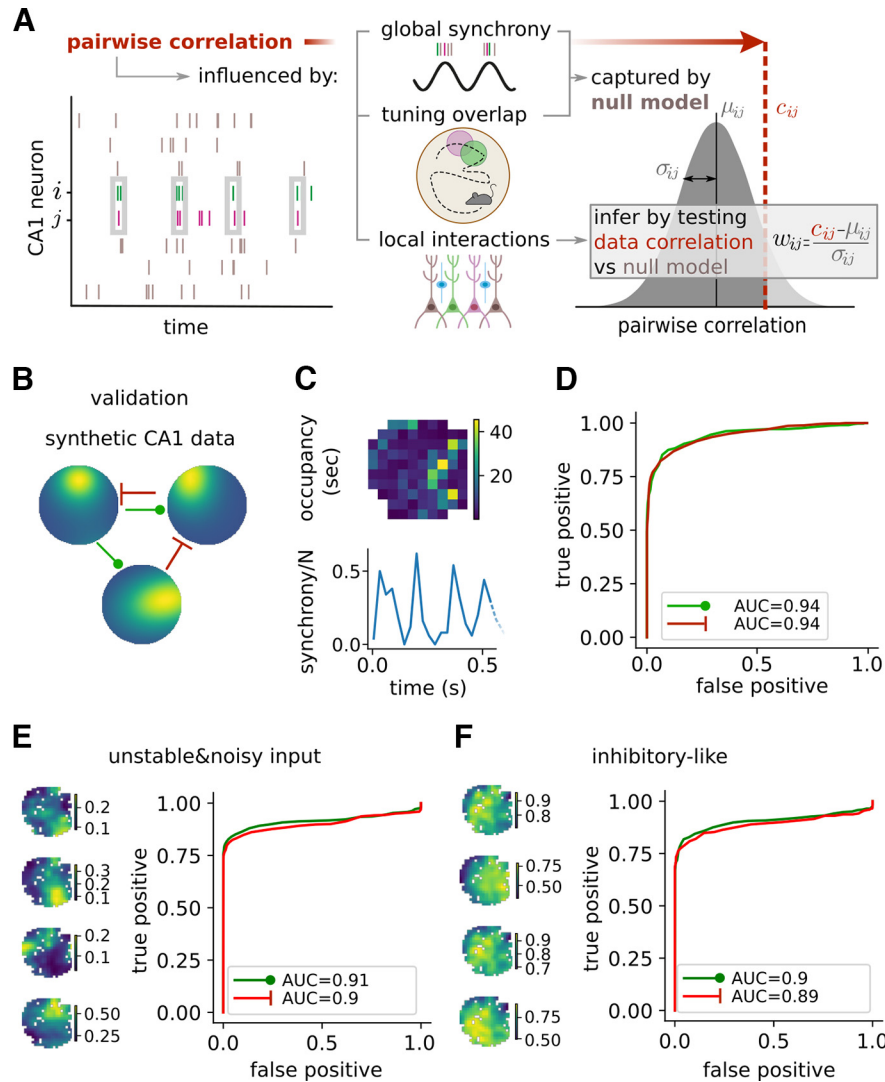


Figure 1. Detecting excess correlations among hippocampal CA1 cells during naturalistic behavior. **A**, Method schematic. A null model for population responses takes into account the inferred place field tuning of each cell and the moment-to-moment global synchrony, but is otherwise maximally unstructured. For each cell pair, this model predicts a null distribution for (total) pairwise correlation (gray distribution), which is compared with the correlation estimate from data (dashed red line). The normalized discrepancy between the data correlation c_{ij} and the null model expectation μ_{ij} for a pair of neurons (i, j) is referred to as “excess correlation,” w_{ij} , and serves as a proxy for cell-cell interaction. **B**, Method validation on synthetic data. Detection accuracy is assessed using simulated data with known positive (green) or negative (red) interactions. **C**, Synthetic data matches a real 20-min exploration session with respect to spatial occupancy (top) and observed synchrony indices (bottom represented here as proportion of simultaneously active cells in a ~ 500 -ms example raster). **D**, Receiver operating characteristic (ROC) shows the probability of correctly detecting positive (green) and negative (red) interactions for different detection thresholds. **E**, Novel-like scenario: unstable and noisy multipeak tuning. Input strength was drawn random from $\mathcal{N}(2, 1)$. Number of peaks chosen as Poisson with average = 3. At each time point, peak weights were re-drawn from a Dirichlet distribution with parameter $\alpha = 2$. The average population firing was always kept to 20% of the population. Left, Example inferred tuning functions. Right, ROC. **F**, Inhibitory-like scenario: high-firing and weak spatial tuning, noisy multipeak tuning. Number of peaks chosen as Poisson with average = 5. At each time point, peak weights were re-drawn from Dirichlet distribution with parameter $\alpha = 2$. The average population firing was kept to 50% of the population. Left, Example inferred tuning functions. Right, ROC.

inputs that are fluctuating over time, but with stable network interactions, which we want to infer. Specifically, the synthetic data were generated using Equation 7 while introducing heterogeneity in input tuning and input strength. In particular, we considered a distribution of input strengths, different for each cell ($h_i \sim \mathcal{N}(2, 1)$), mixture-of-Gaussians input tuning functions with a variable number of components $N_i \sim \text{Pois}(3)$ and components weights w_{ij} that were re-drawn for each iteration of Gibbs sampling from a Dirichlet distribution over N_i -dim positive real vectors with parameter $\alpha = 2$. This setup ensures both significant heterogeneity across cells and a substantial degree of tuning instability.

Single-cell tuning characterization

To describe the tuning properties of single cells, we employed several standard measures: (1) gain, defined as peak firing rate over mean; (2)

sparsity, $\langle \lambda_x \rangle_x^2 / \langle \lambda_x^2 \rangle_x$, where λ_x denotes the average firing at location x , is a measure of how compact the firing field is relative to the recording apparatus (Jung et al., 1994); (3) spatial information, $\langle \frac{\lambda_x}{\lambda} \log \frac{\lambda_x}{\lambda} \rangle_x$, where $\lambda = \langle \lambda_x \rangle_x$, is the leading term of the MI between average spiking and discretized occupancy for small time windows (Skaggs et al., 1992; Souza et al., 2018).

Decoding of spatial position from data

We partitioned the environment in equally spaced 2D bins with bin side length of 20 cm. This choice was because of the fact that, to properly estimate the average co-activation of cells one needs many samples and a finer subdivision of the environment made this task extremely difficult. We refer to each bin with an index $s \in \mathcal{S}$. The data were randomly split in two sets, 75% for training and 25% for testing. Given the training

data, we estimated the average cell-by-cell activation ($\widehat{PV}_s$) and the covariance of the neurons activity ($\widehat{COV}_s$) for each spatial bin s . At test time, decoding was implemented by comparing the cosine similarity of these prototypes to the cell-by-cell average activation (denoted by population vector PV) and the cell-cell covariance (COV) for each nonoverlapping 10 consecutive 25.6-ms time bins and picking the decoded bin as that with the highest similarity in terms of Pearson correlation. For the analyses on larger networks with 100, 200, 500 cells, we also used a Bayesian maximum a posteriori (MAP) decoder for which we inferred the conditional probability of spiking of cell i given location s , $P(y_i|s)$, on $4 \cdot 10^5$ samples. We then tested the decoder on additional 10^5 samples. Cells were assumed to be conditionally independent of each other, hence $P(\mathbf{y}|s) = \prod_i P(y_i|s)$. For simplicity, we used an uniform prior over possible positions. Hence the decoded position given a population vector $\mathbf{y}$ is given by:

$$\hat{s}_{MAP} = \arg \max_s P(s|\mathbf{y}) = \arg \max_s P(\mathbf{y}|s). \quad (8)$$

Linear separability of pairs of stimuli

To investigate the linear separability of population responses to different locations, we randomly selected 500 times two distinct locations in the environment and selected all the 250-ms population responses in a 10-cm surrounding of the two positions. We then found the best hyperplane that separated the two sets of responses by using a soft-margin linear SVM with hinge loss, and reported the training error. We also computed the principal components of the population responses to both locations together, and reported the variance explained by the first PC.

Models of hippocampal population responses with adjustable network structure

Stimulus-dependent MaxEnt model

In order to explore the effects of the noise correlation structure on the coding properties of the hippocampal system, we employed a statistical model of the collective behavior of a population of place cells that allowed us to vary the couplings among cells while keeping fixed the output firing rate. A similar, stimulus-dependent maxent model was introduced in (Granot-Atedgi et al., 2013). A stimulus-independent version was used in (Meshulam et al., 2017) to prove that correlation patterns in CA1 hippocampus are not because of place encoding only, but also to internal structure and pairwise interactions. Our model includes spatially-selective inputs with adjustable strength, $h \in \mathbb{R}$, and noise correlations modeled as a matrix $\mathbf{W}$ describing the strength of interaction between cell pairs. Additionally, we constrained average population firing rates to be the same for each possible choice of h and $\mathbf{W}$, as a way of implementing metabolic resource constraints.

More specifically, consider N neurons with binary activations $\mathbf{y} = (y_1, \dots, y_N)^T$. The distribution of responses $\mathbf{y}$ given a location-stimulus s we considered is

$$p(\mathbf{y}|s) \propto \exp\left(h \sum_i f_i(s) y_i + \sum_{i>j} W_{ij} y_i y_j - h_0 \sum_i y_i\right), \quad (9)$$

where $s \in \{s_1, \dots, s_K\}$ is a spatial position chosen from a set of discrete locations uniformly spaced in the environment (the unit square, $[0, 1] \times [0, 1]$). The spatial input tuning to each cell, $f_i = f_i(s)$, is modeled as a 2D Gaussian bump with continuous boundary conditions, mean randomly drawn from a uniform on $[0, 1] \times [0, 1]$ and fixed covariance $0.1\mathbb{I}$. The parameter h_0 allows us to fix the average population firing rate to 20% of the population size, and is found by grid optimization. Once the input tuning f_i is fixed for each cell, we select the connections W_{ij} for each cell pair by sampling from the data-inferred excess correlations of cell pairs with similar input tuning, and then scaling according to the results found during method validation (Extended Data Fig. 1-1G). We did so separately for familiar and for novel environments.

Inference of input quality from data

We fixed the appropriate parameter h , separately for familiar-like and novel-like network interactions, by matching average marginal statistics of firing rate maps obtained from the model to the data. We used three measures: single-cell spatial information, sparsity, and gain. Separately for familiar and novel environments, we randomly chose 50 tuning functions and selected pairwise connections (rescaled from data, Extended Data Fig. 1-1G) depending on tuning similarity. We then simulated population activity according to real occupancy (as in validation) for $h \in \{0.5, 1, \dots, 5\}$, computed firing rate maps and their statistics. We selected the h value so as to minimize the sum of (absolute) deviations of the averages of the three measures.

Optimization of connections for fixed input and fixed firing rate

Given $h, \{f_i(\cdot)\}$, we optimize the connections $\mathbf{W}$ so as to maximize the mutual information (Shannon, 1948) between population activity and spatial position, $MI(\mathbf{y}; s) = \sum_{\mathbf{y}, s} p(\mathbf{y}|s)p(s) \log \frac{p(\mathbf{y}|s)}{p(\mathbf{y})}$, via sequential least

squares programming (SLSQP; Virtanen et al., 2020). We further constrain the population average firing to 20% of the neural population, and each W_{ij} is restricted to lie in $[-1, 1]$. Both reflect biological resource constraints on the optimal solution.

Most simulations use $N = 10$ neurons, which allows the mutual information to be computed in closed form (by enumerating all possible patterns). Reported estimates are obtained by averaging across 1000 randomly initialized networks (different $f_i(\cdot)$ centers, and initial conditions for the optimization). To ensure that our results generalize to large networks, we also performed limited numerical simulations for $N = 20$ (only for $h = 2$ and $h = 4$, averaging over 10 networks).

Comparison of data-like interactions with optimization results

To measure how “optimal” real data excess correlations are, we randomly sampled 10-cell networks 1000 times as explained above. We measured the MI between population activity and spatial position using data-like interactions (sampled from data excess correlations as in Fig. 3), and compared the to the MI measured on the same network with null interactions; we report average ratio $\pm$ SD in Figure 4B.

Optimal coding for large networks

The exact computation of the mutual information $MI(\mathbf{y}; s)$ is very resource intensive and only applicable to small networks ($N \leq 20$). To investigate the effects of noise correlations at larger scales we need to rely on efficient approximations. The mutual information between population binary responses $\mathbf{y}$ and location-stimulus s can be written as

$$MI(\mathbf{y}; s) = \sum_{\mathbf{y}, s} p(s|\mathbf{y})p(\mathbf{y}) \log p(s|\mathbf{y}) - \sum_{\mathbf{y}, s} p(s|\mathbf{y})p(\mathbf{y}) \log p(s), \quad (10)$$

$$= H(s) - H(s|\mathbf{y}),$$

where H denotes (conditional) entropy. Assuming that $p(s)$ is a uniform distribution over stimuli, we have $H(s) = 2\log B$, where B is the number of bins used to discretize each dimension of the 2D environment. We generally use $B = 16$. The challenge is to compute $H(s|\mathbf{y})$. For a given $\mathbf{y}$, denote with $\hat{h}(\mathbf{y}) := -\sum_s p(s|\mathbf{y}) \log p(s|\mathbf{y})$. Then we have:

$$\begin{aligned} H(s|\mathbf{y}) &= -\sum_{\mathbf{y}, s} p(s|\mathbf{y})p(\mathbf{y}) \log p(s|\mathbf{y}) \\ &= \sum_{\mathbf{y}} p(\mathbf{y}) \hat{h}(\mathbf{y}) \\ &= \sum_s p(s) \sum_{\mathbf{y}} p(\mathbf{y}|s) \hat{h}(\mathbf{y}) \end{aligned} \quad (11)$$

We used the last expression and estimated $H(s|\mathbf{y})$ by drawing 10^6 samples from $p(\mathbf{y}|s)$ for each stimulus s using Gibbs sampling (Geman and Geman, 1984). We reported the estimated average

across stimuli and confidence intervals in the figures. The quantity $\hat{h}(\mathbf{y}) = -\sum_s p(s|\mathbf{y}) \log p(s|\mathbf{y})$ is the entropy of the posterior distribution on stimuli given a certain binary vector. The main obstacle to computing $\hat{h}$ is that, for each stimulus s , we need to know the proportionality constant $Z_s = \sum_{\mathbf{y}} p(\mathbf{y}|s)$ (i.e., the partition function), that makes the probability (Eq. 9) sum up to 1. We computed Z_s exhaustively for $N \leq 20$ by enumerating all the possible binary vectors. For $N \geq 20$ we estimated it using a simple Monte Carlo method by randomly drawing 10^9 independent N -dim binary samples for each stimulus, and then regularizing by applying a mild 2D Gaussian smoothing ($\sigma = 0.5$ bins) on the log-transformed Z_s among neighboring stimuli.

Total versus noise entropy in full versus factorized models

The mutual information can be understood as the difference between total entropy ($H(\mathbf{y})$) and conditional, or noise, entropy ($H(\mathbf{y}|s)$): $MI(\mathbf{y}; s) = H(\mathbf{y}) - H(\mathbf{y}|s)$. The introduction of non-zero second order interactions in Equation 9 will affect differently the two terms. In general, introducing correlations will diminish the total entropy $H(\mathbf{y})$ (which implies that the total possible information is reduced) but can also affect the noise entropy $H(\mathbf{y}|s)$ (whose reduction implies a reduction in uncertainty of the response given stimulus).

“Topology” model simulations

We aimed at characterizing the influence of higher order structure on the coding of the network. We used the same model as in Equation 9 with 50 place cells, but allowed connections to be either $-J$, 0 or $+J$, where $J \in [0, 1]$ is the connection strength. We employed three different strategies to select the units to connect, as described in the main text, based on their tuning similarity. We kept fixed the number of positive ($+J$) and negative ($-J$) couplings to 6% and 3% respectively. For each choice of tuning, connectivity rule and strength J we used the parameter h_0 to enforce the population average firing to be 20% of the population size.

Network analysis

Graph theoretical measures

All the measures were conducted using the library NetworkX (release 2.4) in Python 3.7. We considered unweighted and nondirected graphs where each cell was a vertex and an edge connected each cell pair that had a significant excess correlation ($|w_{ij}| > 4.5$). A graph $G = (V, E)$ formally consists of a set of vertices V and a set of edges E between them. An edge e_{ij} connects vertex v_i with vertex v_j . The neighborhood for a vertex v_i is defined as its immediately connected neighbors: $N_i = \{v_j : e_{ij} \in E \vee e_{ji} \in E\}$ and its size will be denoted by $k_i = |N_i|$.

We measured:

1. Clustering coefficient: This measure represents the average clustering coefficient of each node, which is defined as the fraction of detected triangles that include that node as a vertex. Formally, the local clustering coefficient c_i for a vertex v_i is given by the proportion of links between the vertices within its neighborhood divided by the number of links that could possibly exist between them, hence measuring how close its neighborhood is to forming a clique. If a vertex v_i has k_i neighbors, $\frac{k_i(k_i - 1)}{2}$ edges could exist among the vertices within the neighborhood. Thus, the local clustering coefficient for vertex v_i can be defined as

$$c_i = \frac{2|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}$$

and the average clustering coefficient as

$$c_G = \frac{1}{n} \sum_{v_i \in V} c_i$$

1. Average shortest path length: This measure can be computed only if the graph is connected. If not, we computed this measure on the largest connected subgraph.

$$a_G = \sum_{u,v \in V} \frac{d(u, v)}{n(n-1)}$$

where u, v are distinct vertices, $d(u, v)$ is the shortest path length between u, v and n is the size of the graph G .

Analysis of local network motifs

We tested for the over-expression of particular interaction patterns by counting the number of triangles (i.e., three all-to-all interacting cells) composed by three inhibitory cells, two inhibitory and one excitatory, one inhibitory and two excitatory or three excitatory cells. We tested these counts against the counts from the same networks with shuffled edges. We employed an edge-shuffling procedure that preserved both the total number of edges and the number of incident edges per node, separately for the EE, EI, and II subnetworks (i.e., an edge connecting two excitatory cells could be exchanged only with another edge connecting two excitatory edges, etc.). To do this, we randomly selected two edges of each subnetwork, say AB and CD . If $A \neq C \neq D$ and $B \neq C \neq D$ we removed the two edges and inserted the “swapped” ones, AC and BD . We repeated this procedure 100 times for each subnetwork to yield one shuffled network. We repeated this procedure 1000 times, which gave us a null distribution to test the original counts against. In Extended Data Figure 2-4, we reported the counts of each pattern, separately for familiar and novel environments, normalized against our null distribution.

Betti numbers

We computed the Betti numbers of the clique-complex induced by the graphs. These are distinct from the graphs Betti numbers (Giusti et al., 2015). A clique in a graph is an all-to-all connected set of vertices. The clique complex $X(G)$ of an undirected graph G is an abstract simplicial complex (that is, a family of finite sets closed under the operation of taking subsets), formed by the sets of vertices in the cliques of G . Intuitively, the clique-topology can be characterized by counting arrangements of cliques which bound holes. Formally, the dimensions of the homology groups $H_m(X(G), \mathbb{Z}_2)$ yield the Betti numbers b_m (Giusti et al., 2015). Given our low connectivity (9%), b_m was almost always zero for $m \geq 2$. On the other side, b_0 simply counts the number of connected components, so in our analysis we focused on b_1 . This is the number of cycles, or holes, that are bounded by 1D cliques. Graphically, these are four edges that form a square, or five edges that form a pentagon, etc. Notice that three edges that form a triangle do not count toward b_1 , because they represent a 2D clique (i.e., three vertices that are all-to-all connected). This is why a higher clustering coefficient (i.e., more triangles) implies a lower b_1 .

Scripts and code availability

The code used for the detection of interactions, together with the code used for analyses throughout the paper, is available at https://github.com/Savin-Lab-Code/CA1_network_interactions.

Results

Measuring excess correlations without stimulus repeats

To investigate cell-cell interactions between CA1 neurons and their role in spatial information coding, we devised a procedure to infer excess correlations from simultaneous tetrode recordings of hundreds of isolated units in dorsal hippocampus of freely behaving rats. Our approach starts by constructing a null model for population responses that exactly accounts for the measured spatial selectivity of each recorded neuron as well as for the moment-to-moment measured global neural synchrony, but is otherwise maximally unstructured (Fig. 1A; Extended Data Fig. 1-1A-E). Given this null model, surrogate neural rasters can be sampled (Fig. 1A, left) and their statistics directly compared with those of real data (Savin and Tkačik, 2017; Extended Data Fig. 1-1D-F). Unique to our solution is a rigorous propagation of uncertainty about the inferred single-cell place tuning (see

Materials and Methods) as a way of taking into account data limitations on the estimation of neural tuning. For every cell pair, the model predicts the expected distribution of pairwise correlations against which the measured total correlation for that pair can be tested for significance (Fig. 1A, right); we report “excess correlation” w as the (normalized) amount of total correlation that is not explained by the null model. Since our approach explicitly discounts for correlations arising from overlapping place fields and sources of global modulation (e.g., because of locking to the underlying brain oscillations or influence of behavioral covariates such as running velocity; Extended Data Fig. 1-1A,B), it differs from previous attempts to use total correlations to probe the intrinsic network mechanisms (Gava et al., 2021). Although it does not necessarily detect individual synaptic connections (Das and Fiete, 2020), this is in keeping with traditional definitions of noise correlations, which estimate correlations conditioned on the stimulus. It isolates the fine structure of statistical dependencies beyond global shared co-variability and potentially offers a clearer window into the net effects of local circuit interactions on neural activity.

We validated our detection method by constructing synthetic datasets of spiking CA1 neurons whose responses were modulated by the position of an artificial agent and by an assumed network of interactions (Fig. 1B, left; see Materials and Methods). We ensured that the synthetic data matched overall firing rates, the moment by moment synchrony and the highly irregular occupancy observed in a real 20-min exploration session (Fig. 1B, middle). Excess correlations identified by our method strongly overlap with the ground truth, as measured by the area under the receiver operating characteristic (Fig. 1B, right). The inferred excess correlations were also well aligned with the ground truth interaction strengths (Extended Data Fig. 1-1G). We did not find any tendency of cells that are more (or less) similarly tuned to show higher (or lower) inferred w_{ij} s (linear regression, $r = 0.023$, $p = 0.24$). Inference quality did not change when using weaker or stronger inputs (Extended Data Fig. 1-1H). Introducing instability and additional noise in the tuning of single cells (Fig. 1E) or high-firing inhibitory-like cells (Fig. 1F) did not alter our ability to detect network interactions. Overall, our estimator outperformed standard noise correlation detection (Extended Data Fig. 1-1I) and proved robust across parameter variations that capture documented familiar versus novel environment differences in CA1 activity.

We next analyzed CA1 tetrode recordings of six wild-type rats exploring familiar and novel 2D environments separated by a short period of rest (Fig. 2A,D; Kaefer et al., 2019; Stella et al., 2019). The two environments were of similar size (Extended Data Fig. 2-1A) and different in color and geometry. Putative units were filtered by using several clustering quality measures, based on the Mahalanobis distance and the interspike interval (Harris et al., 2000; Schmitzer-Torbert et al., 2005; Hill et al., 2011), to ensure that they were well isolated (Extended Data Fig. 2-1F; see Materials and Methods). Our subsequent analyses only included cells active in both environments (with > 0.25 spike/s to ensure reasonable statistical power). As expected, spatial occupancy, behavior, and firing fields differ across the two environments (Extended Data Fig. 2-1A–C). Nonetheless, our null-model did not differ across familiarity levels in terms of model fit quality, as estimated by marginal log-likelihood (Extended Data Fig. 2-1D). Further, the differences in behavior between conditions are unlikely to bias excess correlations, as downsampling the familiar data to match novel occupancy and speed exploration statistics did not affect the estimates (Extended Data Fig. 2-1E). Considering only

pairs of cells that were simultaneously recorded on different tetrodes, our final dataset includes a total of 9511 excitatory-excitatory (EE), 7848 excitatory-inhibitory (EI), and 1612 inhibitory-inhibitory (II) pairs (see Materials and Methods for animal-by-animal details). We set the threshold to declare an excess correlation significant at $|w| > 4.5$ (corresponding to a strict Bonferroni correction for the largest number of pairs analyzed, see Materials and Methods). We detected both positive and negative excess correlations among cell pairs (Fig. 2B,C). Interestingly, cell pairs with negative excess correlations can have positive total correlations (Fig. 2C), illustrating the idea that total correlations are not necessarily a good predictor of local circuit interactions.

Excess correlation networks in familiar and novel environments

What is the structure of the inferred excess correlations? We found a generally sparse inferred connectivity pattern in the excitatory-excitatory (EE) subnetwork, with $\sim 6\%$ (624/9511) of pairs in the familiar and $\sim 9\%$ (860/9511) of pairs in the novel environment showing significant excess correlations (Fig. 2E, left). This connectivity level is substantially above the Bonferroni corrected detection threshold, but also implies that the null model accounts for most of the observed total correlation structure, which makes excess correlations nontrivial to estimate by simpler means. The fraction of interactions is larger among excitatory-inhibitory (EI) cell pairs (Fig. 2E, middle; 1056/7848 significant pairs in familiar, 1143/7848 significant pairs in novel environment), where, as expected, negative interactions dominate; the fraction is highest at $\sim 30\%$ among positive interactions in the inhibitory-inhibitory (II) subnetwork (Fig. 2E, right; 1008/1612 significant pairs in familiar, 952/1612 significant pairs in novel environment).

We next focused on interaction changes induced by a switch from the familiar to the novel environment (Fig. 2D). We observed an increase in EE significant excess correlations, possibly mediated by decreased inhibition during novelty. Consistently, we found putative inhibitory cells to be less synchronous and slightly less active in novel environments (Extended Data Fig. 2-1H), in line with previous reports (Nitz and McNaughton, 2004), while excitatory neurons were more synchronous but did not differ in terms of their average firing rates (Extended Data Fig. 2-1G). Total correlations were generally stronger in novel environments (Extended Data Fig. 2-1I; Cheng and Frank, 2008; Gava et al., 2021). Consistent with the view that CA1 circuit modifications during spatial learning originate in altered spike transmission among connected excitatory and inhibitory neurons (Dupret et al., 2013; McKenzie et al., 2021), we found an increase in positive EI significant excess correlations, while their negative counterpart remained unchanged. This increase could not be attributed to increased reliability of monosynaptic EI connections (Extended Data Fig. 2-2E), especially since cell pairs on the same tetrode were excluded from this analysis (Csicsvari et al., 1998). We did not observe significant changes in the number of II interactions.

Are interaction changes explained by fine time scale oscillatory modulation of hippocampal firing? We measured the similarity of cell-pair theta-locking histograms, and compared it against the inferred w_{ij} s. We found no significant relation in the EE subnetwork and no significant changes across familiarity levels (Extended Data Fig. 2-2A,B, left). The correlation between theta-phase locking similarity and excess correlation was small but significant for EI cell pairs (Extended Data Fig. 2-2A,B, middle), and was strongest within the II subnetwork (Extended Data

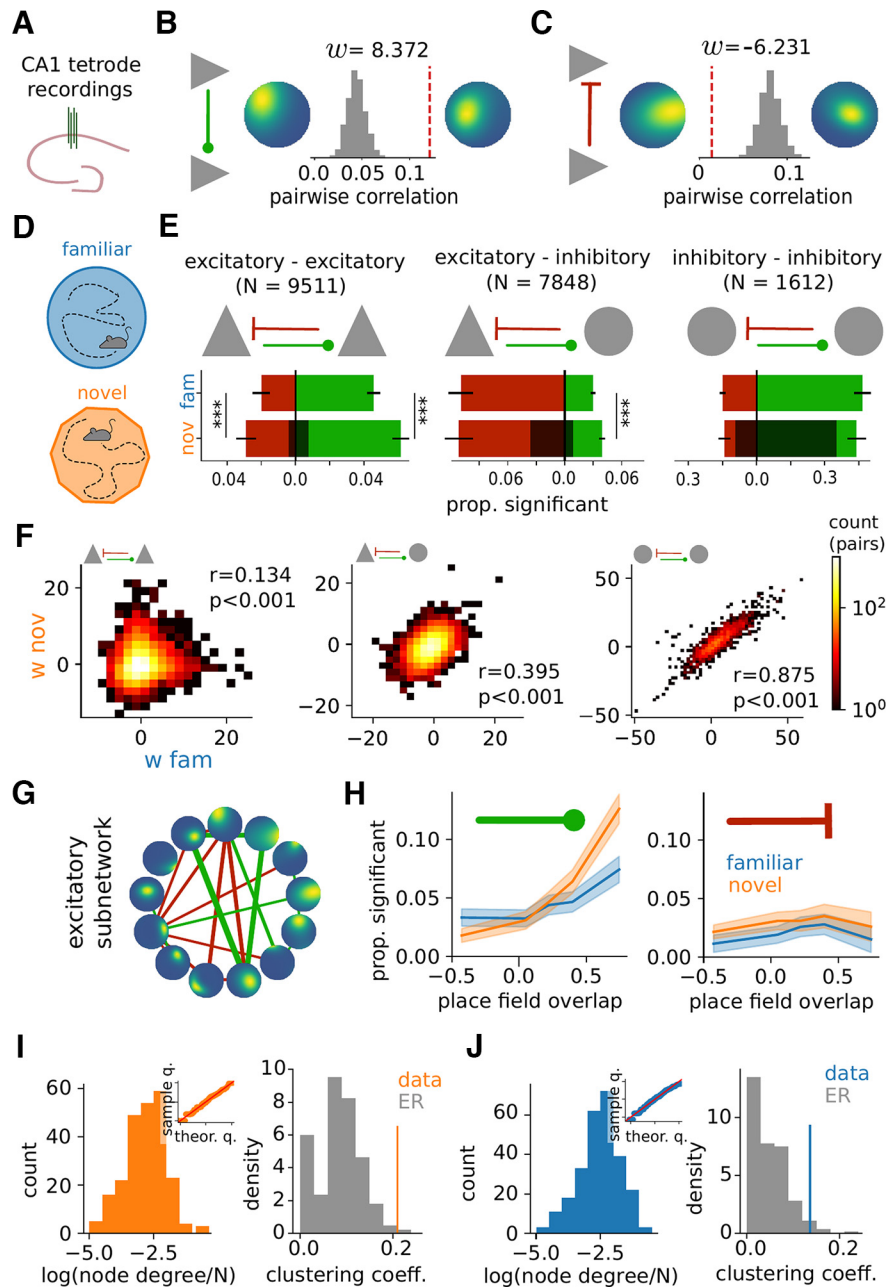


Figure 2. Excess correlations across familiar and novel environments. **A**, Neural activity was recorded using tetrodes implanted in the dorsal CA1 area of the hippocampus. **B**, **C**, Example pairs of pyramidal cells simultaneously recorded during free exploration of a familiar environment with significant positive (**B**) and negative (**C**) excess correlation w (gray histogram, distribution of correlation coefficients derived from the null model; red dashed line, measured raw pairwise correlation). **D**, Experimental paradigm. Axi animals explored a familiar environment, then rested in a sleep box (rest data not used), after which they explored a novel environment (20–40 min for each condition). Each animal contributed to one experimental session. **E**, Proportion of significant excess correlations for different cell types (triangle, pyramidal cell; circle, putative interneuron), positive (green) and negative (red) excess correlations, for both the familiar (top row, blue) and the novel (bottom row, orange) environment (stars, significant difference under binomial test at $p < 0.001$). Error bars indicate SEM across $N = 6$ animals. Shaded regions mark the fraction of interactions detected in the familiar environment that remain in the novel environment. **F**, Paired comparison (colormap, binned pair count) between excess correlations w_{ij} detected in familiar versus novel environment for each cell-pair within EE (left), EI (middle), and II (right) subnetworks. **G**, Example of an estimated excitatory subnetwork. Circles show the place field of each neuron, with edges showing significant excess correlations (green, positive; red, negative excess correlations); line thickness corresponds to interaction strength. **H**, Left, Proportion of significant excess correlations in the excitatory subnetwork increases with place field overlap, measured as the Pearson correlation of 2D firing rate maps, for positive interactions (blue, familiar environment; orange, novel environment; shaded area, 99th percentile confidence interval for the mean). Right, Analogous plot for negative excess correlations. **I**, **J**, Left, Distribution of log node-degree of E cells normalized by the total number of E cells in each session, for the novel (**I**) and familiar (**J**) environment. Inset, Quantile-quantile plot comparing this distribution to the normal expectation. Right, Excitatory subnetwork has a significantly higher clustering coefficient (orange/blue line, data) compared with the expected distribution for an Erdos-Renyi (ER) network with a matched connection density. See Extended Data Figures 2-1, 2-2, 2-3, 2-4 and Extended Data Tables 2-1, 2-2 for more information.

Fig. 2-2A,B, right). Nonetheless, this relation was not significantly different across experience (Extended Data Fig. 2-2B). Finally, since excitatory cells further exhibit theta-phase precession, which might be itself experience dependent (Cheng and Frank, 2008), co-processing cells could lead

to spurious excess correlations. We controlled for this effect by introducing different degrees of spike jitter, and repeating the analysis. Despite this disruption of individual cell precession, we found that inferred interactions could still be inferred reliably (Extended Data Fig. 2-2C,D). Overall, the

changes in excess correlation structure across experience do not seem to be trivially explained in terms of known oscillatory structure in neural responses.

How conserved are significant excess correlation pairs across environments? The largest overlap was found in the II subnetwork, where 77.5% of detected interactions were preserved in both familiar and novel. EI excess correlations showed less overlap (31.1%), while the overlap was weakest (16.8%) in the EE subnetwork. All reported overlaps were statistically significant under a permutation test (1000 random shuffles of cell labels; $p < 10^{-3}$ for all subnetworks). Significance was confirmed by comparing the Jaccard similarity of the adjacency matrices of familiar and novel subnetworks against the null distributions constructed from random graphs with the same numbers of vertices and edges (1000 Erdos–Renyi graphs; $p < 10^{-3}$ for II and EI subnetworks, $p = 0.009$ for EE). Moreover, the similarity of excess correlations across the two environments extended beyond the binary presence or absence of significant edges. Figure 2F compares the strength of excess correlations, w , in the familiar versus novel environment for EE, EI, and II cell pairs. For all subnetworks, w are significantly correlated across the two environments, with the reported correlation strength related to the network overlap (Fig. 2E). Taken together, these findings corroborate the idea that cell-cell interactions do not reconfigure randomly across environments, likely because of underlying network constraints (English et al., 2017; Levy et al., 2023; Huszár et al., 2022).

Because spatial information is encoded predominantly by pyramidal cells (Skaggs et al., 1992; Frank et al., 2001), we analyzed the EE subnetwork in further detail (Fig. 2G). Our key statistical observation is shown in Figure 2H: the proportion of significant excess correlations increases nonlinearly with place field overlap for positive interactions, and is roughly constant for negative interactions. In the novel environment, the proportion of positive excess correlations increases ~ 3 -fold over the observed range of place field overlap. In the familiar environment, the modulation with place field overlap is less pronounced, possibly indicating a shift toward a more decorrelated representation of space (Karlsson and Frank, 2008). These effects are consistent on an animal-by-animal basis (Extended Data Fig. 2-3C) and persist even after excluding the first 5 min of exploration, known for presenting unstable place fields (Wilson and McNaughton, 1993; Extended Data Fig. 2-3A–C).

We further characterized the topology of familiar and novel EE excess correlation networks. The number of interactions that a neuron engages in (its node degree) appears to be log-normally distributed in both environments, with clustering coefficients that are significantly higher than expected from matched independently randomly connected (Erdos–Renyi) graphs (Fig. 2I,J). This effect was more pronounced during novelty (Extended Data Fig. 2-4A).

Accordingly, interacting excitatory triplets were overrepresented, more strongly so in the novel environment (Extended Data Fig. 2-4C). Finally, we found a linear relationship between the log-number of nodes and the shortest path length (i.e., the minimal distance between randomly chosen node pairs, Extended Data Fig. 2-4B), which is a defining feature of small-world networks (Watts and Strogatz, 1998). These results add to the array of documented small-world topology seen across brain scales (Bassett and Sporns, 2017), from anatomic connectivity (Perin et al., 2011), to human imaging functional connectivity (Bassett et al., 2008).

Effects of network interactions on CA1 single-cell and population spatial coding

To explore how the network structure affects spatial information encoding at the population level, we constructed a statistical model of interacting excitatory cells responding to spatial inputs (Fig. 3A). Our model, a version of pairwise-coupled, stimulus-driven maximum entropy distribution over binary spiking units (see Materials and Methods; Granot-Atedgi et al., 2013) allows us to vary cell-cell excess correlations (to study the effect of network topology and interaction strength) as well as the strength of the spatial inputs (to study the effect of differences between the novel vs familiar environment) while maintaining a fixed average firing rate for the population. Mechanistically, the model describes pairwise network interactions as the source of the excess correlations measured in the data and spatially selective inputs as the primary driver of cell selectivity, with tuning further refined by the local circuit interactions. This marks an important departure from classic theoretical analyses of the effects of neural correlations on neural coding. In standard approaches, marginal statistics as fixed and the second order interactions are manipulated independently of tuning. In contrast, here we acknowledge that circuit interactions necessarily also shape marginal statistics and thus the two cannot be controlled independently in the brain. We used our model to assess spatial coding at the population level as a function of the structure of network interactions.

We contrasted spatial coding in two networks which were identical in their input tuning and average firing rates, but different in the pattern of network interactions. Interactions in the “structured” network followed the relationship between place field overlap and excess correlation w observed in real data (note that spatial inputs and place fields overlap are strongly correlated in our statistical regime, see Fig. 3C and Extended Data Fig. 3-1A); interactions in the “random” network were drawn from the same data-derived distribution for w , but independently of the overlap in (input) tuning (Fig. 3A). Finally, in keeping with traditional neural coding analyses on the role of correlations, we also considered the “no interactions” case in the absence of any covariability of neural responses (Extended Data Fig. 3-2). For each case, we further simulated the effects of familiar versus novel environment by adjusting the strength of the feed-forward spatial input, in analogy to previous experimental results (Brun et al., 2002, 2008; Leutgeb et al., 2004; Hafting et al., 2005; Barry et al., 2012; Cohen et al., 2017). In the simplest version of the model, higher input strength corresponds to higher spatial informativeness of input tuning, which is why we refer to this parameter as “input quality.” We adjusted the input quality to best resemble, on average, familiar versus novel data in terms of various marginal statistics [spatial information, measuring single-cell informativeness (Skaggs et al., 1992; Souza et al., 2018); sparsity, capturing size, and sharpness of place fields (Jung et al., 1994); and peak-over-mean firing values; see Materials and Methods and Extended Data Fig. 3-1B; see Extended Data Fig. 3-1C,D for distributional model-data match].

We quantified the coding performance of our networks by estimating the mutual information between population activity and spatial position and by estimating the average decoding error. For tractability, we simulated populations of 50 place cells (Fig. 3). As expected, higher input quality in the familiar environment leads to overall higher information values (Fig. 3B) and lower decoder error (Extended Data Fig. 3-2B). Less trivial are the effects of network connectivity: in both environments,

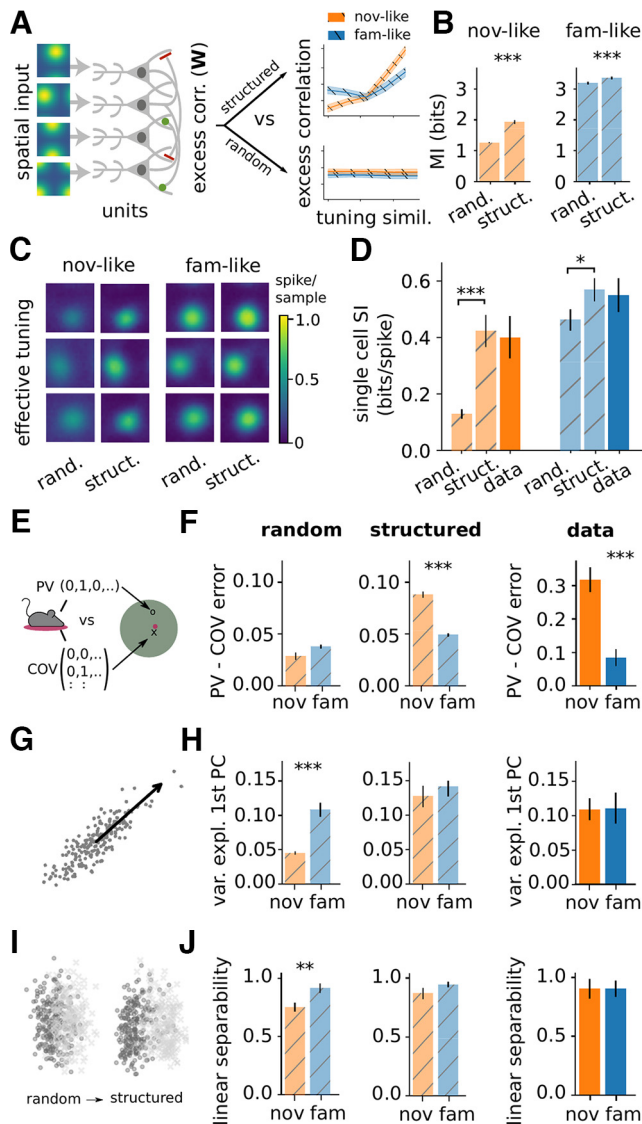


Figure 3. Effects of network interactions on spatial encoding. **A**, A schematic of the circuit model with variable excess correlations and fixed population firing rate (see Materials and Methods). Two connectivity patterns are compared: “structured” (mimicking the inferred excess correlation vs tuning similarity relationship) versus “random.” **B**, Estimated spatial information (MI; error bar, 99th percentile CI for the mean) using structured and random interactions, in the novel-like and familiar-like scenario (see text). Structured interactions significantly increase the spatial information ($***p < 0.001$ or $**p < 0.01$ under a nonparametric Mann–Whitney U test). **C**, Effective single-cell tuning for random versus structured models in the novel-like (orange; left) and familiar-like (blue; right) scenarios. **D**, Average single place cell spatial information (I_{spike}) for random versus structured interactions (dashed bars) and data (solid bar) in the novel-like (orange; left) and familiar-like (blue; right) scenarios. **E**, The performance of a simple population vector (PV) decoder is compared with the performance of a decoder which takes into account the co-variability of neurons (COV). **F**, Improvement in decoding performance, measured as error of PV decoder minus error of COV decoder, evaluated on $4 \cdot 10^4$ samples for random (left), data-like structure (center), and data (right). Random 50% of data were used for train/test. Decoding error measured in spatial bins. Error bars and significance tests as in **B**. **G**, Fraction of variance explained by the first principal component for responses to random pairs of stimuli, used to measure the effect of structured interactions on population responses to different stimuli. **H**, Fraction of variance explained by the first principal component of population vectors for 10^3 random pairs of locations in the maze for random (left) versus data-like structure (center) versus data (right). The fraction is unchanged between the novel and familiar environments on structured network and on real data but differs significantly on the random network (error bars and significance tests as in **B**). **I**, Linear separability of responses to pairs of stimuli. The schematic depicts the 2D PCA projection of the responses to a random pair of stimuli (light gray vs dark gray) when using random (left) and structured (right) excess correlations. **J**, Linear SVM

structured (data-like) interactions significantly outperform random ones, with larger improvements seen in the novel environment. This suggests that network interactions among hippocampal cells adjust to maintain a high-fidelity spatial representation even when they receive lower quality, noisy inputs. Improved decodability offered by structured interactions, probed with different decoders, persists in larger networks with 100, 200, and 500 cells (PV decoder, Mann–Whitney U test, all $ps < 0.001$; MAP decoder, Mann–Whitney U test, all $ps < 0.001$). Qualitatively similar results are observed with more realistic inputs, where single cells have irregular spatial input tuning (Epsztein et al., 2011) and variable input strengths as considered in Extended Data Figure 3-1C (Mann–Whitney U test, $p < 0.001$). These population benefits are also reflected in a general improvement of single-cell spatial tuning (Fig. 3C), and single neuron spatial information (Fig. 3D; Souza et al., 2018).

Do the structured interactions better predict other population-level aspects of the real hippocampal code? We assessed the importance of pairwise (co-firing) statistics for the decoding performance, highlighted by previous work (Stefanini et al., 2020; Fig. 3E). For the random network, the decoding performance improvement with co-firing statistics relative to population-vector decoding is small and comparable in the novel versus familiar environment. In contrast, for the structured network and data, the improvement is significantly larger in the novel environment (Fig. 3F); the improvement reaches 3-fold in novel relative to the familiar environment on real data, perhaps because of the larger population size.

The structure of the population responses also affects the ability of downstream circuits to read out stimulus relevant information. To assess this, we focused on the task of discriminating between randomly selected pairs of locations. We began by measuring the fraction of variance explained by the first principal component of the population activity patterns for the two locations (Fig. 3G,H). For the structured network model (and the data), the fraction was ~ 0.1 across experience. In contrast, in the random network the variance along the first principal component was 2-fold lower in the novel versus familiar environment. Thus, structured interactions appear to organize neural responses in the novel environment so that the code maintains a collective correlated response even when the input quality is weak. This effect might aid downstream areas to better differentiate responses to different stimuli (Ni et al., 2018). Therefore, we assessed the linear separability of spatial positions based on neural population responses, as a measure of whether information is readily available to downstream brain areas (Fig. 3I,J). For the random network, the performance of a linear classifier trained to discriminate random positions is significantly worse in the novel environment. In contrast, the performance is restored to a high value (~ 0.9) regardless of the environment by data-like interactions in the structured model, matching observations on real data (see Extended Data Fig. 3-2F for separability of positions as a function of their distance).

Taken together, our results suggest an important coding role for the interaction patterns inferred in Figure 2H and the corresponding structured networks explored in Figure 3 and Extended Data Figure 3-2. In comparison to the random

←

classification accuracy of the responses to random pairs of stimuli trained on 1000 pairs of same versus different positions for random (left) versus data-like structure (center) versus data (right). See Extended Data Figures 3-1, 3-2, 3-3 for more information.

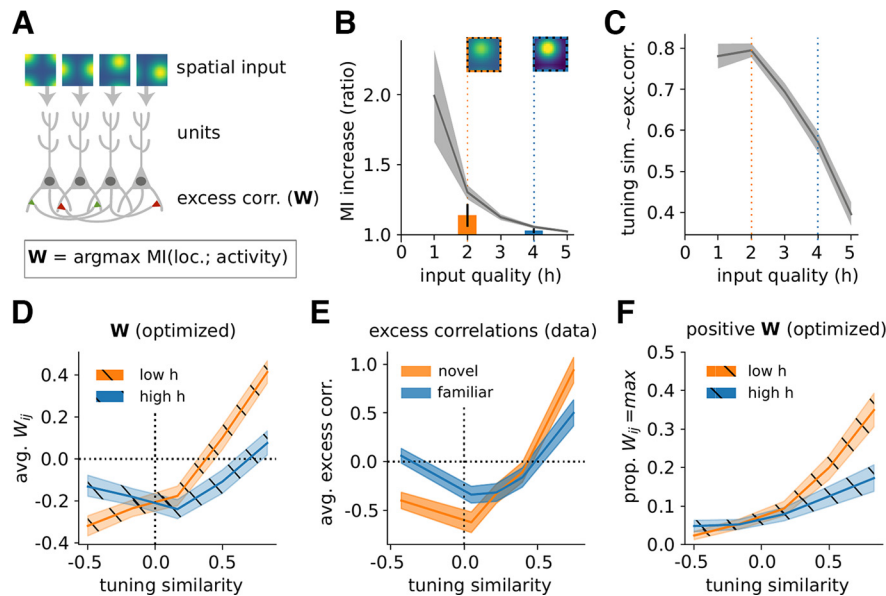


Figure 4. Predicted optimal network interactions. **A**, A schematic of the circuit model. Individual neurons, which receive spatially tuned inputs (with overall strength controlled by parameter h), are pairwise connected with interactions W_{ij} ; interactions are numerically optimized to maximize the mutual information between spatial position and population responses while constraining population mean firing rates and $|W_{ij}| \leq w_{\max}$ (here, $w_{\max} = 1$). **B**, Average ratio between mutual information (MI) in optimized versus noninteracting ($W = 0$) networks. Dashed vertical lines denote two chosen input quality levels, together with firing rate map of an example cell ("low quality" $h = 2$, orange, resembling novel environment; "high quality" $h = 4$, blue, resembling familiar environment). In all simulation plots we show averages over 1000 replicate optimizations with random initial assignments of place fields (see Materials and Methods); shaded area, 95th percentile CI for the mean. Vertical bars represent mean $\pm$ SD of MI ratio for networks using data-like interactions; t test against baseline: novel env. $p < 0.001$, familiar env. $p < 0.01$. **C**, Average alignment (Spearman's correlation) between pairwise input similarity and optimal W_{ij} as a function of input quality. **D**, Average magnitude of optimal W_{ij} as a function of tuning similarity for the two environments. **E**, Data proxy for **D**, computed using the excitatory-excitatory excess correlations w_{ij} estimated from data versus place field overlap. Note the vertical scale difference between **D** and **E**: excess correlations w_{ij} are a statistical proxy for the true interactions W_{ij} ; the two are expected to be correlated but not identical (Extended Data Fig. 1–1G). **F**, Proportion of optimal $W_{ij} = w_{\max} = 1$ as a function of tuning similarity. See Extended Data Figure 4–1 for more information.

network, the data-like, structured network (1) encodes more information about position especially when the input is of low quality; (2) improves single-cell spatial information and allows for easy retrieval with different decoders (population vector, covariance decoder, naive Bayes); (3) preserves linear separability under changes of input quality. Consistent conclusions hold for the comparison between the data-like, structured network and an uncoupled population (Extended Data Fig. 3–2).

CA1 interactions match predictions of an optimal coding model

While Figure 3 suggests that interactions between cells self-organize to improve spatial information coding relative to a random or an unconnected (Extended Data Fig. 3–2) network, it is not clear whether the observed organization is in any sense consistent with an optimal organization. To address this question, we numerically optimized cell-cell interactions among a population of place cells, so as to maximize the mutual information between the population activity and spatial position (Fig. 4A). In essence, this amounts to finding "efficient coding" solutions for network structure given inputs to individual cells that are correlated because of place field overlaps (Tkacik et al., 2010). As before, an important control parameter is the overall magnitude (quality) of the input spatial tuning, h , which we now vary parametrically. Resource constraints were simulated by constraining the optimization to keep the average population firing rate constant and the interactions bounded, $|W_{ij}| \leq w_{\max} = 1$ (see Materials and Methods).

As the input quality increases, the information gain because of optimal interactions decreases, indicating that optimization benefits novel environments (with noisy spatial inputs) more

than familiar environments (with reliable spatial inputs; Fig. 4B), providing an explanation for the differences observed in Figure 3B. When comparing the performance of data-like interactions (Fig. 3) against the optimal ones, in terms of mutual information improvements relative to a corresponding unstructured network, we found that the data estimated performance is close to optimal in both in familiar and novel environments (Fig. 4B; within optimal CI: familiar 10.6%, novel 17.2%; avg. improvement relative to optimal: fam. 0.371, nov. 0.583). We further find that an overlap in tuning similarity between two cells correlates with optimal pairwise interaction between them when input quality is low, but this correlation grows weaker with increasing input quality (Fig. 4C), consistent with theoretical expectations (Tkacik et al., 2010).

How do interactions improve spatial coding? We found that optimal interactions reduce the entropy of the stimulus-dependent population responses ("noise entropy") while largely preserving the total output entropy (Extended Data Fig. 3–3A). This explains the improvement at the level of single-cell place fields (Fig. 3C,D). Considering a factorized distribution, where each (improved) place field is preserved, but cells are made independent, one observes an even higher difference between total and noise entropy (Extended Data Fig. 3–3B). Accordingly, shuffling position-dependent responses so as to keep single-cell tuning fixed while removing excess correlations (Hazon et al., 2022) improves decoding accuracy (Extended Data Fig. 3–3C). These results explain a seeming paradox of shuffling-based analyses usually reporting negative effects of correlations (Hazon et al., 2022): if place cells could preserve their (improved) place fields but be independent, then the spatial information would benefit.

Does optimization predict a clear relationship between the tuning similarity and interaction strength for pairs of cells? Figure 4D shows two such relationships, for high and low input

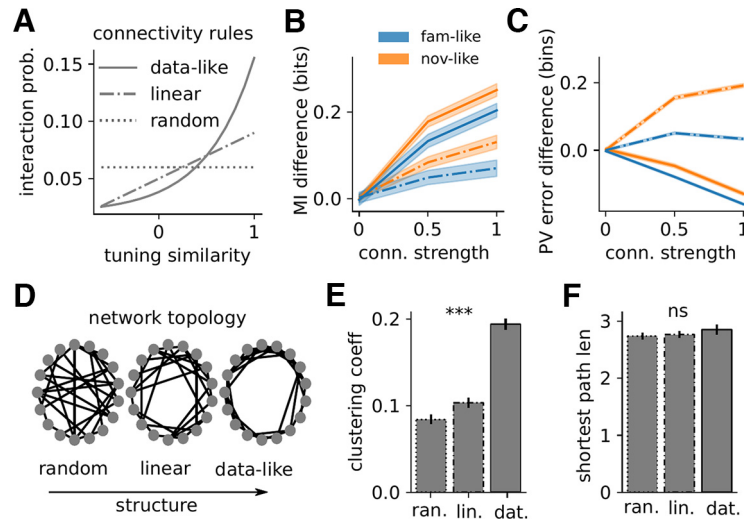


Figure 5. Data-like interaction pattern is sufficient to generate small-world networks with improved spatial coding. **A**, Connectivity rules for positive connections in a simulated place cell network with 50 units. **B**, Mutual information (MI) increase for data-like (solid) and linear (dashed) connectivity rule relative to the random connectivity, for familiar-like (blue) and novel-like (orange) quality input. Shaded areas show the 95th percentile confidence interval for the mean. **C**, Average decoding error increase for data-like (solid) and linear (dashed) connectivity rule relative to random connectivity. Error measured as distance between real and decoded spatial bin. **D**, Example network topologies obtained by using different connectivity rules from **A**. Nearby nodes have high tuning similarity. **E**, Average clustering coefficient for the three connectivity rules from **A** (error bars, SE; significance, one-way ANOVA test, *** $p < 0.001$ for or n.s. for $p > 0.05$). **F**, Average shortest path length for the three connectivity rules from **A**. See Extended Data Figure 5-1 for more information.

quality, predicted *ab initio* by maximizing spatial information. The optimal relationships closely resemble the analogous curves inferred from data (Fig. 4E). A similar resemblance is not observed if one maximizes spatial information carried by individual cells, highlighting the importance of information coding at the population, not individual-cell, level.

As an alternative comparison to experiments, we studied the proportion of optimized couplings that reached maximal allowable strength (positive: Fig. 4F; negative: Extended Data Fig. 4-1B). In the data, cells are deemed to be interacting when their excess correlation exceeds a threshold, and so Figure 2H represents a direct counterpart to our theoretical prediction. We find a clear qualitative match that includes the decrease in proportion of strong interactions for familiar environments (Extended Data Fig. 4-1C). We further observe that the proportion of optimal couplings reaching the constraint $w_{\max}$ scales nonlinearly with the tuning similarity, as in the data; the shape of the nonlinearity depends on the imposed $w_{\max}$ (Extended Data Fig. 4-1A). This shows how constraints shape the optimal solution, and suggests that CA1 prioritizes interactions among similarly tuned cells to counterbalance the limits on biological resources.

Although our simulations use a coarse-grained and down-scaled model of a real neural population (precluding exact comparisons), we observe an excellent qualitative match between theoretical predictions and the data. Taken together, these modeling results suggest that the network interactions in familiar versus novel environments observed in the CA1 data can be parsimoniously explained by the requirement of improving the fidelity of the population-level spatial representation across experience.

Data-like nonlinearity generates small-world networks with improved spatial coding

So far, our analysis of data as well as of optimized networks has identified a consistent pattern: the nonlinear dependence of interaction probability on tuning similarity (Figs. 2H, 4F). Figure 3 further showed that the pattern is necessary, since coding benefits were absent in randomized networks. The key remaining

question is whether the observed connectivity pattern is not only necessary, but also sufficient, to convey spatial coding benefits and generate networks with data-like topology.

To address this question, we generated model networks of 50 place cells, as before, but limited their connection strengths to three possible values, $\{-J, 0, +J\}$, where $J \in [0, 1]$ could be varied parametrically. We now used the interaction pattern of Figure 2H as a direct connectivity rule: we selected 6% of pairs (as in data) to have a positive connection $+J$ and connected them according to their tuning similarity as in data (Fig. 5A, “data-like”). To assess the role of the nonlinearity, we compared this with networks where the connection probability was linear in tuning similarity (“linear”) or where it was constant (“random”). In each of the three cases, a randomly chosen 3% of the place cell pairs (as in data) were connected with a negative strength, $-J$. As before, we fixed the average firing rate, and considered two levels of input quality, mimicking the familiar and novel environments (see Materials and Methods). This setup removed all structure (specifically, by making all connections have the same magnitude) except for that generated by the connectivity rule, allowing us to test for sufficiency.

First, we find that the data-like connectivity rule consistently improves mutual information between the population responses and position for increasing J , especially for novel-like input quality (Fig. 5B). This improvement is larger for the nonlinear, data-like connectivity than for the linear one. Extended Data Fig. 5-1C further suggests that connectivity alone accounts for a large fraction of mutual information gain, without the need for the fine-tuning of the interaction strengths. The data-like connectivity rule also improves the performance of a simple population vector decoder relative to random connectivity, in stark contrast to the linear dependence, which performs worse than the random one (Fig. 5C).

Finally, we asked whether different connectivity rules leave a strong signature on the network topology (Fig. 5D). To this end, we randomly generated 1000 networks according to the three different rules (Fig. 5A). The average clustering coefficient was substantially higher in networks created using the data-like rule

(Fig. 5E, similar to real data in Fig. 2I,J) compared with both the random and linear connectivity rules, without significantly affecting the distribution of incident edges (Extended Data Fig. 5-1A) or the average shortest path length (Fig. 5F). Additional analysis on the clique-complexes of the connectivity graphs revealed that the 1D Betti numbers are significantly smaller for the synthetic networks generated using the data-like rule compared with the two alternative rules, and are comparable with the data (Extended Data Fig. 5-1B). These analyses are consistent with the overexpression of triangles (Extended Data Fig. 2-4C) and high clustering coefficients (Fig. 2I,J) observed in the data-derived network. Overall, the nonlinear, data-like connectivity rule is sufficient to generate small-world network topologies with high spatial information.

Discussion

Statistical challenges limit our understanding of how experience shapes circuit interactions and, consequently, information coding during animal-driven behavior. While the use of pairwise correlations as a window into network interactions is not new (O'Neill et al., 2008; Bassett and Sporns, 2017; Humphries, 2017), the statistical problem of separating network interactions from other shared factors remains unsolved. Previous approaches based on stimulus-averaged correlations (Mathis et al., 2013), shuffles (Tocker et al., 2015), or joint model fits (Dunn et al., 2015; Posani et al., 2017) each suffer from statistical limitations (weak sample efficiency, strong model assumptions) which limit their applicability. Hence, most analyses of hippocampal collective behavior rely on total correlations (Harris et al., 2003; O'Neill et al., 2008; Lopes-dos Santos et al., 2013; El-Gaby et al., 2021; Gava et al., 2021). Unfortunately, total correlations potentially conflate changes in coding with changes in nuisance variables; even if the representation is unchanged, a shift in animal behavior (e.g., with experience) would change collective interactions defined based on total correlations. Furthermore, theta oscillations (Colgin, 2013), as well as the animal's speed (McNaughton et al., 1983; Fuhrmann et al., 2015), increase global synchrony and introduce spurious correlations. Only by factoring out all these known sources of covariability, compactly captured by spike synchrony, the fine structure of pairwise cell interactions can be revealed.

To reliably detect circuit interactions, we developed a novel statistical test rooted in the maximum entropy framework (Savin and Tkačik, 2017) and demonstrated its robustness in a range of simulated scenarios. Our work differs from previous maxEnt approaches in that we do not directly fit second-order moments (Loaiza-Ganem et al., 2017) or connectivity parameters (Bittner et al., 2021), but rather test their significance against a null model (Elsayed and Cunningham, 2017; Savin and Tkačik, 2017). These "excess correlations," are not intended to reflect monosynaptic interactions (Das and Fiete, 2020), but rather circuit effects, likely mediated by lateral inhibition (Klausberger and Somogyi, 2008). When applying our method to tetrode recordings from dorsal CA1 of freely behaving rats, we found stark differences between familiar and novel environments, especially in the EE subnetwork. Furthermore, we detected increased EI interactions in novel environments, unexplained by the higher reliability of direct excitatory-inhibitory connections (Csicsvari et al., 1998). Inhibition is generally weaker in a novel versus a familiar environment (Nitz and McNaughton, 2004; Arriaga and Han, 2019; Geiller et al., 2020), which may enhance learning by promoting

synaptic plasticity in excitatory neurons (Whitlock et al., 2006; Arriaga and Han, 2019). This in turn is linked to the sparsification of CA1 excitatory responses with increasing familiarity (Karlsson and Frank, 2008), although here this effect was minimized by only selecting cells active in both environments. Our observations in the novel environment are likely to derive from an increased dendritic excitability of pyramidal cells (Sheffield et al., 2017), believed to be necessary for place field formation and stabilization (Pedrosa and Clopath, 2020).

Our key statistical observation can be distilled as a monotonic nonlinear dependence of interaction probability on place field overlap for positive interactions among excitatory cells. This effect was observed across experience, but more prominently during novelty. We analyzed its coding implications in a model where we could separately control the informativeness of circuit inputs and the strength and structure of local interactions. We found that data-like interactions offered improvements in spatial information content and decoding. These improvements are visible at the single-cell level, as a sharpening of place fields and improved single cells' spatial information and at the population level, by reducing the conditional entropy of stimulus-driven responses while largely preserving the total output entropy (Tkacik et al., 2010). Coding advantages were higher during novelty, when CA1 needs to cope with worse quality input from CA3 (Leutgeb et al., 2004) and MEC (Hafting et al., 2005; Barry et al., 2012) and may expand to other situations where spatial input is noisy (e.g., darkness). We also found that data-like interactions improved stimulus discriminability, corroborating previous findings (da Silveira and Berry, 2014).

Efficient coding yields optimal solutions in which similarly tuned neurons have a higher probability of interacting positively. This is especially prominent for lower-quality inputs, where the predicted relation between interaction probability and tuning similarity is clearly nonlinear, as observed in the novel environment data. Simulated networks where this observed relationship is elevated to an actual connectivity rule show that (1) the observed relationship is sufficient to improve population spatial coding, and (2) the resulting network topology is small-world (Watts and Strogatz, 1998; Bullmore and Sporns, 2009). Small-world networks offer advantages for distributed information transfer and processing (Bassett and Bullmore, 2006) while being economical, i.e., they tend to minimize wiring while supporting high dynamical complexity (Karbowski, 2001); moreover, this topology is altered in disease (Bassett et al., 2008; Harrington et al., 2015). Nonetheless, while our results point toward small-worldness as one consequence of the particular CA1 connectivity structure, they do not provide evidence of intrinsic coding benefits from small-worldness per se (Latora and Marchiori, 2001; Gallos et al., 2012).

Although pairwise interactions only indirectly reflect underlying synaptic connectivity (Das and Fiete, 2020), together with the neuron tuning function they offer an accurate statistical description of a neural population output (Schneidman et al., 2006; Tkačik et al., 2014; Meshulam et al., 2017). Moreover, pairwise interactions can be studied using well established tools from information theory, which critically rely on the differentiation between stimulus selectivity overlap and network interactions to assess the amount of information that a population carries about a stimulus (Tkacik et al., 2010). We derived and tested an efficient coding hypothesis for a network of interacting place cells, by maximizing the mutual information between the animal's location (the stimulus) and the population response, while

holding input tuning and overall firing rate fixed. We found that network interactions adapt to different levels of input quality by employing different interaction versus tuning similarity strategies. In particular, for low input quality (i.e., at low signal-to-noise ratio mimicking the novel environment), optimal network interactions are strongly aligned with the tuning similarity of the interacting cells. When input quality is higher (i.e., at higher signal-to-noise ratio mimicking the familiar environment), this relation weakens yet remains detectable. These predictions closely resemble the data, suggesting that the CA1 circuit is close to an optimal operating regime across experience. This notion of optimality is different from traditional approaches which assume that noise correlation structure can be manipulated independently from individual tuning properties (Stefanini et al., 2020; Hazon et al., 2022). In this latter sense, removing noise correlations does further improve decoding performance also in our model, as previously demonstrated for CA1 responses on a linear track (Hazon et al., 2022). While such a manipulation is possible as a data analysis technique, it is hard to imagine how such a scenario could come about mechanistically in the brain, as disabling local interactions to remove noise correlations would automatically affect single-cell tuning as well; on the other hand, keeping sharp marginal single-cell tuning fixed while removing (optimal) local interactions that generated it is equivalent to a trivial improvement in the informativeness of the inputs. These considerations highlight the importance of mechanistic constraints for considering optimality hypotheses, and suggest that optimal interactions predominantly improve spatial coding in the hippocampus by sharpening the marginal tuning at the cost (and with the experimental signature) of injecting some amount of undesired, structured noise correlation.

Theory predicts that the optimal interactions and tuning overlap should negatively correlate at very high signal-to-noise ratios (Tkacik et al., 2010). This causes the neural population to decorrelate its inputs, a regime that is characteristic for coding in the sensory periphery (Simoncelli and Olshausen, 2001). While our numerical simulations reproduce this decorrelation regime at very high signal-to-noise ratio inputs, our inferences and data analyses suggest that it is not relevant for the hippocampal place code. This is likely because the overall noise levels are higher in the spatial navigation circuits compared with the sensory periphery, and partially because of the intrinsic differences in the statistics of the signal to be encoded (position vs natural images). Further work is needed to quantitatively relate the experimentally measured noise in CA1 inputs and responses to the effective “input quality” parameter that enters our predictions.

Are there previous reports where efficient coding predictions do not lead to decorrelation? A classic analysis in the retina correctly predicted that the receptive fields should lose their surrounds and switch to spatial averaging at low light (Atick and Redlich, 1990). A detailed study of retinal mosaics suggested that even during day vision receptive field centers of ganglion cells should (and do) overlap, increasingly so as the noise increases, leading to a residual redundancy in the population code (Doi and Lewicki, 2007; Borghuis et al., 2008), as reported (Puchalla et al., 2005). These findings support a more nuanced view of retinal coding (H. Barlow, 2001) than the initial redundancy reduction hypothesis (H.B. Barlow et al., 1961), precisely because they take into account the consequences of noise in the input and circuit processing (Linsker, 1989; van Hateren, 1992; Karklin and Simoncelli, 2011). Another study in fly vision focused on an interaction between two identified neurons, to find that its magnitude increased as the visual input

became more and more noisy, as theoretically predicted by information maximization (Weber et al., 2012). Psychophysics of texture sensitivity that arises downstream of the primary visual cortex further suggested that the relevant neural mechanisms operate according to the efficient coding hypothesis, yet in the input-noise-dominated regime where decorrelation is not optimal (Hermundstad et al., 2014). In light of these examples and our results, efficient coding, understood more broadly as information maximization (Chalk et al., 2018) rather than solely in its noiseless decorrelating limit, should be revisited as a viable candidate theory for representations in the central brain. More generally, our approach enables a synergistic interplay between statistical analysis, information theory, graph theory and traditional neural coding, and opens new ways for investigating neural coding during complex/naturalistic behavior in other systems.

References

- Arriaga M, Han EB (2019) Structured inhibitory activity dynamics in new virtual environments. *Elife* 8:e47611.
- Atick JJ, Redlich AN (1990) Towards a theory of early visual processing. *Neural Comput* 2:308–320.
- Barlow H (2001) Redundancy reduction revisited. *Network* 12:241–253.
- Barlow HB, et al. (1961) Possible principles underlying the transformation of sensory messages. *Sens Commun* 1:217–234.
- Barry C, Ginzberg LL, O’Keefe J, Burgess N (2012) Grid cell firing patterns signal environmental novelty by expansion. *Proc Natl Acad Sci U S A* 109:17687–17692.
- Bassett DS, Bullmore E (2006) Small-world brain networks. *Neuroscientist* 12:512–523.
- Bassett DS, Sporns O (2017) Network neuroscience. *Nat Neurosci* 20:353–364.
- Bassett DS, Bullmore E, Verchinski BA, Mattay VS, Weinberger DR, Meyer-Lindenberg A (2008) Hierarchical organization of human cortical networks in health and schizophrenia. *J Neurosci* 28:9239–9248.
- Bittner SR, Palmigiano A, Piet AT, Duan CA, Brody CD, Miller KD, Cunningham J (2021) Interrogating theoretical models of neural computation with emergent property inference. *Elife* 10:e56265.
- Borghuis BG, Ratliff CP, Smith RG, Sterling P, Balasubramanian V (2008) Design of a neuronal array. *J Neurosci* 28:3178–3189.
- Brun VH, Otnass MK, Molden S, Steffenach HA, Witter MP, Moser MB, Moser EI (2002) Place cells and place recognition maintained by direct entorhinal-hippocampal circuitry. *Science* 296:2243–2246.
- Brun VH, Leutgeb S, Wu HQ, Schwarcz R, Witter MP, Moser EI, Moser MB (2008) Impaired spatial representation in ca1 after lesion of direct input from entorhinal cortex. *Neuron* 57:290–302.
- Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10:186–198.
- Buzsáki G, Moser EI (2013) Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat Neurosci* 16:130–138.
- Chalk M, Marre O, Tkacik G (2018) Toward a unified theory of efficient, predictive, and sparse coding. *Proc Natl Acad Sci U S A* 115:186–191.
- Cheng S, Frank LM (2008) New experiences enhance coordinated neural activity in the hippocampus. *Neuron* 57:303–313.
- Cohen JD, Bolstad M, Lee AK (2017) Experience-dependent shaping of hippocampal ca1 intracellular activity in novel and familiar environments. *Elife* 6:e23040.
- Colgin LL (2013) Mechanisms and functions of theta rhythms. *Annu Rev Neurosci* 36:295–312.
- Csicsvari J, Hirase H, Czurko A, Buzsáki G (1998) Reliability and state dependence of pyramidal cell-interneuron synapses in the hippocampus: an ensemble approach in the behaving rat. *Neuron* 21:179–189.
- Csicsvari J, Hirase H, Czurko A, Mamiya A, Buzsáki G (1999) Oscillatory coupling of hippocampal pyramidal cells and interneurons in the behaving rat. *J Neurosci* 19:274–287.
- Csicsvari J, Hirase H, Mamiya A, Buzsáki G (2000) Ensemble patterns of hippocampal ca3-ca1 neurons during sharp wave-associated population events. *Neuron* 28:585–594.

- Das A, Fiete IR (2020) Systematic errors in connectivity inferred from activity in strongly recurrent networks. *Nat Neurosci* 23:1286–1296.
- da Silva RA, Berry MJ (2014) High-fidelity coding with correlated neurons. *PLoS Comput Biol* 10:e1003970.
- Doi E, Lewicki MS (2007) A theory of retinal population coding. *Adv Neural Inf Process Syst* 19:353.
- Dunn B, Mørreanet M, Roudi Y (2015) Correlations and functional connections in a population of grid cells. *PLoS Comput Biol* 11:e1004052.
- Dupret D, O'Neill J, Csicsvari J (2013) Dynamic reconfiguration of hippocampal interneuron circuits during spatial learning. *Neuron* 78:166–180.
- Eichenbaum H (2000) A cortical–hippocampal system for declarative memory. *Nat Rev Neurosci* 1:41–50.
- El-Gaby M, Reeve HM, Lopes d, Santos V, Campo-Urriza N, Perestenko PV, Morley A, Strickland LA, Lukács IP, Paulsen O, Dupret D (2021) An emergent neural coactivity code for dynamic memory. *Nat Neurosci* 24:694–704.
- Elsayed GF, Cunningham JP (2017) Structure in neural population recordings: an expected byproduct of simpler phenomena? *Nat Neurosci* 20:1310–1318.
- Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top–down processing. *Nat Rev Neurosci* 2:704–716.
- English DF, McKenzie S, Evans T, Kim K, Yoon E, Buzsáki G (2017) Pyramidal cell–interneuron circuit architecture and dynamics in hippocampal networks. *Neuron* 96:505–520.e7.
- Epszstein J, Brecht M, Lee AK (2011) Intracellular determinants of hippocampal CA1 place and silent cell activity in a novel environment. *Neuron* 70:109–120.
- Flaxman S, Wilson A, Neill D, Nickisch H, Smola A (2015) Fast Kronecker inference in Gaussian processes with non-gaussian likelihoods. *International Conference on Machine Learning*, Vol 37, pp 607–616. PMLR, July 07–09 2015, Lille, France.
- Frank LM, Brown EN, Wilson MA (2001) A comparison of the firing properties of putative excitatory and inhibitory neurons from CA1 and the entorhinal cortex. *J Neurophysiol* 86:2029–2040.
- Fuhrmann F, Justus D, Sosulina L, Kaneko H, Beutel T, Friedrichs D, Schoch S, Schwarz MK, Fuhrmann M, Remy S (2015) Locomotion, theta oscillations, and the speed-correlated firing of hippocampal neurons are controlled by a medial septal glutamatergic circuit. *Neuron* 86:1253–1264.
- Gallo LK, Makse HA, Sigman M (2012) A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proc Natl Acad Sci U S A* 109:2825–2830.
- Gava GP, McHugh SB, Lefèvre L, Lopes d, Santos V, Trouche S, El-Gaby M, Schultz SR, Dupret D (2021) Integrating new memories into the hippocampal network activity space. *Nat Neurosci* 24:326–330.
- Geiller T, Vancura B, Terada S, Troullinou E, Chavlis S, Tsagakakis G, Tsakalides P, Ócsai K, Poirazi P, Rózsa BJ, Losonczy A (2020) Large-scale 3d two-photon imaging of molecularly identified CA1 interneuron dynamics in behaving mice. *Neuron* 108:968–983.e9.
- Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell* 6:721–741.
- Giusti C, Pastalkova E, Curto C, Itskov V (2015) Clique topology reveals intrinsic geometric structure in neural correlations. *Proc Natl Acad Sci U S A* 112:13455–13460.
- Granat-Atedgi E, Tkačik G, Segev R, Schneidman E (2013) Stimulus-dependent maximum entropy models of neural population codes. *PLoS Comput Biol* 9:e1002922.
- Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436:801–806.
- Harremoës P (2001) Binomial and Poisson distributions as maximum entropy distributions. *IEEE Trans Inf Theory* 47:2039–2041.
- Harrington DL, Rubinov M, Durgarian S, Mourany L, Reece C, Koenig K, Bullmore E, Long JD, Paulsen JS; PREDICT-HD investigators of the Huntington Study Group; Rao SM (2015) Network topology and functional connectivity disturbances precede the onset of Huntington's disease. *Brain* 138:2332–2346.
- Harris KD (2005) Neural signatures of cell assembly organization. *Nat Rev Neurosci* 6:399–407.
- Harris KD, Henze DA, Csicsvari J, Hirase H, Buzsáki G (2000) Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J Neurophysiol* 84:401–414.
- Harris KD, Csicsvari J, Hirase H, Dragoi G, Buzsáki G (2003) Organization of cell assemblies in the hippocampus. *Nature* 424:552–556.
- Hazon O, Mincses VH, Tomàs DP, Ganguli S, Schnitzer MJ, Jercog PE (2022) Noise correlations in neural ensemble activity limit the accuracy of hippocampal spatial representations. *Nat Commun* 13:4276.
- Henze DA, Borhegyi Z, Csicsvari J, Mamiya A, Harris KD, Buzsáki G (2000) Intracellular features predicted by extracellular recordings in the hippocampus in vivo. *J Neurophysiol* 84:390–400.
- Hermundstad AM, Briguglio JJ, Conte MM, Victor JD, Balasubramanian V, Tkačik G (2014) Variance predicts salience in central sensory processing. *Elife* 3:e03722.
- Hill DN, Mehta SB, Kleinfeld D (2011) Quality metrics to accompany spike sorting of extracellular signals. *J Neurosci* 31:8699–8705.
- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A* 79:2554–2558.
- Humphries MD (2017) Dynamical networks: finding, measuring, and tracking neural population activity using network science. *Netw Neurosci* 1:324–338.
- Huszár R, Zhang Y, Blockus H, Buzsáki G (2022) Preconfigured dynamics in the hippocampus are guided by embryonic birthdate and rate of neurogenesis. *Nat Neurosci* 25:1201–1212.
- Johnson O (2007) Log-concavity and the maximum entropy property of the Poisson distribution. *Stochast Process Appl* 117:791–802.
- Jung MW, Wiener SI, McNaughton BL (1994) Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *J Neurosci* 14:7347–7356.
- Kaefer K, Malagon-Vina H, Dickerson DD, O'Neill J, Trossbach SV, Korth C, Csicsvari J (2019) Disrupted-in-schizophrenia 1 overexpression disrupts hippocampal coding and oscillatory synchronization. *Hippocampus* 29:802–816.
- Kanitscheider I, Coen-Cagli R, Pouget A (2015) Origin of information-limiting noise correlations. *Proc Natl Acad Sci U S A* 112:E6973–E6982.
- Karbowski J (2001) Optimal wiring principle and plateaus in the degree of separation for cortical neurons. *Phys Rev Lett* 86:3674–3677.
- Karklin Y, Simoncelli EP (2011) Efficient coding of natural images with a population of noisy linear-nonlinear neurons. *Adv Neural Inf Process Syst* 24:999–1007.
- Karlsson MP, Frank LM (2008) Network dynamics underlying the formation of sparse, informative representations in the hippocampus. *J Neurosci* 28:14271–14281.
- Klausberger T, Somogyi P (2008) Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. *Science* 321:53–57.
- Latora V, Marchiori M (2001) Efficient behavior of small-world networks. *Phys Rev Lett* 87:198701.
- Leutgeb S, Leutgeb JK, Treves A, Moser MB, Moser EI (2004) Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* 305:1295–1298.
- Levy ERJ, Carrillo-Segura S, Park EH, Redman WT, Hurtado JR, Chung SY, Fenton AA (2023) A manifold neural population code for space in hippocampal coactivity dynamics independent of place fields. *Cell Reports*, Vol 42. Elsevier.
- Linsker R (1989) An application of the principle of maximum information preservation to linear systems. *Adv Neural Inf Process Syst* 1:186–194.
- Loaiza-Ganem G, Gao Y, Cunningham JP (2017) Maximum entropy flow networks. *arXiv 1701.03504*. <https://doi.org/10.48550/arXiv.1701.03504>.
- Lopes-dos Santos V, Ribeiro S, Tort AB (2013) Detecting cell assemblies in large neuronal populations. *J Neurosci Methods* 220:149–166.
- Mathis A, Herz AV, Stemmler MB (2013) Multiscale codes in the nervous system: the problem of noise correlations and the ambiguity of periodic scales. *Phys Rev E Stat Nonlin Soft Matter Phys* 88:022713.
- McKenzie S, Huszár R, English DF, Kim K, Christensen F, Yoon E, Buzsáki G (2021) Preexisting hippocampal network dynamics constrain optogenetically induced place fields. *Neuron* 109:1040–1054.e7.
- McNaughton BL, Barnes CA, 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.
- McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser MB (2006) Path integration and the neural basis of the 'cognitive map.' *Nat Rev Neurosci* 7:663–678.
- Meshulam L, Gauthier JL, Brody CD, Tank DW, Bialek W (2017) Collective behavior of place and non-place neurons in the hippocampal network. *Neuron* 96:1178–1191.e4.

- Mizuseki K, Buzsaki G (2014) Theta oscillations decrease spike synchrony in the hippocampus and entorhinal cortex. *Philos Trans R Soc Lond B Biol Sci* 369:20120530.
- Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A (2014) Information-limiting correlations. *Nat Neurosci* 17:1410–1417.
- Morris RG, Garrud P, Rawlins JN, O'Keefe J (1982) Place navigation impaired in rats with hippocampal lesions. *Nature* 297:681–683.
- Moser EI, Paulsen O (2001) New excitement in cognitive space: between place cells and spatial memory. *Curr Opin Neurobiol* 11:745–751.
- Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR (2018) Learning and attention reveal a general relationship between population activity and behavior. *Science* 359:463–465.
- Nitz D, McNaughton B (2004) Differential modulation of ca1 and dentate gyrus interneurons during exploration of novel environments. *J Neurophysiol* 91:863–872.
- O'Keefe J, 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'Neill J, Senior T, Csicsvari J (2006) Place-selective firing of ca1 pyramidal cells during sharp wave/ripple network patterns in exploratory behavior. *Neuron* 49:143–155.
- O'Neill J, Senior TJ, Allen K, Huxter JR, Csicsvari J (2008) Reactivation of experience-dependent cell assembly patterns in the hippocampus. *Nat Neurosci* 11:209–215.
- Panzeri S, Moroni M, Safaai H, Harvey CD (2022) The structures and functions of correlations in neural population codes. *Nat Rev Neurosci* 23:551–567.
- Pedrosa V, Clopath C (2020) The interplay between somatic and dendritic inhibition promotes the emergence and stabilization of place fields. *PLoS Comput Biol* 16:e1007955.
- Perin R, Berger TK, Markram H (2011) A synaptic organizing principle for cortical neuronal groups. *Proc Natl Acad Sci U S A* 108:5419–5424.
- Posani L, Cocco S, Jezdek K, Monasson R (2017) Functional connectivity models for decoding of spatial representations from hippocampal ca1 recordings. *J Comput Neurosci* 43:17–33.
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes in c++. *Art Sci Comput* 2:1002.
- Puchalla JL, Schneidman E, Harris RA, Berry MJ (2005) Redundancy in the population code of the retina. *Neuron* 46:493–504.
- Rasmussen CE (2003) Gaussian processes in machine learning. In: *Summer school on machine learning*, pp 63–71. New York: Springer.
- Rumyantsev OI, Lecoq JA, Hernandez O, Zhang Y, Savall J, Chrapkiewicz R, Li J, Zeng H, Ganguli S, Schnitzer MJ (2020) Fundamental bounds on the fidelity of sensory cortical coding. *Nature* 580:100–105.
- Savin C, Tkacik G (2016) Estimating nonlinear neural response functions using gp priors and Kronecker methods. *Adv Neural Inf Process Syst* 29:3603–3611.
- Savin C, Tkacik G (2017) Maximum entropy models as a tool for building precise neural controls. *Curr Opin Neurobiol* 46:120–126.
- Schmitzer-Torbert N, Jackson J, Henze D, Harris K, Redish A (2005) Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* 131:1–11.
- Schneidman E, Berry MJ, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440:1007–1012.
- Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 20:11–21.
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423.
- Sheffield ME, Dombeck DA (2019) Dendritic mechanisms of hippocampal place field formation. *Curr Opin Neurobiol* 54:1–11.
- Sheffield ME, Adoff MD, Dombeck DA (2017) Increased prevalence of calcium transients across the dendritic arbor during place field formation. *Neuron* 96:490–504.e5.
- Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
- Skaggs WE, McNaughton BL, Gothard KM (1992) An information-theoretic approach to deciphering the hippocampal code. *Adv Neural Inf Process Syst* 5:1030–1037.
- Souza BC, Pavao R, Belchior H, Tort AB (2018) On information metrics for spatial coding. *Neuroscience* 375:62–73.
- Stefanini F, Kushnir L, Jimenez JC, Jennings JH, Woods NI, Stuber GD, Kheirbek MA, Hen R, Fusi S (2020) A distributed neural code in the dentate gyrus and in ca1. *Neuron* 107:703–716.e4.
- Stella F, Baracska P, O'Neill J, Csicsvari J (2019) Hippocampal reactivation of random trajectories resembling Brownian diffusion. *Neuron* 102:450–461.e7.
- Tkacik G, Prentice JS, Balasubramanian V, Schneidman E (2010) Optimal population coding by noisy spiking neurons. *Proc Natl Acad Sci U S A* 107:14419–14424.
- Tkacik G, Marre O, Amodei D, Schneidman E, Bialek W, Berry MJ (2014) Searching for collective behavior in a large network of sensory neurons. *PLoS Comput Biol* 10:e1003408.
- Tocker G, Barak O, Derdikman D (2015) Grid cells correlation structure suggests organized feedforward projections into superficial layers of the medial entorhinal cortex. *Hippocampus* 25:1599–1613.
- Valente M, Pica G, Bondanelli G, Moroni M, Runyan CA, Morcos AS, Harvey CD, Panzeri S (2021) Correlations enhance the behavioral readout of neural population activity in association cortex. *Nat Neurosci* 24:975–986.
- van Hateren JH (1992) A theory of maximizing sensory information. *Biol Cybern* 68:23–29.
- Virtanen P, et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 17:261–272.
- Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393:440–442.
- Weber F, Machens CK, Borst A (2012) Disentangling the functional consequences of the connectivity between optic-flow processing neurons. *Nat Neurosci* 15:441–448.
- Whitlock JR, Heynen AJ, Shuler MG, Bear MF (2006) Learning induces long-term potentiation in the hippocampus. *science* 313:1093–1097.
- Wilson M, McNaughton B (1993) Dynamics of the hippocampal ensemble code for space. *Science* 261:1055–1058.