J Comput Neurosci (2016) 41:269–293
DOI 10.1007/s10827-016-0617-5

Neural assemblies revealed by inferred connectivity-based models of prefrontal cortex recordings

G. Tavoni[1,2,3] · **S. Cocco**[1] · **R. Monasson**[2]

Received: 12 December 2015 / Revised: 7 July 2016 / Accepted: 12 July 2016 / Published online: 28 July 2016

Abstract We present two graphical model-based approaches to analyse the distribution of neural activities in the prefrontal cortex of behaving rats. The first method aims at identifying cell assemblies, groups of synchronously activating neurons possibly representing the units of neural coding and memory. A graphical (Ising) model distribution of snapshots of the neural activities, with an effective connectivity matrix reproducing the correlation statistics, is inferred from multi-electrode recordings, and then simulated in the presence of a virtual external drive, favoring high activity (multi-neuron) configurations. As the drive increases groups of neurons may activate together, and reveal the existence of cell assemblies. The identified groups are then showed to strongly coactivate in the neural spiking data and to be highly specific of the inferred connectivity network, which offers a sparse representation of the correlation pattern across neural cells. The second method relies on the inference of a Generalized Linear Model, in which spiking events are integrated over time by neurons through an effective connectivity matrix. The functional connectivity matrices inferred with the two approaches are compared. Sampling of the inferred GLM distribution allows us to study the spatio-temporal patterns of activation of neurons within the identified cell assemblies, particularly their activation order: the prevalence of one order with respect to the others is weak and reflects the neuron average firing rates and the strength of the largest effective connections. Other properties of the identified cell assemblies (spatial distribution of coactivation events and firing rates of coactivating neurons) are discussed.

Keywords Cell assemblies · Replay · Statistical inference · Ising model · Generalized linear model

Action Editor: Alain Destexhe

This work was funded by the [EU-]FP7 FET OPEN project Enlightenment 284801.

✉ G. Tavoni
tavoni@sas.upenn.edu

[1] Laboratoire de Physique Statistique, Ecole Normale Supérieure, CNRS, PSL Research, Sorbonne Université UPMC, Paris, France

[2] Laboratoire de Physique Théorique, Ecole Normale Supérieure, CNRS, PSL Research, Sorbonne Université UPMC, Paris, France

[3] Present address: Department of Physics & Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

1 Introduction

Cell assemblies, defined as groups of coactivating neurons, are thought to be essential units of neural representations and memories. Yet, identification of cell assemblies from the spiking data of neural populations, e.g. obtained from multielectrode recordings, is a complicated task. Among the currently available methods are template matching techniques (Carr et al. 2011; Diba and Buzsáki 2007; Foster and Wilson 2006; Lee and Wilson 2002), often combined with decoding methods (Brown et al. 1998; Johnson and Redish 2007). Those methods are well adapted to the detection of cell assemblies with known sensory correlates, typically

hippocampal cell assemblies of place cells. Other techniques have been proposed when correlates are unknown. Examples include the Markov Stability method (Billeh et al. 2014), which dynamically searches for the partitions of the neural (pairwise) correlation graph, and Principal Component Analysis-based methods, which assimilate cell assemblies to correlation modes of the data, extracted from each signal component of the covariance matrix (Peyrache et al. 2009; Peyrache et al. 2010b; Benchenane et al. 2010; Chapin and Nicolelis 1999) or from linear combinations of those signal components, e.g. in the Assembly Vector estimation method (Lopes-dos Santos et al. 2011; Lopes-dos Santos et al. 2013). In this work, we propose two methods to investigate the properties of cell assemblies, based on the inference of graphical models (capturing the functional connectivity) for the distribution of the neuronal activity in the spike recordings. Both methods are applied to recordings of the prefrontal cortex of behaving rats, during performance of a decision-making task and the preceding and following sleep epochs.

The first approach, inspired by statistical physics, consists in inferring functional interaction network models that reproduce the low-order statistics of snapshots (within a few tens of ms time bin) of the population activity. In a previous publication (Tavoni et al. 2015) we have analyzed the 'structural' changes in the networks between the sleep epochs, and how these changes are correlated with the network associated to the Task epoch. Here we extend our previous work and exploit the probabilistic power of the inferred graphical model to identify groups of coactivating neurons (assemblies) in single epochs. Indeed, once the model has been inferred (in each epoch) it can be simulated in the presence of an external drive, which transiently increases network excitability. We introduce the notion of neural susceptibility, which measures the change in the activity of the neuron due to a small change in the input drive. The susceptibility of a neuron depends on the full distribution of the activity of the other recorded neurons. Neural cells whose susceptibilities peak at the same value of the drive are used to define the cell assemblies. We apply this approach to the prefrontal cortex activity of behaving rats, recorded during the learning of a task (Task epoch) and during the sleeping epochs before (Sleep Pre) and after (Sleep Post) learning. The putative cell assemblies identified with our method are proved to correspond to groups of neurons with strong coactivation in the firing data. As a particular case, our method allows us to confirm the existence of groups of neurons coactivated in Task and reactivated during Sleep Post, representing the neuronal substrate of the 'replay' phenomenon, known to be important for memory consolidation, and previously identified through a comparison of the functional connectivity networks across the three epochs in (Tavoni et al. 2015). We discuss in details the significance of the drive in the simulations of the neural network, its possible implementations in the brain, both during Task and Sleep, and more generally the meaning and potential use of the method.

As the first method above relies on the inference of the model distribution of snapshots of the activity, and cannot give information about the temporal sequence of activation on longer time scales, we extend it through the introduction of dynamical, connectivity-based models. This second approach allows us to study the existence of specific orderings in the spiking sequence of the neurons in the identified assemblies. We investigate these temporal aspects by inferring and sampling a Bernoulli-Generalized Linear Model. In this model, the inferred connectivity matrix is not necessarily symmetric, and can be used to search for an ordering in the activation events. For the prefrontal cortex activity studied here, the asymmetry of the connectivity matrix inferred from the data is, however, weak. Sampling of the inferred model distribution of spatio-temporal patterns reveals only a partial ordering in the activation of neurons in the cell assemblies.

2 Materials and methods

2.1 Cortical recordings of behaving rats

We consider recordings of the activity of tens of neurons in the prefrontal cortex of behaving rats (Peyrache et al. 2009). Each recording session is divided in three $\simeq$ 30-minute epochs. During the Task epoch the rat has to learn a rule, such as go left, right, in the arm where the light is on, or off, in a Y–shaped maze; the rule is changed as soon as the rat has learned it. In addition, the activity is recorded during two Sleep epochs, one before (Sleep Pre) and one after (Sleep Post) the Task epoch. We consider in the present work four representative sessions (taken from three different rats) called 1, 2, 3 and 4 in the following; the first three correspond to respectively sessions A, H and D of (Tavoni et al. 2015). The numbers of recorded neurons in these sessions are, respectively, N =37, 10, 45 and 34.

Spiking times are binned within time windows of $\Delta t = 10$ ms; the activity configuration vectors $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, ..., \sigma_N)$ are snapshots of the neuronal activity, where σ_i takes value one or zero depending on whether the i-th neuron is respectively active or inactive in the time bin. We compute the N individual spiking probabilities, $f_i = \langle \sigma_i \rangle$, and the $\frac{1}{2}N(N-1)$ pairwise spiking probabilities, $f_{ij} = \langle \sigma_i \sigma_j \rangle$, within a time bin, where the average $\langle \cdot \rangle$ is computed over the binned activity configurations. In addition to providing access to the spiking times, the recorded activity identifies, through the analysis of the spectral properties of the EEG, specific periods

of the Sleep epochs, called Slow-Wave-Sleep (SWS), in which sharp-wave ripples are present. Those periods are known to be important for memory consolidation (Roumis and Franck 2015) through the reactivation of Task-related cell assemblies.

2.2 Snapshots of neural activities: the Ising model distribution

We consider the maximum entropy (least constrained) distribution P over the 2^N configurations $\boldsymbol{\sigma}$ in a time bin that reproduces all the measured f_i and f_{ij} (Schneidman et al. 2006). This distribution for 'snapshots' of the neural activity configurations $\boldsymbol{\sigma}$ defines a graphical model, called Ising model in statistical physics:

$$P(\boldsymbol{\sigma}\,|\{h_i\},\{J_{ij}\}) = \frac{1}{Z}\exp\left(\sum_{i<j} J_{ij}\,\sigma_i\,\sigma_j + \sum_i h_i\,\sigma_i\right) \tag{1}$$

where Z ensures the normalization of the distribution. The $N(N+1)/2$ parameters h_i and J_{ij} are fitted to reproduce all the experimental individual and pairwise firing probabilities

$$\begin{aligned} f_i &= \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}\,|\{h_i\},\{J_{ij}\})\,\sigma_i\ , \\ f_{ij} &= \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}\,|\{h_i\},\{J_{ij}\})\,\sigma_i\,\sigma_j, \end{aligned} \tag{2}$$

within their sampling errors. Local inputs h_i capture each neuron's overall tendency to discharge, while couplings J_{ij} define statistical pairwise interactions between the cells and allow us to disentangle direct (two cells fire together because they interact) from indirect (mediated by other recorded cells) correlations. Inferring the field and coupling parameters from the spiking data, i.e. solving the self-consistent (2), is a hard computational task. In practice, we resort to the Adaptive Cluster Expansion of (Cocco and Monasson 2011, 2012, Barton and Cocco 2013), an approximate but accurate procedure to get the Ising model parameters; the inference code is available from (Barton et al. 2016).

Once the parameters defining P are inferred, the distribution can be sampled with the Monte Carlo Markov Chain (MCMC) procedure, e.g. with the Metropolis-Hastings updating rule. We may in particular compute high-order correlations, such as the probability that a given triplet of neurons fire in the same time bin, from the model and compare the outcomes to their experimental counterparts estimated from the data. We observe that the inferred model is able to accurately predict higher order statistical properties of the population activity, see (Tavoni et al. 2015) for details. In addition, the network of couplings turns out to be a sparser representation of the interaction structure compared to the network of correlations as a result of the extraction of the signal from the background sampling noise in the inference procedure (Cocco and Monasson 2012) and of the fact that couplings can be seen as direct correlations obtained after removing the indirect ones which are mediated by the network. The inferred sparse coupling structure seems to reflect a sparse property of the neural code, in agreement with (Ganmor et al. 2011a; 2011b) and theoretical results (Ganguli and Sompolinsky 2012; Abbeel et al. 2006).

2.3 Modifying the activity distribution: input drive H

From the model probability distribution (1), we define the log ratio of the probabilities that neuron i is active or silent given the other neuron activities $\{\sigma_j\}$ with $j \neq i$:

$$V_i(\{\sigma_j\}) = \log\left[\frac{P(\sigma_1,\ldots,\sigma_i=1,\ldots,\sigma_N|\{h\},\{J\})}{P(\sigma_1,\ldots,\sigma_i=0,\ldots,\sigma_N|\{h\},\{J\})}\right]. \tag{3}$$

It is easy to check that, with distribution P defined in Eq. (1), this log ratio coincides with

$$V_i(\{\sigma_j\}) = \sum_{j(\neq i)} J_{ij}\sigma_j + h_i\ , \tag{4}$$

and can be interpreted as the total input to neuron i, due to the other neurons and to the local field h_i. If V_i is positive neuron i has a higher probability of being active than silent, while if V_i is negative neuron i is most likely silent.

To sample rare events with many active neurons from the distribution P we modify the log ratios V_i according to the transformation

$$V_i(\{\sigma_j\}) \to V_i(\{\sigma_j\}) + H, \tag{5}$$

where H is a control parameter. Under this transformation the distribution of neural activities is changed into

$$P(\boldsymbol{\sigma}\,|\{h_i\},\{J_{ij}\}) \to P\left(\boldsymbol{\sigma}\,\middle|\,\{h_i+H\},\left\{J_{ij}\right\}\right)\ . \tag{6}$$

The parameter H, hereafter called drive, is an additive global input acting on all neurons. For the small time bins $\Delta t = 10$ ms used to infer the model, P in Eq. (1) is maximal for the all-silent neuron configuration, $\sigma_i = 0$ for all $i = 1, ..., N$, as neurons typically have low firing rates and are frequently silent at that time scale (giving rise to negative h_i). Nevertheless, self-sustaining patterns in which multiple neurons are active can emerge following transient increases in network excitability, as may be induced by e.g. sharp waves in slow oscillations of sleep. In order to reproduce assembly-generating transients we boost uniformly the h_i beyond the values inferred with the choice $\Delta t = 10$ ms, through the addition of the global drive H (Litwin-Kumar and Doiron 2012) to all neurons. For particular values of the drive an avalanche process may be triggered in the presence

of sufficiently strong interactions, which results in the rapid activation of several neurons representing the recorded part of a cell assembly, as illustrated in Fig. 1 and caption.

2.4 Neural susceptibilities measure changes in activation under the drive

The application of a drive changes the average values of the single neuron activities,

$$\langle \sigma_i \rangle (H) = \sum_{\boldsymbol{\sigma}} P\left(\boldsymbol{\sigma} \,\middle|\, \{h_i + H\}, \left\{J_{ij}\right\}\right) \sigma_i \ , \tag{7}$$

compared to their zero-drive counterparts, i.e. $\langle \sigma_i \rangle (H = 0) = f_i$. We define the (single neuron) susceptibility as the change in the probability of firing (in a time bin of width Δt) of neuron i resulting from a small increase of the external drive H:

$$\chi_i(H) = \frac{\partial \langle \sigma_i \rangle}{\partial H}(H) \ . \tag{8}$$

Susceptibilities can be directly computed through MCMC sampling at two close values of the drive, H and $H + \epsilon$. However, estimates of derivatives are notoriously difficult to obtain accurately due to the sampling noise. We therefore resort to the fluctuation-dissipation theorem (Chandler 1987), which relates responses to perturbation (here, H) to connected correlations. More precisely, we have

$$\begin{aligned} \chi_i(H) &= \sum_{j=1}^{N} \frac{\partial \langle \sigma_i \rangle}{\partial h_j}(H) \\ &= \sum_{j=1}^{N} \left[\langle \sigma_i \, \sigma_j \rangle (H) - \langle \sigma_i \rangle (H) \, \langle \sigma_j \rangle (H) \right] \ . \end{aligned} \tag{9}$$

This formula allows us to estimate the susceptibilites $\chi_i(H)$ in a reliable way.

In the case of an independent neuron, say, $i = 1$, with local field h_1, the average activity is

$$\langle \sigma_1 \rangle (H) = \frac{e^{h_1 + H}}{1 + e^{h_1 + H}} \ , \tag{10}$$

and the susceptibility is given by

$$\chi_1(H) = \frac{\partial \langle \sigma_1 \rangle (H)}{\partial H} = \frac{e^{h_1 + H}}{(1 + e^{h_1 + H})^2} \ . \tag{11}$$

This function has an upward bell-like shape, ranging between 0 and the maximal value 0.25, reached for $H = -h_1$. The baseline $\chi = 0.25$ therefore represents a null value for the maximal susceptibility, corresponding to neurons essentially independent from any other neuron in the population.

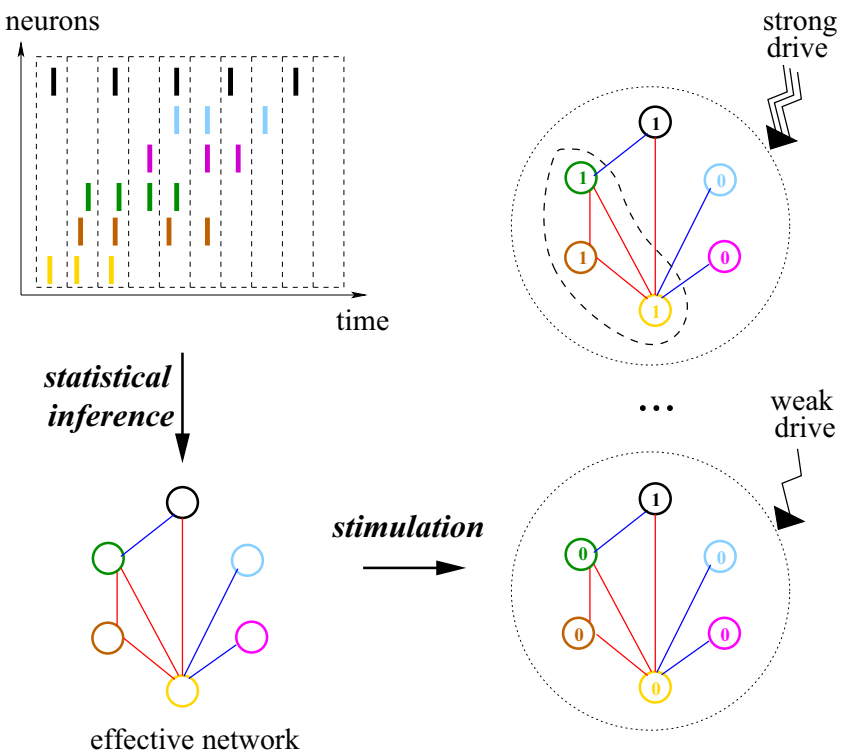

Fig. 1 Ising model for the neural activity: definition and simulation. Spiking times are binned into time bins of width Δt; each neuron i is assigned the variable $\sigma_i = 1$ or 0, if it is active in the time bin or not *(top, left)*. A model of the neural activity distribution (P in Eq. (1)) is inferred to reproduce the 1- and 2-cell spiking probabilities of this binned data *(bottom, left)*; red and blue links correspond, respectively, to positive and negative effective couplings J_{ij} in the inferred network. The model distribution is then analysed with the addition of an increasing drive (stimulation), which favors configurations with more and more active neurons *(right)*. As the drive increases a group of neurons (comprised in the dashed contour) may abruptly coactivate, defining a cell assembly

2.5 Self-sustaining activity configurations

Of particular interest are self-sustaining activity patterns, defined as the set of configurations $\boldsymbol{\sigma}$ such that

$$(\sigma_i = 1 \text{ and } V_i > 0) \quad \text{or} \quad (\sigma_i = 0 \text{ and } V_i < 0) \tag{12}$$

for all neurons i. Informally speaking, self-sustaining patterns are activity configurations in which each neuron reads the activity of the other cells, responds in a way consistent with its total input V_i, and in turn participates as an input to the other cells, a concept closely related to Hebb's classical definition of a cell assembly as "a diffuse structure [...] capable of acting briefly as a closed system" (Hebb 1949).

Self-sustaining patterns coincide with the local maxima of the model distribution P in Eq. (1). Indeed, consider a configuration $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_i, ..., \sigma_N)$ and change the activity of one neuron, say, i, to obtain configuration $\boldsymbol{\sigma}^{(i)} = (\sigma_1, ..., 1 - \sigma_i, ..., \sigma_N)$. The resulting variation in the log probability is

$$\begin{aligned} \log P(\boldsymbol{\sigma}^{(i)}|\{h\},\{J\}) \; - \; & \log P(\boldsymbol{\sigma}|\{h\},\{J\}) \\ &= \log\left[\frac{P(\boldsymbol{\sigma}^{(i)}|\{h\},\{J\})}{P(\boldsymbol{\sigma}|\{h\},\{J\})}\right] \\ &= (1-2\sigma_i)\; V_i(\{\sigma_j\}), \end{aligned} \tag{13}$$

according to Eq. (3). Hence, $\boldsymbol{\sigma}$ is a local maximum of P if and only if $1 - 2\sigma_i$ and V_i have opposite signs for all neurons i, or equivalently, if and only if it is a self-sustaining configuration.

In the presence of a drive H, finding self-sustaining configurations is relatively fast, and does not require MCMC equilibration, unlike computation of the susceptibilities. In practice, we start with the all-silent neuron configuration ($\sigma_i = 0$ for $i = 1, \ldots, N$). If the configuration is self-sustaining, the algorithm has found a maximum of P and halts. If one or more neurons are not self-sustaining, i.e. their values σ_i do not agree with the signs of their total inputs $V_i + H$, we pick up uniformly at random one of them, say, i, and flip its value σ_i (from silent to active, or vice-versa). This asynchronous updating is iterated until the configuration is self-sustaining. The procedure is guaranteed to converge as the log-probability of the configuration increases after each updating step.

Re-running the dynamics may, however, produce different maxima (for a given H), due to the stochasticity in the choice of the (non self-sustaining) neuron to flip at each step. In practice, after about one hundred runs new maxima are not found, and the dynamics ends up in already encountered maxima. To be reasonably sure that all self-sustaining configurations have been found, we set the number of runs to 10^6.

2.6 Temporal patterns of neural activities: the Bernoulli-Generalized-Linear model distribution

The Ising model gives a time-independent probability distribution of all possible neuronal configurations in a time bin, and does not allow us to study spatio-temporal patterns of the activity. To do so, we introduce a Generalized Linear Model (GLM) (Truccolo et al. 2005), with Bernoulli variables $\sigma_{i,t}$ representing the activity of neuron i in time bin t. In the Bernoulli-GLM, the probability of the activity variable $\sigma_{i,t} = 0, 1$ is given by

$$p(\sigma_{i,t}) = \frac{\lambda_{i,t}^{\sigma_{i,t}}}{1+\lambda_{i,t}}\,, \tag{14}$$

where $\lambda_{i,t}$ is a function of the activities at previous times, defined through

$$\log \lambda_{i,t} = h_i + \sum_{j\neq i} J_{ij} \sum_{\tau=0}^{T_1} e^{-\tau/\tau_1}\sigma_{j,t-\tau} + J_{ii}\sum_{\tau=1}^{T_2} e^{-\tau/\tau_2}\sigma_{i,t-\tau} \tag{15}$$

with parameters $\{h_i, J_{ij}, J_{ii}\}$. We choose time bins of width $\Delta t = 5$ ms. We set the synaptic integration time constant to $\tau_1 = 2$ time bins, which is a biophysically plausible value (McCormick et al. 1985), and the integration time constant of a neuron past history to $\tau_2 = 8$ time bins. This value is chosen according to the histogram of the inter-spike-intervals (ISI) for each neuron: in a rough approximation the probability of ISI is proportional to a decreasing exponential of ISI with time constant τ_2; τ_2 can therefore be estimated from the typical slope of the histograms of ISIs in log-scale. We impose a cut-off equal to e^{-4} in the second and third addends of Eq. (15), that is, we set $T_1 = 8$ and $T_2 = 32$ for the upper limits of the sums over past time bins.

The pseudo-log-likelihood of the spiking data under this model is given by:

$$\log P = \sum_{i,t} (\sigma_{it}\log\lambda_{it} - \log(1+\lambda_{it})) \tag{16}$$

Note that $\log P$ is not normalized because it includes instant terms (with $\tau = 0$) in the second addend of Eq. (15); hence P should be interpreted as the product of the conditional single-variable probabilities (pseudo likelihood).

As $\log P$ is a concave function of the B-GLM parameters $\boldsymbol{\psi} = \{h_i, J_{ij}, J_{ii}\}$ (Shlens 2014), inference can be done with Newton's method. We compute the gradient and the Hessian of $\log P$, i.e. the matrix $\mathcal{M}(\boldsymbol{\psi}^{(n)})$ of its second derivatives in $\boldsymbol{\psi}^{(n)}$, and recursively update $\boldsymbol{\psi}$ through

$$\boldsymbol{\psi}^{(n+1)} = \boldsymbol{\psi}^{(n)} - \left[\mathcal{M}(\boldsymbol{\psi}^{(n)})\right]^{-1} \cdot \nabla \log P(\boldsymbol{\psi}^{(n)})$$

until $\|\boldsymbol{\psi}^{(n+1)} - \boldsymbol{\psi}^{(n)}\| < \epsilon$ (we choose $\epsilon = 10^{-3}$ in practice). Once convergence has been reached the diagonal

elements of the inverse of $\mathcal{M}$ can be used to estimate the error bars on the inferred parameters.

2.7 Coactivation factor

We define the coactivation ratio (CoA) of a group G of neurons over the time scale τ through

$$\mathrm{CoA}(G,\tau) = \frac{f(G,\tau)}{\prod_{i\in G} f_i(\tau)} \tag{17}$$

where $f(G,\tau)$ is the probability that all the neurons in the group are active within the time scale τ, and the denominator is the product of the individual spiking probabilities over the same time scale. For a group of independent cells the CoA is on average equal to unity. As a relevant coactivation event contributing to $f(G,\tau)$ should correspond to a sequence of spikes, in which each neuron 'reads' the preceding spike and triggers in turn the next spike, we expect τ to be not larger than $n \times \Delta t$, where n is the number of neurons in G and $\Delta t = 10$ ms is the time-bin duration used for the inference, capturing the interactions between pairs of neurons.

The statistical validity of the CoA defined in Eq. (17) can be assessed as follows. Assuming a Poisson distribution for the coactivation events, the standard deviation of the CoA is estimated to be $\mathrm{CoA}(G,\tau)/\sqrt{N_G(\tau)}$, where $N_G(\tau)$ is the number of coactivation events for the cells in G over the time scale τ. Note that simultaneous-firing events (contributing to $f(G,\tau)$) are unlikely to be found, and the CoA is likely to be zero, if the duration of the recording is small, e.g. compared to $T_{min} = \tau / \prod_{i\in G} f_i(\tau)$.

3 Results

3.1 Sub-groups of neurons show maximal or minimal susceptibilities for similar values of the drive

Single-neuron susceptibilities $\chi_i(H)$ are shown as functions of the drive H for session 1 in Fig. 2, top. For many neurons the susceptibilities remain small, in the range [0;0.25] compatible with what is expected for independent cells (Methods). However, some neurons show large positive susceptibilities or even negative values, a result of the presence of strong effective excitatory or inhibitory couplings in the inferred network. We observe in particular the existence of subgroups of neurons, whose susceptibilites show large positive peaks at the same value (or for very close values) of the drive H.

We classify the observed behaviours into three main types summarized in Fig. 2, bottom. The neural susceptibility may have an upward bell-like shape with a global maximum (**a**) or a downward bell-like shape with a global

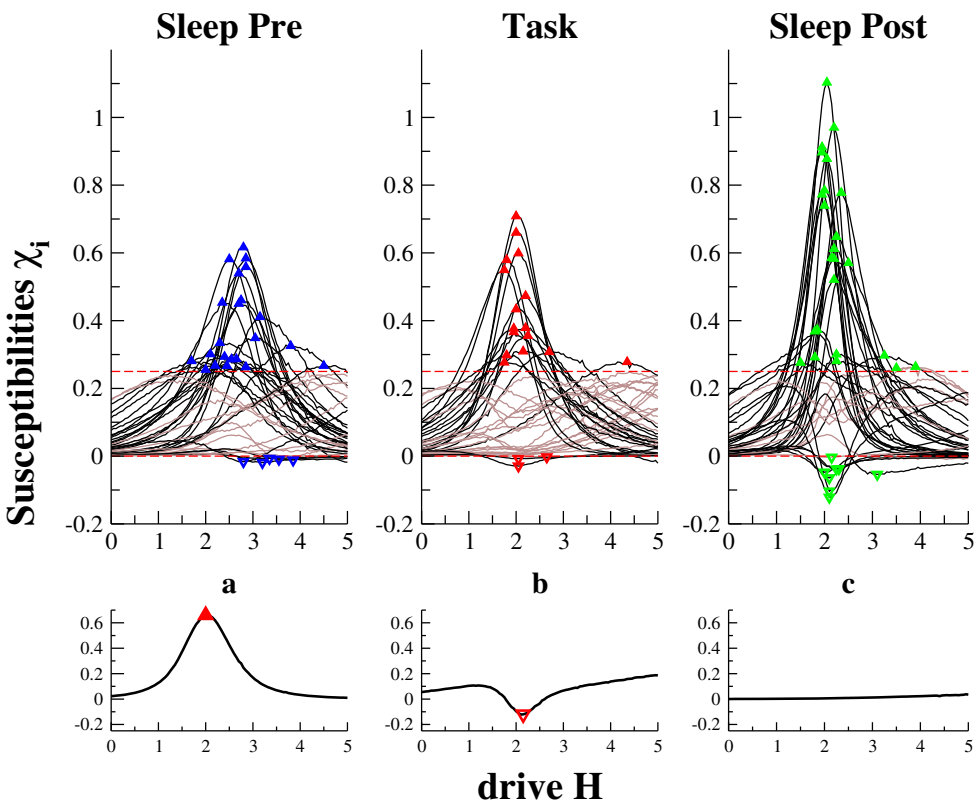

Fig. 2 Susceptibilities as functions of the drive. Top: susceptibilities $\chi_i(H)$ across the three epochs of session 1. Upward triangles locate the top susceptibility maxima ($\chi_i^+ > 0.25$), while downward triangles show the lowest susceptibility minima ($\chi_i^- < 0$) of Sleep Pre (blue), Task (red) and Sleep Post (green). Susceptibilities with such maxima and minima are shown in black, while the other ones are shown in light brown. Bottom: representative examples of neuron susceptibilities found in the Task epoch of session 1: (**a**) upward bell-shaped susceptibility with a maximum (neuron 26); (**b**) downward bell-shaped susceptibility with a minimum (neuron 27); (**c**) almost flat susceptibility over the range of H values considered here (neuron 31). Similar susceptibility shapes are found in the other epochs and sessions

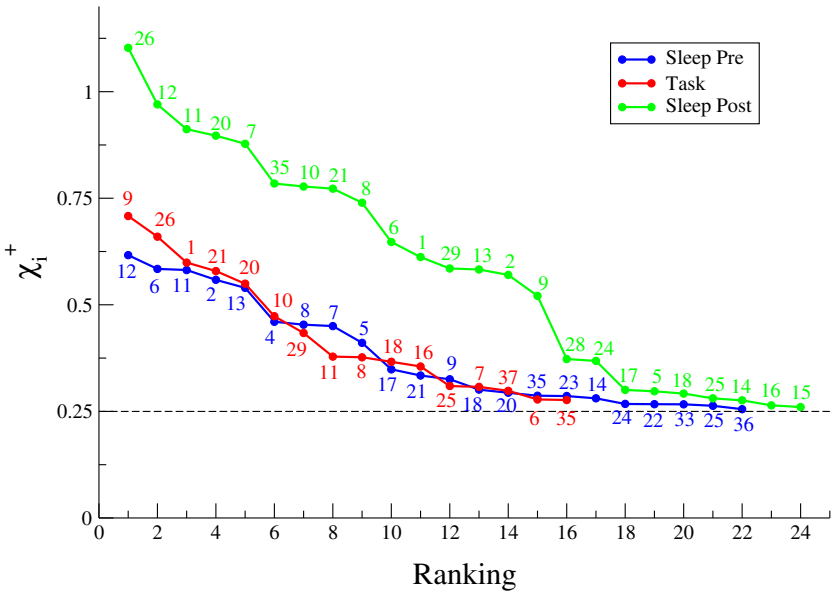

Fig. 3 Ranking of the maxima $\chi_i^+ > 0.25$ of the neural susceptibilities in the three epochs of Session 1 (same color code as in Fig. 2). Labels indicate the corresponding neuron numbers

minimum (**b**), or may not show any marked minimum or maximum over the range of H considered here (**c**).

For each neuron i with a susceptibility of type **a** or **b**, we define the input drives

$$H_i^+ = \text{argmax}\ \chi_i(H) \quad \text{or} \quad H_i^- = \text{argmin}\ \chi_i(H)\,, \qquad (18)$$

corresponding to, respectively, the maximum or minimum of the susceptibility:

$$\chi_i^+ = \max\ \chi_i(H) \quad \text{or} \quad \chi_i^- = \min\ \chi_i(H)\,. \qquad (19)$$

Type-**a** corresponds to neurons i whose activity largely increases when the drive value crosses H_i^+, while type-**b** neurons undergo a sharp decrease in activity when $H \simeq H_i^-$. In Fig. 2, top, pairs (H_i^+, χ_i^+) for type-**a** neurons are represented as upward triangles and pairs (H_i^-, χ_i^-) for type-**b** neurons as downward triangles; maxima χ_i^+ lower than 0.25, the top value expected for independent neurons (Methods), are not shown. Both susceptibility maxima and minima are present at approximately the same value of the input H, which indicates that the drive is able to elicit the activation of a group of neurons and the inhibition of another group of neurons.

We show in Fig. 3 the ranking of the maxima $\chi_i^+ > 0.25$ of the neural susceptibilities for the three epochs of Session

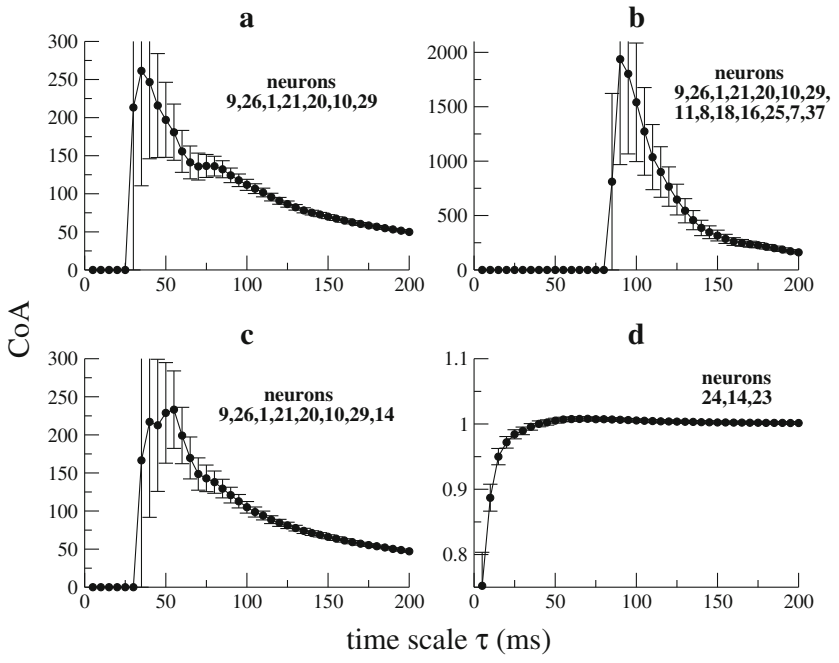

Fig. 4 Coactivation ratio (17) for groups of neurons in the Task epoch of Session 1. Panels (**a**) and (**b**) show two groups whose neurons have large susceptibilities, see Fig. 3. Panel (**c**) shows the CoA of the group in panel (**a**) after inclusion of a low-susceptibility neuron. The three neurons in panel (**d**) have low susceptibilities; their CoA is close to unity as expected for independent cells

1. Interestingly, maxima may reach a much higher value in the Sleep Post epoch than in Sleep Pre. In addition, we observe that some neurons with low susceptibilities in Sleep Pre, e.g. neurons 20 and 26, have large maxima in Sleep Post. Those neurons generally have large χ_i^+ values also in Task, and are part of the 5-cell replay group identified in (Tavoni et al. 2015) through a comparison of the inferred functional networks across epochs.

3.2 Groups of neurons with large susceptibilities strongly coactivate

We now show that the neurons with large susceptibility maxima (at similar values of the drive) identified in the previous section indeed strongly coactivate in the spiking data. To this aim we consider the coactivation ratio (CoA) of a group G of neurons over the time scale τ defined in Eq. (17). For a group of independent cells the CoA is on average equal to unity, while it is expected to be much larger for coactivating neurons. Figure 4 shows the CoA for different groups of neurons with large susceptibilities (panels (**a**) and (**b**)). The first group (**a**) includes the seven neurons with top susceptibilities (and includes as a subgroup the 5-cell replay group identified in (Tavoni et al. 2015)), while the second group is made of the fourteen cells with top susceptibilities. We observe that the CoA takes very high values, much larger than expected for independently firing cells. These high CoA values are also much higher than the statistical error bars estimated from a simple Poisson model (Methods). Note that the time scale at which the CoA is maximal seems to increase linearly with the group size, as expected. Similar results are obtained with other groups of neurons with high susceptibilities (not shown).

The two bottom panels in Fig. 4 serve as controls to our detection of coactivation based on susceptibilities. In panel (**c**) we show the CoA of the group of panel (**a**) upon inclusion of neuron 14, whose susceptibility curve as a function of H is similar to the one of an independent neuron (Methods). We observe that the CoA is left unchanged by the addition of this neuron. In panel (**d**) we show the CoA of a group of three neurons with low susceptibilities ($\chi_i^+ < 0.25$). The CoA saturates to a value close to unity at very short time scales τ.

In Appendix B we show the results of the application of this method to sessions 2, 3 and 4. In all these sessions we were able to identify cell assemblies, confirmed by computation of the coactivation ratios. In session 2 a large replay group is sampled, like in session 1, while session 3 is characterized by substantially different cell assemblies in the recordings of the three epochs and session 4 by the presence of a conserved group of neurons, coactivating in all epochs.

3.3 Properties of neurons in self-sustaining configurations are indicative of their susceptibilities

The results above can be found back from the analysis of the self-sustaining activity configurations under a drive (Methods). The advantage with respect to the analysis with the neuron susceptibilities is that finding self-sustaining configurations is faster than computing susceptibilities because it does not require to sample the distribution P (Methods). The disadvantage is that the analysis based on self-sustaining patterns is approximate in that it does not take into account fluctuations of the activity around the local maxima of P; on the contrary the analysis with neuron susceptibilities, relying on the full ditribution P, is more accurate since real neural networks are noisy: stochasticity might arise from e.g. ion channels, synaptic transmission and inputs from the surrounding environment (including non-recorded neurons). Here we illustrate the simplified, yet faster procedure to identify cell assemblies.

The number of active neurons in the self-sustaining patterns as a function of H is shown in Fig. 5 for the three epochs of session 1. When $H = 0$, the only self-sustaining configuration is the all-silent neuron configuration. As H increases, neurons start to activate one after the other, in decreasing order of their fields h_i, as expected. As more and more neurons i get activated, contributions to the total inputs V_j of the other neurons j build up, facilitating ($J_{ij} > 0$) or hindering ($J_{ij} < 0$) their activations. Discontinuous 'jumps' in the number of active neurons are observed at special values of H (arrows in Fig. 5), corresponding to the coactivation of a group of neurons under the drive. These jumps correspond to the presence of two (or more) coexisting local maxima of P, or self-sustaining patterns, one corresponding to a low level of activity and the other one to a higher level of activity: the difference between the two local maxima thus represents a group of neurons coactivating in the avalanche, that is, a putative cell assembly.

To assess the participation of each neuron to a jump we consider its conditional average activity, given the other neuron activities, in the coexisting self-sustaining configurations. The conditional average activity of a neuron i is simply derived from the Ising distribution in Eq. (1):

$$\langle \sigma_i \rangle \left(\{\sigma_{j(\neq i)}\} \right) = \frac{P(\sigma_1,\ldots,\sigma_i=1,\ldots,\sigma_N)}{P(\sigma_1,\ldots,\sigma_i=0,\ldots,\sigma_N)+P(\sigma_1,\ldots,\sigma_i=1,\ldots,\sigma_N)} = \frac{1}{1+e^{-V_i}} . \quad (20)$$

Variations $\delta\langle\sigma_i\rangle$ between the conditional activities $\langle\sigma_i\rangle$ in the low- and the high–activity self-sustaining patterns corresponding to the jumps of session 1 are shown for all recorded neurons in the insets of Fig. 5. A high variation $\delta\langle\sigma_i\rangle$ indicates that neuron i is strongly interacting with all (other) neurons activated in the jump and takes part in

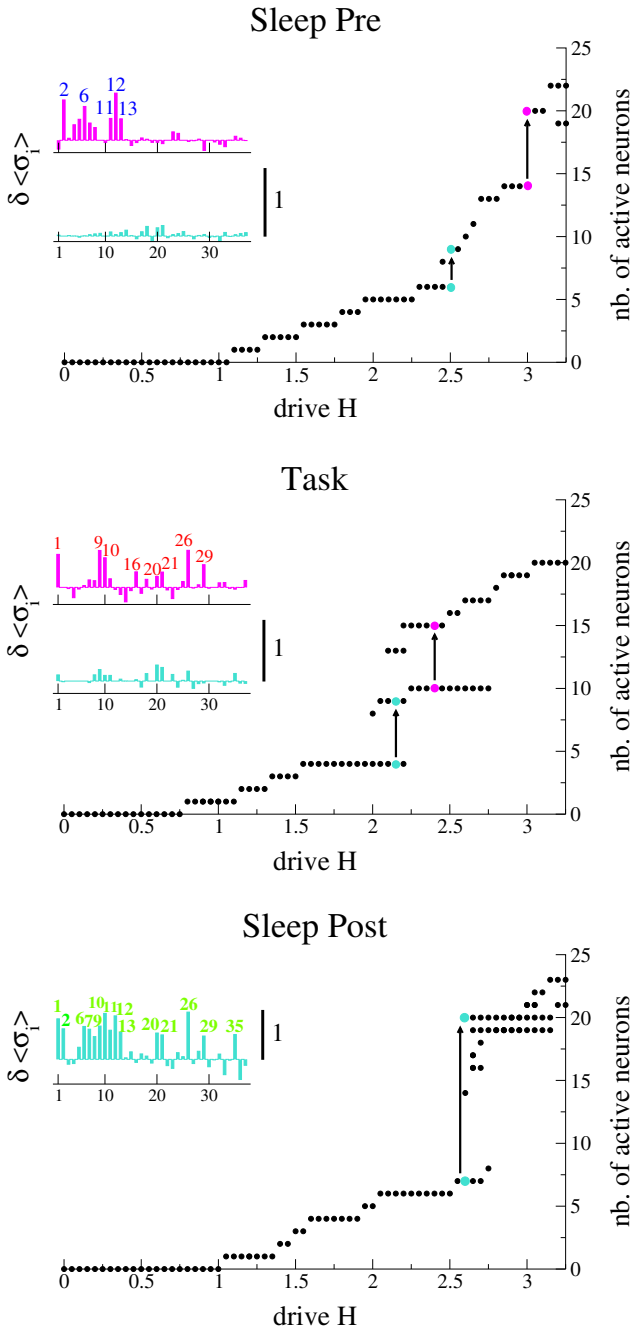

Fig. 5 Session 1: number of active neurons in the self-sustaining patterns of the distribution of activity configurations as a function of the drive H. Large changes in the number of active neurons between two self-sustaining configurations at the same value of H (*colored dots*) are shown by arrows. Variations of the neuron conditional average activities $\delta\langle\sigma_i\rangle$ between the low and the high activity configurations corresponding to each jump are shown in insets with the same colors. Labels indicate the neurons with the largest $\delta\langle\sigma_i\rangle$ and correspond with good approximation to the neurons with the largest χ_i^+ (compare with Fig. 3)

the cell assembly. We observe that the neurons with high $\delta\langle\sigma_i\rangle$ largely coincide with the ones with large susceptibility maxima, see Fig. 3.

The variations of the neuron conditional average activities $\delta\langle\sigma_i\rangle$ across the jumps between self-sustaining configurations are in close connection with the single-neuron susceptibility maxima, χ_i^+, and minima, χ_i^-. Figure 6 shows that neurons with the highest susceptibility maxima (lowest susceptibility minima) largely coincide with neurons having the largest (smallest) $\delta\langle\sigma_i\rangle$ in the jumps. Correlation between χ_i^- and negative $\delta\langle\sigma_i\rangle$ is less strong than correlation between χ_i^+ and positive $\delta\langle\sigma_i\rangle$, indicating that the more accurate analysis of the susceptibilities is required to precisely characterize inhibited neurons. It is worth noting that the point at which the susceptibility peaks cross the $\chi_i = 0.25$ line (representing the reference for independent neurons) corresponds to $\delta\langle\sigma_i\rangle = 0$, that is to independent neurons in the self-sustaining configuration analysis.

3.4 Can cell assemblies arise in randomly-connected neural networks?

To assess the statistical significance of the cell assemblies above we have performed control simulations on random connectivity models. More precisely, we have performed 100 random permutations of the Task couplings of session 1 and for each new coupling set we plotted the number of active neurons in the self-sustaining patterns as a function of drive H. None of the reshuffled datasets gave results similar

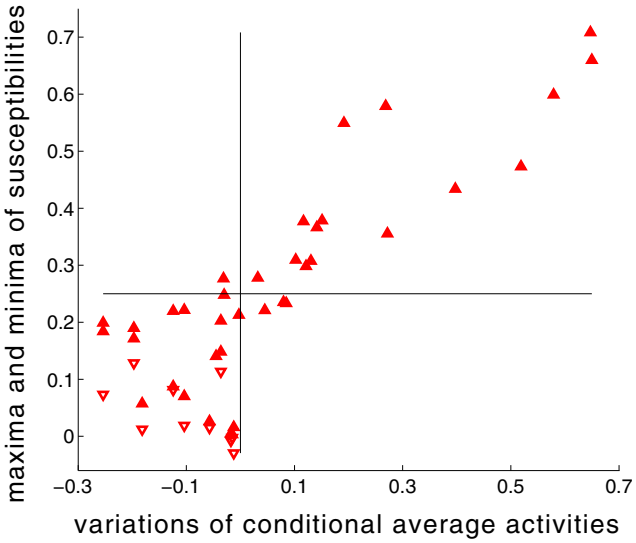

Fig. 6 Scatter plot of the maxima χ_i^+ (upward full triangles) and minima χ_i^- (downward empty triangles) of single-neuron susceptibilities vs. variations $\delta\langle\sigma_i\rangle$ of the neuron conditional average activities in the second jump of Fig. 5 (middle panel). Horizontal and vertical lines show the values expected for independent neurons, that is, respectively, $\chi_i = 0.25$ and $\delta\langle\sigma_i\rangle = 0$. Data correspond to the Task epoch of Session 1. Similar results are obtained for the other epochs and sessions

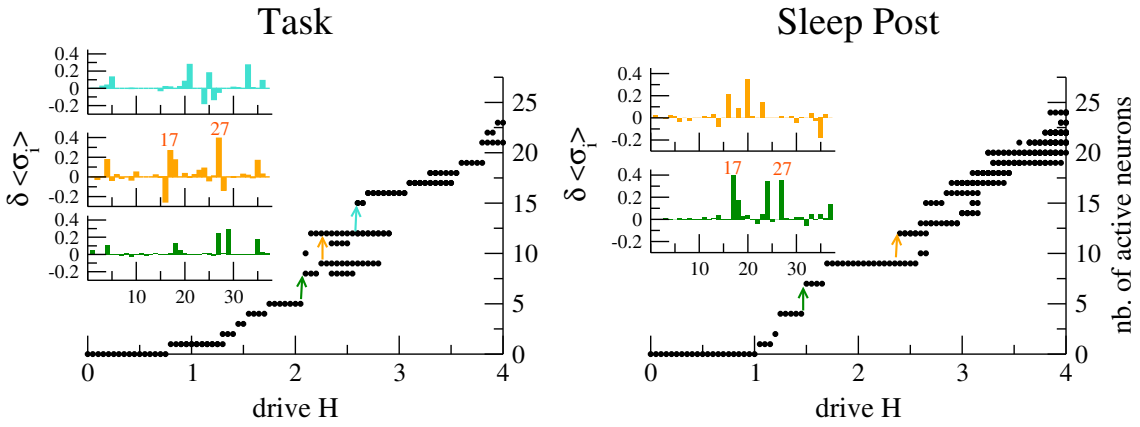

Fig. 7 Number of active neurons in the self-sustaining configurations vs. drive H for two datasets of reshuffled couplings obtained from the Task and Sleep Post epochs of session 1. Each dataset results from a random permutation of the couplings inferred in each epoch; the two datasets illustrated here have the largest jumps in the Task relative to a sample of 100 random datasets. The jumps in the number of active neurons and the variations $\delta\langle\sigma_i\rangle$ of the neuron conditional average activities (insets) show that only very little cell assemblies ($\lesssim 3$ neurons) and no significant replay group (at most 2 neurons, e.g. 17-27) are found in a set of 100 sessions with random coupling networks

to the ones of the original dataset: no jump or only very little jumps ($\lesssim 3$ neurons) were found in the reshuffled datasets. This confirms that the original coupling network contains groups of strongly interconnected neurons that have very little chance to be found in random networks. We then selected the 10 reshuffled datasets having the largest jumps in Task as our best candidates for the detection of replay groups. We then performed the same permutations of the couplings in the Sleep Post epoch, and looked for cell assemblies and replay groups in these candidate datasets. Across all candidates, 2 neurons at most were coactivated in both the Task and Sleep Post reshuffled datasets. Figure 7 shows the results obtained for the 'best candidate' dataset. Therefore, over 100 different random networks, no replay group with more than 2 cells is found. We conclude that the presence of a 5-cell replay group in session 1 (extracted from a dataset of $\sim$ 100 sessions) arises from specific Task-related modification of the Sleep Post network with respect to the Sleep Pre network. A similar reasoning applies to the 4-neuron replay group recorded in session 2.

3.5 Functional connectivity networks in the Bernoulli-GLM and Ising models are similar

To go beyond the stationary description of the Ising model distribution we introduced a Bernoulli-GLM with logistic link function (Truccolo et al. 2005; Gerwinn et al. 2010), see Methods. To test the goodness-of-fit of this model we compare the pseudo-log-likelihood of the data computed with the inferred model to the distribution of pseudo-log-likelihoods of spatio-temporal configurations generated by the inferred model with MCMC simulations. A spatio-temporal pattern is here defined as a set of activity variables $\{\sigma_{it}\}$ with $i = 1, \ldots, N_{tot}$ ($N_{tot} = 37$ = number of

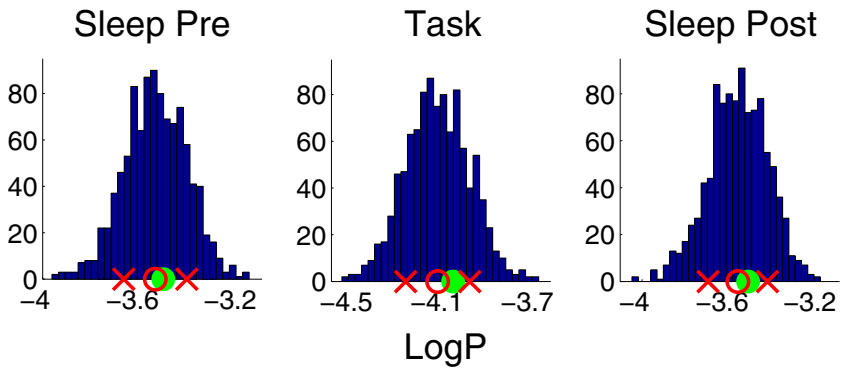

Fig. 8 Distributions of the pseudo-log-likelihoods of spatio-temporal patterns generated by the inferred models, and comparison with spiking data. Histograms represent the distributions of $\log P$ per time bin for the configurations generated by Monte Carlo simulations of the models inferred from the Sleep Pre (*left*), the Task (*middle*) and the Sleep Post (*right*) epochs. Green full dots represent $\log P$ of the recorded spiking data (*per time bin*) with the respective epoch model; red empty dots (*red crosses*) indicate the mean (mean $\pm$ one standard deviation) of $\log P$ values (*per time bin*) of configurations generated by the models

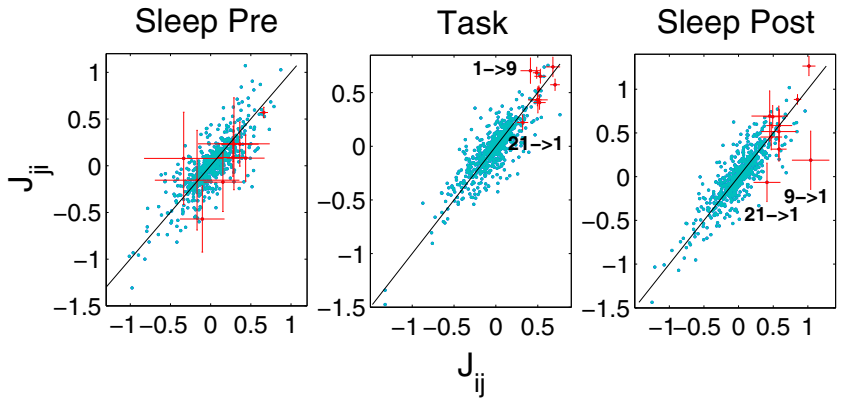

Fig. 9 Scatter plots of couplings J_{ij} vs. couplings J_{ji} in the three epochs of session 1. Black line: diagonal; red dots: couplings between the neurons of the replay group (1-9-20-21-26). For clarity of the figure, error bars are shown only for the replay group couplings. Almost all couplings have $J_{ij} - J_{ji}$ compatible with zero. In the Sleep Post, couplings 1-9 and 1-21 of the replay group are significantly asymmetric; in the Task, 1-9 is also slightly asymmetric, but with opposite sign, while 1-21 is basically symmetric

recorded neurons) and $t = 1, \ldots, T_{tot}$ ($T_{tot} = 640$, in time bins Δt of 5 ms). At each iteration of the MCMC, we update a variable chosen randomly in the full spatio-temporal pattern according to the Metropolis acceptance rule (enforcing detailed balance): if the variable flip causes a positive variation of $\log P$ in Eq. (16), the flip is accepted; if the variable flip causes a negative variation $\Delta \log P$ of the pseudo-log-likelihood, the flip is accepted with probability $e^{\Delta \log P}$. For the sake of definiteness we choose periodic boundary conditions over time for each $N_{tot} \times T_{tot}$–dimensional pattern; this choice is justified by the fact that sampled configurations are substantially independent of the boundary condition, since time dependence in the model extends back to $T_2 = 32$ time bins, and we consider spatio-temporal patterns 20 times longer. Each set of $N_{tot} T_{tot}$ consecutive flip trials defines a Monte Carlo step; every 2^5 Monte Carlo steps (decorrelation time), we compute the pseudo-log-likelihood per time bin, $(\log P)/T_{tot}$, of the sampled pattern. The MCMC is halted after 2^{15} steps, collecting a total of $(2^{15}/2^5) = 1,024$ patterns. We plot the histogram of the pseudo-log-likelihoods per time bin corresponding to those patterns in Fig. 8. The figure shows results obtained for the three epochs of session 1, compared to the pseudo-log-likelihood of the real data (green dot). The latter agrees within one standard deviation with the average value of the pseudo-log-likelihood over the patterns sampled from the Bernoulli-GLM. This indicates that the Bernoulli-GLM is a good representation of the recorded data.

We first study the asymmetry in the inferred coupling matrix. For each pair of neurons ij, the Bernoulli-GLM defines two directional couplings, J_{ij} and J_{ji}, together with their error bars, which we can compare in each epoch. Results are shown, for session 1, in Fig. 9. Almost all couplings are equal in the two directions up to the statistical error bars. However, in Sleep Post, we identify two couplings of the 5-cell replay group which are significantly

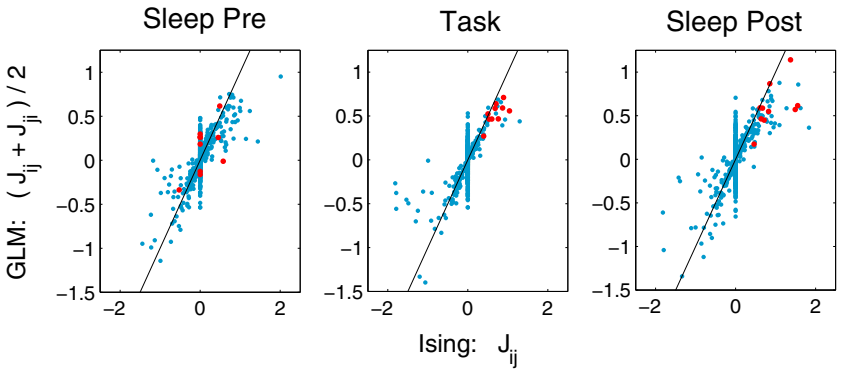

Fig. 10 Scatter plots of $(J_{ij} + J_{ji})/2$ inferred with the Bernoulli-GLM vs. J_{ij} inferred with the Ising model for the three epochs of session 1. Black line: diagonal; red dots: neurons of the replay group. Both sets of couplings are very correlated. The GLM couplings tend to be slightly smaller in absolute value than the Ising couplings, since their effect is summed over a neuron past spiking history. Note that the Ising coupling network is sparser than the GLM network, as a result of the ACE procedure

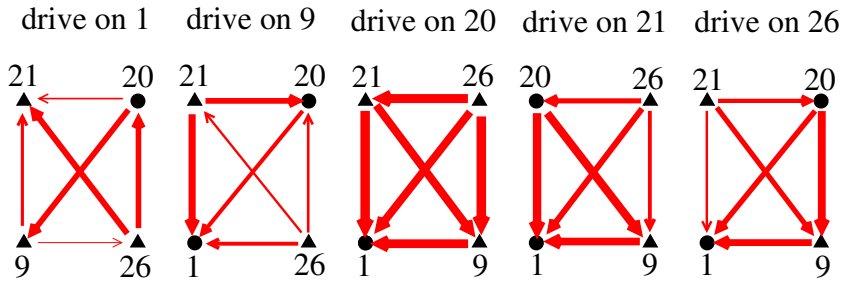

Fig. 11 Networks representing the most frequent activation order of the non-stimulated neurons in the spatio-temporal patterns obtained from simulations of the B-GLM with oscillatory drive on one neuron at a time (indicated in the title). Width of the arrows is proportional to $|AR|$ and the direction of the arrows is given by the sign of AR. Triangles denote pyramidal neurons, circles stand for undetermined cell types

asymmetric: couplings 1-9 and 1-21. In Task, 1-9 is also slightly asymmetric, but with opposite sign, while 1-21 is basically symmetric. Asymmetric couplings are potentially related to the sequential activation ordering of neurons in an assembly, see next Section. Figure 10 shows that the average between the two couplings inferred with the Bernoulli-GLM for each pair of neurons is in good agreement with the respective couplings inferred with the Ising model.

3.6 Distribution of temporal patterns of activity

The Bernoulli-GLM can be used to test the presence of favored temporal activation sequences in the coactivated groups unveiled with the Ising model under a drive (we focus here on the activation order of neurons in the replay group of session 1, identified in (Tavoni et al. 2015) and confirmed by the analysis of the neuron susceptibilities). We have sampled the spatio-temporal patterns of the Bernoulli-GLM inferred from the Sleep Post epoch data, in the presence of an external oscillatory drive acting on a single neuron of the replay group. More precisely, we apply to each neuron at a time a square pulse through an additive contribution to Eq. (15). This driving pulse is equal to H for $8\Delta t = 40$ ms, followed by no input for the subsequent $40\Delta t = 200$ ms; the stimulation is repeated 13 times. A MCMC simulation is then run for this Bernoulli-GLM under periodic drive. The square pulse drive mimics the robust activation of one neuron and allows us to study if the ordered chain of activation of replay neurons which follows depends more on the directionality of the GLM couplings or on intrinsic

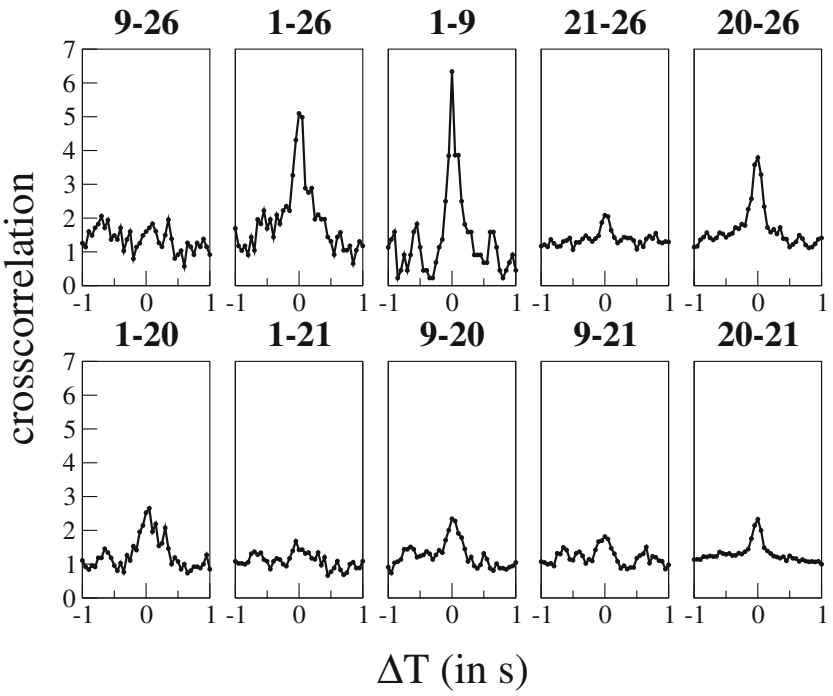

Fig. 12 Cross-correlation histograms between neurons of the 5-cell replay group of Session 1 (1-9-20-21-26) in Sleep Post. Time windows of histograms: 50ms

neuron excitability. Periodicity of the square pulse drive is imposed simply to sample several spatio-temporal patterns and to study their ordering statistics.

From the sampled spatio-temporal patterns of activity we have computed the conditional probabilities $p_{ji}(\tau)$ that neuron j spikes at time $t+\tau$ (in units of 5 ms time bins) given that neuron i is active at time t, averaged over all the time bins t in the spatio-temporal patterns. We have then computed the Asymmetry Ratios (AR) through

$$AR_{ji} = \frac{\sum_{\tau=-40}^{-1} p_{ji}(\tau) - \sum_{\tau=1}^{40} p_{ji}(\tau)}{\sum_{\tau=-40}^{-1} p_{ji}(\tau) + \sum_{\tau=1}^{40} p_{ji}(\tau)} \tag{21}$$

AR_{ji} is a measure of the asymmetry between the probability for j to spike in a time window between 5 ms and 200 ms before a spike of neuron i and the probability for j to spike in a time window of the same duration after a spike of neuron i. The larger the probability for the neurons in the pair to spike in a precise temporal order, the larger the absolute value of AR; note that $AR_{ij} = -AR_{ji}$.

In practice AR values are found to be small (lower than 0.2 in absolute values), and statistically compatible with zero for most (but not all) the pairs. Figure 11 shows all AR values and their signs as arrows (see caption) among the 4 non-stimulated neurons of the 5-cell replay group. Though most of the $|AR|$ are small, a tendency for each neuron to activate early or later in the firing patterns emerges from Fig. 11. Neuron 26, which has the largest Ising couplings with all the other neurons, tends to be activated first, i.e. it is the first neuron to respond to the stimulation of the driven neuron. The other three non-stimulated neurons simply tend to activate in order of their decreasing firing rates. In conclusion, the prevalence of one order of activation with respect to the others is weak and does not necessarily reflect a conditional dependence between the neurons resulting from the (small) asymmetries in the GLM-couplings but rather reflects the neuron average firing rates (similarly to the finding of (Peyrache et al. 2010a)) and, to a lesser extent, the strength of the undirected (Ising) couplings. This results extends what can be directly inferred from the pairwise cross-correlation histograms of Fig. 12, which do not show any clear asymmetry in the firing order of pairs of neurons.

Our study indicates that, differently from what reported in the hippocampus and in neocortical areas (Luczak et al. 2007), like visual, motor, somatosensory and parietal cortices (Ikegaya et al. 2004; Qin et al. 1997; Hoffman and McNaughton 2002; Ji and Wilson 2007), and even in a study of medial PFC neurons (Euston et al. 2007), replay events in this experiment are highly synchronous reactivations of strongly interacting neurons, which probably carry non-sequential information about the Task.

4 Discussion

The discussion section is mainly devoted to the meanings and implications of the drive H. General conclusions are given at the end of the section.

4.1 Sampling of rare coactivation events

The presence of a drive shifts the log-likelihood of a neural configuration $\boldsymbol{\sigma}$ by $H\sum_i \sigma_i$, as can be seen from Eqs. (1) and (6). In other words, the probabilities of neural configurations $\boldsymbol{\sigma}$ with a large number of active neurons $N(\boldsymbol{\sigma}) = \sum_i \sigma_i$ are boosted by a multiplicative factor, increasing exponentially with $N(\boldsymbol{\sigma})$. The same effect can be *a priori* obtained in a much simpler way, without inferring any Ising model. We start from the histogram of snapshots of the neuron activities (in a, say, 10 ms time bin) computed from the spiking data,

$$P_{data}(\boldsymbol{\sigma}) = \frac{B(\boldsymbol{\sigma})}{B}\,, \tag{22}$$

where $B(\boldsymbol{\sigma})$ is the number of time bins where the neural configuration is $\boldsymbol{\sigma}$, and B is the total number of time bins. We then introduce the biased histogram:

$$P_{data}(\boldsymbol{\sigma}|H) = \frac{1}{Z_{data}(H)}\, P_{data}(\boldsymbol{\sigma})\, e^{H\,N(\boldsymbol{\sigma})}\,, \tag{23}$$

where $Z_{data}(H)$ ensures normalization of the distribution. We have computed the single-neuron susceptibilities for several values of H using this formula, and compared them to those obtained from MCMC sampling of the Ising distribution in the presence of the same drive H, with the results shown in Fig. 13, top, for 2 representative neurons of the replay group (9 and 21) and another neuron (14), substantially independent of the others. The susceptibilities obtained from the biased histogram (23) are shown in red, and overlap with those obtained from the rescaled Ising distribution, shown in black, just for very low values of the rescaling parameter H (typically, $\lesssim 1.5$). For larger values the behaviours of the susceptibilities computed from the biased histogram and from the Ising model depart from each other, especially in strongly interacting neurons such as those of the replay group; only independent neurons have similar shaped curves.

This comparison highlights that firing events important for cell-assembly activations, which are unveiled by the addition of a drive in the inferred Ising model, are very rarely found in the recorded activity (on the small time scale Δt). In (Ganmor et al. 2011a; 2009), the authors notice that the Ising model, defined from pairwise correlations only (which can be reliably estimated from a relatively short recording), gives a better prediction of network activity patterns than what can be achieved by empirical sampling.

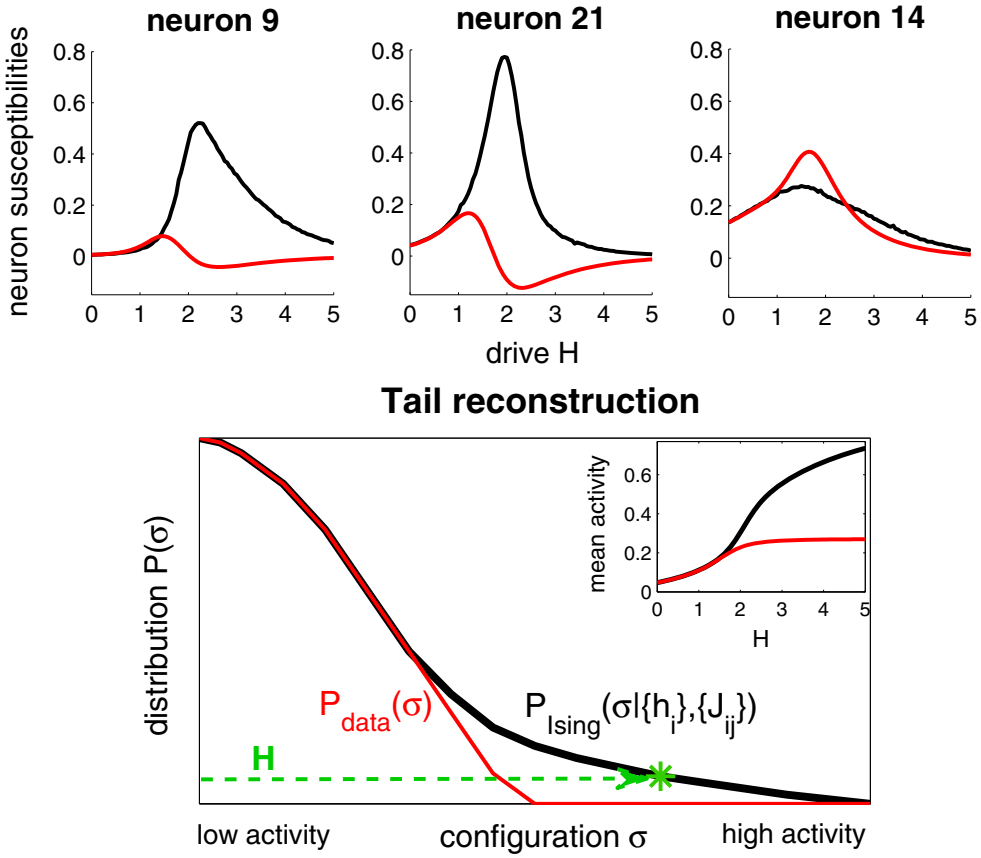

Fig. 13 The model distribution in the presence of a drive is able to detect rare coactivated configurations. Top: single-neuron susceptibilities computed from the Ising model distribution in the presence of the global drive H (*black*), compared to the corresponding susceptibilities computed from the reweighted data distribution Eq. (23) (red). Susceptibilities of three representative neurons (9 and 21 belonging to the replay group, 14 substantially independent of the other neurons) are shown for session 1, Sleep Post epoch; the reweighted data distribution Eq. (23) cannot be used to detect cell assembly neurons. Bottom: In contradistinction with the reweighted data distribution (*red*), the Ising distribution Eq. (6) (*black*) reconstructs accurately the tail of the distribution of all activity patterns potentially generated in such a network. Inset: mean global activity as a function of drive H, computed from MCMC sampling of distribution Eq. (6) in the presence of H (black), and from the reweighted data distribution (*red*); curves refer to the Sleep Post epoch of session 1. The red curve saturates at ~ 0.25, corresponding to the maximal activity configuration observed in the data on a $\Delta t = 10$ ms time scale, while the black curve keeps growing as H increases. Note that, for values of H necessary to uncover coactivation of replay neurons ($H > 2.1$), the two curves do not coincide

Here we push this consideration further: inferring the interaction network of the data not only allows us to have an accurate estimate of the distribution of activity patterns in the conditions at which the network is inferred, but also to reconstruct activity patterns generated by this effective network at different time scales and in the presence of a real input. Drive H pushes the network into a state of increased neuronal excitability and triggers avalanche processes of neuron activations that can occur in the real system (Fig. 1). The ability of the Ising model to capture these network effects is schematically illustrated in Fig. 13, bottom: the inferred model reconstructs the tail of the distribution of all configurations potentially observable in the system; rare and high-activity configurations are accessible when the distribution is biased by a drive H. In this sense, a cell-assembly activation emerges from the structural properties of the pairwise interaction network. In Appendix C we show that our method can detect coactivation events with a resolution as high as the minimal coactivation frequency ascribable with certainty to effective pairwise correlations (beyond the apparent ones due to sampling noise).

4.2 Time scales

While introducing a global drive H in the model allows us to reconstruct the tail of the configuration distribution and eventually to unveil cell-assembly activations, the practical significance of the drive H in real neural systems remains to be clarified. From CoA estimates we observe that cell-assembly activations take place over time scales larger than

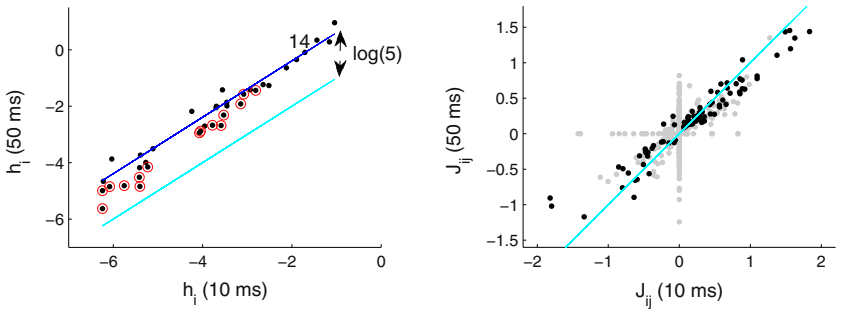

Fig. 14 Scatter plots of inputs h_i and couplings J_{ij} inferred with 10 ms time bins vs. 50 ms time bins, in the Sleep Post epoch of session 1. Cyan lines represent diagonals. Local inputs h_i of nearly independent neurons (e.g. neuron 14) undergo approximately a uniform logarithmic increase with the time-bin ratio ($h_i^{(50)} \sim h_i^{(10)} + \log\left(\frac{50}{10}\right)$, blue line), while inputs of coactivating cells (*red circles*) tend to increase less than predicted by the independent neuron approximation. Reliable couplings in the two sets ($|J_{ij}|$ larger than three times their statistical standard deviations, black dots) do not show a systematic dependence on the time bin. These results hold for a wide range of time-bin widths $\Delta t^{(2)} > 10$ ms

10 ms, which is the time-bin width at which firing rates and pairwise correlations are computed to infer the Ising model parameters. It is thus reasonable that the drive H could represent, at least partially, this time-scale gap.

For a system of independent neurons with firing rates f_i, we can compute with very good approximation the relation between H and the time scale. The values of the local inputs h_i in an Ising model (with zero couplings) are related to the

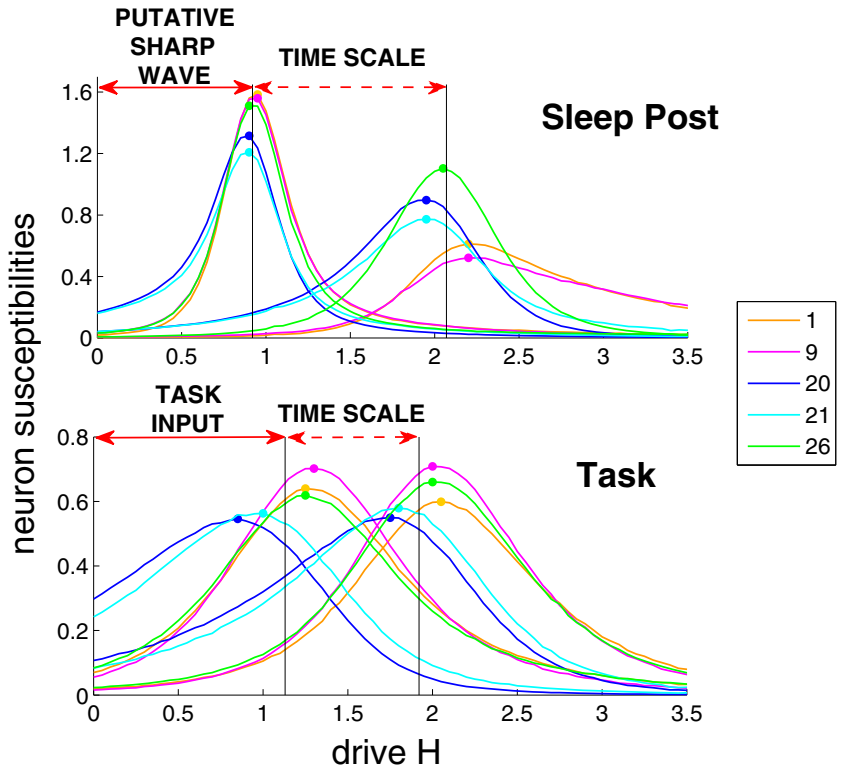

Fig. 15 Susceptibilities of replay neurons of session 1 (Sleep Post and Task phases), computed through MCMC sampling of the Ising distribution (in the presence of the global drive H) with parameters inferred with 10 ms time bins (*right curves*) and time bins corresponding to the time scale of neuron coactivation in the data (*left curves*): 40 ms for the Sleep Post and 30 ms for the Task. Each color identifies a neuron (*see legend*). Vertical bars indicate the mean value of H_i^+ of the 5 neuron susceptibility peaks relative to the two time-bin widths. Increase in the time bin results in susceptibility shift towards smaller values of H (*dashed red arrows*). A global drive of $\sim$ 0.9 in Sleep Post and $\sim$ 1.1 in Task (*full red arrows*) has yet to be applied to see susceptibility peaks, even if the inference of Ising parameters is done at the same time scale as the one of coactivation of the replay group in the spiking data. We believe that this residual part of the drive represents a real input necessary to trigger activation of the cell assembly

probabilities $p_i(\Delta t)$ that each neuron i is active in a time bin of width Δt, computed from the spike recordings, through

$$p_i(\Delta t) \sim f_i \cdot \Delta t = \frac{e^{h_i}}{1+e^{h_i}} \tag{24}$$

from which we obtain

$$h_i \simeq \log(f_i \cdot \Delta t) \tag{25}$$

for small Δt: local inputs h_i vary logarithmically with Δt. If we change the time-bin duration from $\Delta t^{(1)}$ to $\Delta t^{(2)}$ the local inputs undergo a uniform shift:

$$h_i^{(2)} - h_i^{(1)} \sim \log\left(\frac{\Delta t^{(2)}}{\Delta t^{(1)}}\right) \tag{26}$$

Therefore, in a system of independent neurons, adding a uniform drive H to each local input h_i (inferred with time bins of width $\Delta t^{(1)}$) would be equivalent to rescaling the time-bin width by a factor e^H.

For real recordings, where neurons are not independent, we observe the presence of deviations from the uniform logarithmic increase of the inferred local inputs h_i with the time-bin ratio. While the inferred interaction structure is substantially stable with changes in the time-bin width, as expected (Fig. 14, right), inputs h_i of nearly independent cells (e.g. neuron 14) undergo the shift predicted by the independent neuron approximation, Eq. (26), and inputs h_i of coactivating neurons increase less than predicted by Eq. (26), see Fig. 14, left. The reason is that the activity of these neurons is mainly driven by the couplings with other cells rather than their spontaneous firings and Eq. (24) becomes:

$$f_i \cdot \Delta t \sim \left\langle \frac{e^{h_i + \sum_j J_{ij}\sigma_j}}{1+e^{h_i+\sum_j J_{ij}\sigma_j}} \right\rangle \tag{27}$$

where the average is calculated over the activities σ_j of the other neurons. For small Δt, the argument of the exponential must be large and negative, and the denominator in the above formula can be approximated to unit. In addition, if we replace the average value of the exponential with the exponential of the average value (the so-called mean-field approximation), we simply obtain

$$h_i^{(2)} - h_i^{(1)} \sim \log\left(\frac{\Delta t^{(2)}}{\Delta t^{(1)}}\right) - \sum_j J_{ij}\left(\langle\sigma_j\rangle^{(2)} - \langle\sigma_j\rangle^{(1)}\right) \tag{28}$$

where $\langle\sigma_j\rangle^{(2)} - \langle\sigma_j\rangle^{(1)}$ is non negative. The shift in the local inputs tends thus to be smaller than the logarithm of the time-bin ratio. In general, the most negative h_i show the largest deviations from the logarithmic behaviour: this suggests that interaction effects are, on average, stronger for neurons with low firing rates than for neurons with high firing rates, see also Section 4.4.

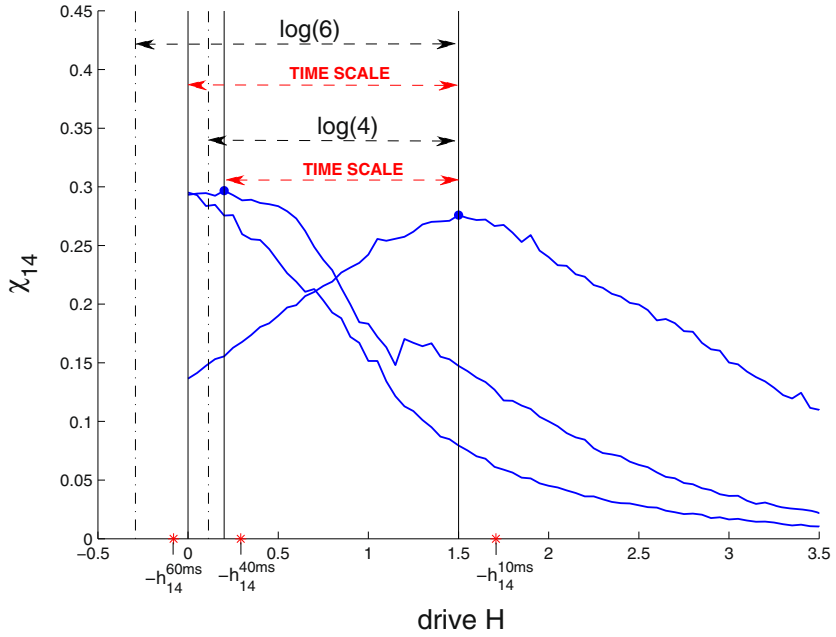

Fig. 16 Addition of drive H may be interpreted as a time-scale increase for independent neurons. Susceptibility of neuron 14, in the Sleep Post epoch of session 1, as a function of the global drive H, computed through sampling of the Ising distribution inferred with $\Delta t = 10$ ms (*right curve*), 40 ms (*middle curve*) and 60 ms (*left curve*) time bins, the last one corresponding to the inverse of the firing rate of neuron 14. Susceptibility peaks are just above 0.25 and their H_i^+ (*full vertical bars*) are close to the opposite of the local input h_i (indicated by a red star for each time-bin width), see Methods. Differences in the time bins result in susceptibility shifts (red dashed arrows) well accounted for by the shift $\sim \log(\Delta t^{(2)}/\Delta t^{(1)})$ (*black dashed arrows*) expected for an independent neuron, see Eq. (26)

4.3 Magnitude of the input

We have estimated the magnitude of the real input required to elicit activation of the replay group (1-9-20-21-26) of session 1, in the following way. We have computed the susceptibilities of replay neurons through MCMC sampling of the Ising distribution (in the presence of drive H) inferred with time bins equal to the time scale on which coactivation of those neurons is observed in data. Neuron susceptibilities would be peaked around $H = 0$ if neurons were *on average* coactivated on that time scale. This is not the case, and the susceptibility peaks are located at $H > 0$ both in Task and in Sleep Post, see Fig. 15. The drive H corresponding to those susceptibility peaks gives an estimate of the real input necessary to trigger transient coactivations of replay neurons.

As a control, we have plotted in Fig. 16 the susceptibilities vs. H for neuron 14 (almost independent) in Sleep Post, for three different values of the time bin in the inference: 10 ms, 40 ms and 60 ms, the last one being the typical time scale of neuron activation computed from data ($1/f$). The susceptibility peak is progressively shifted to the left by $\sim \log(\Delta t^{(2)}/\Delta t^{(1)})$ (see caption for more details). For $\Delta t = 60$ ms the maximum value of the susceptibility is for $H = 0$, confirming that H has simply the meaning of a time scale for independent neurons.

On the contrary, values H_i^+ of the susceptibility peaks of the replay neurons, obtained sampling the Ising model inferred with $\Delta t = 10$ ms, account for both a time-scale increase from 10 ms to 40 ms in Sleep Post (or 30 ms in Task), that is the time scale of group coactivation in the data (peak of the CoA), and a real input which triggers the coactivation event, according to the avalanche process illustrated in Fig. 1.

In Sleep Post this input ($H \sim 0.9$) could represent sharp waves coming from hippocampus: indeed, coactivation of replay neurons is observed during SWS periods, when sharp-wave ripples reach prefrontal cortex. Sharp wave-related inputs could also be present when the rat is awake, i.e. in the Task epoch (Roumis and Franck 2015), but in periods of rest, and could be related to the position of the rat in the maze. We show in Fig. 17 the frequency of coactivation P_{CoA} of the replay group in session 1 as a function of the location of the rat in the maze for the three sessions

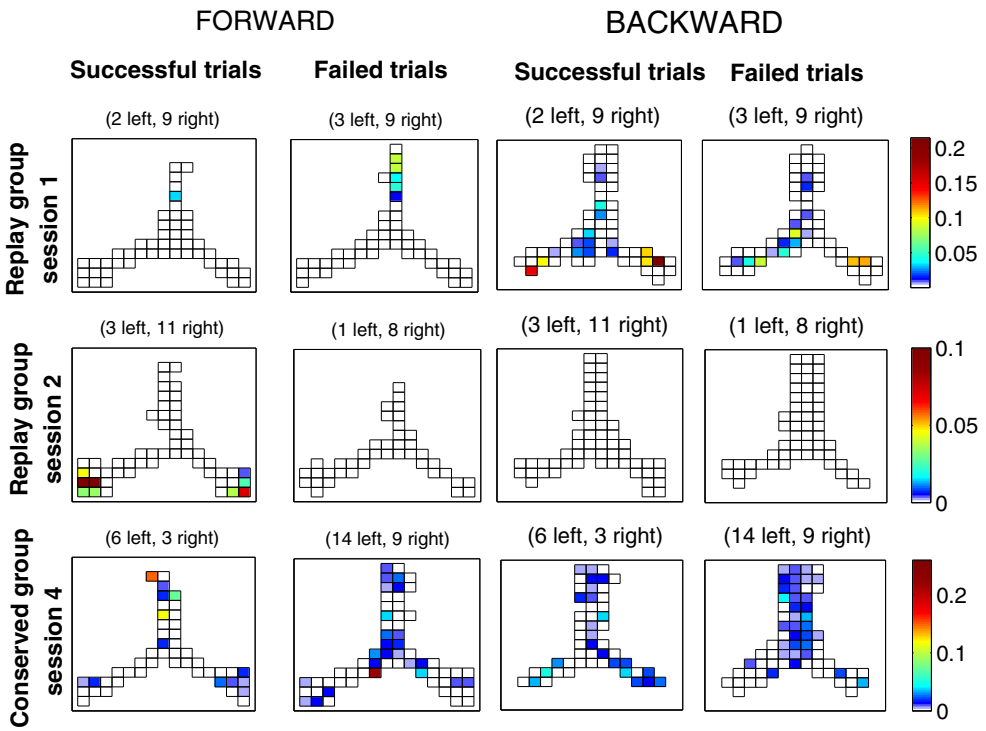

Fig. 17 Heat map of the probability P_{CoA} (computed as the number of coactivation events in each patch divided by the number of times the rat crosses that patch) that the replay group in sessions 1 and 2 (top and middle panels) and the conserved group in session 4 (*bottom panels*) is coactivated in each patch of the maze crossed by the rat (*black rectangles in every panel*). We consider coactivation events on a $\lesssim 100$ ms time-scale for the 5 neuron replay group of session 1, and on a $\lesssim 70$ ms time-scale for the 4 neuron replay group of session 2 and the 3–neuron conserved group of session 4. Results are shown for the forward path (first and second columns, corresponding to successful and failed trials respectively) and for the backward path (third and fourth columns, corresponding to successful and failed trials respectively). The probability distribution of coactivation events for the replay groups is much more localized than for the conserved group and peaked at the final point in successfull trials, where the rat receives the reward

studied here. P_{CoA} for the replay group of session 1 is peaked at the beginning of the maze (starting point of the forward path) and more markedly close to the end of the maze (at the beginning of the backward path), where the rat stops (and sharp-waves can occur) and where P_{CoA} is stronger in successful trials; this suggests that the coactivation of these replay neurons could be related to reward. Localization of coactivation events at the end of the maze is also observed for the replay group of session 2 (second row of panels in Fig. 17): interestingly, this group is found to coactivate only at the final point and in all 14 successful trials, while it does not coactivate in any of the 9 failed trials. This result also holds for session L of (Tavoni et al. 2015) (not shown), which was recorded the day after session 2: replay groups of these two sessions (which probably have very close recorded neurons) could be activated by sharp waves during reward presentation. This observation suggests that PFC could play a role in the representation of reward, together with adjacent frontal areas, where reward-related activity has been reported (Chang et al. 2013). On the contrary, the probability distribution of coactivation events for the conserved group of session 4, which is coactivated in all epochs and can not carry new information about the task, is more spread all over the maze and in both successful and failed trials. More statistics would however be needed to draw a clear conclusion about reactivation and reward in the Task epoch.

4.4 Properties of coactivating neurons

For an independent neuron with firing rate f, the value H^+ of the drive at which the susceptibility is maximal is approximately given by $\log\left(\frac{1}{f\cdot\Delta t}\right)$, where $\Delta t = 10$ ms is the time-bin width, see Eqs. (11) and (25). In general, we observe an approximately monotonic relationship between the shift $H^+ - \log\left(\frac{1}{f\cdot\Delta t}\right)$ and the height of the susceptibility peak χ^+, see Fig. 18, left. Neurons with strong positive interactions with the other cells have negative shifts $H^+ - \log\left(\frac{1}{f\cdot\Delta t}\right)$, since positive interactions reduce the value of the input necessary for their activations.

This shift can be related to other neural properties, such as the neuron firing rates. Figure 18, right, shows that the shift tends to decrease in absolute value with the neuron firing rate. Our method is particularly powerful in detecting those rarely active neurons or, in other words, in catching the deviations from the average of the neuron activity, caused by network effects. Drive H can be seen as a global forcing on the system which allows the underlying interaction structure of the network to manifest itself, in the form of transient coactivation events that would not occur otherwise, due to neuron low firing rates.

A careful comparison of the properties of the replay neurons in Sleep Post and Task unveils a difference between the two epochs: in Sleep Post replay neurons are characterized

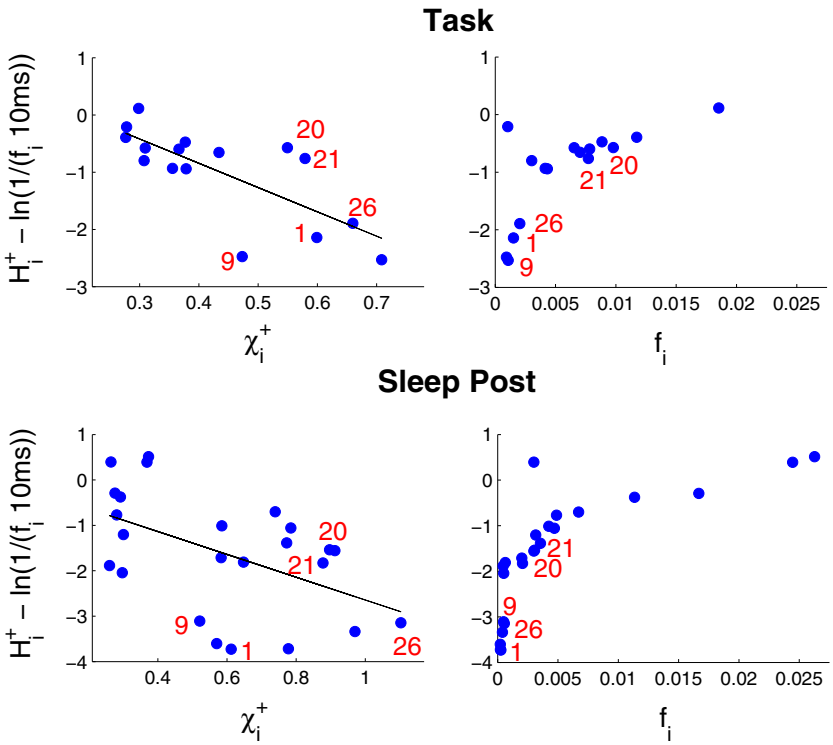

Fig. 18 Properties of coactivating neurons. Y-axis: difference between H_i^+ (value of the drive corresponding to the peak of the MCMC susceptibility, computed with $\Delta t = 10$ ms time bins in the inference) and $\log\left(\frac{1}{f_i\cdot\Delta t}\right)$ (value of the drive at the susceptibility peak for an independent neuron). X-axis: susceptibility peaks χ_i^+ (*left*); neuron firing rates (*right*). Only neurons with susceptibilities having peak values > 0.25 in the Task and Sleep Post epochs of session 1 are shown; red labels refer to neurons of the replay group

by larger shifts $H_i^+ - \log\left(\frac{1}{f_i \cdot \Delta t}\right)$ and lower firing rates compared to the Task epoch, particularly neurons 20 and 21. This is indicative of the fact that in Task coactivation events are less rare than in Sleep Post, which has a biological reason. In Task several coactivation events are needed to form new couplings important for encoding information, whereas in Sleep Post few offline pattern reactivations are necessary to consolidate the newly formed couplings.

4.5 Conclusion

The present study builds on and extends a recent work (Tavoni et al. 2015), where we identified putative groups of neurons involved in task-related memory consolidation from the comparison of the inferred functional networks across different epochs of the same session. Here, groups of coactivating neurons are found based on the sampling of the inferred distribution P of neural activities in the presence of a drive. This drive mimics, in a very approximate way, the inputs to the network, or change in excitability of the network across time. We observe that groups of neurons activate together (for the same value of the drive) in the model, and that those neurons are found to coactivate in the spiking data. While sampling the distribution of activity is justified by the fact that real neuronal networks are noisy (due to e.g. ion channels, synaptic transmission and inputs from the surrounding environment, including non-recorded neurons), we show that a simpler, faster computational approach focusing on the local maxima of P only is able to capture essentially the same information about neural coactivation.

Our Ising-model approach allows us to span efficiently the spectrum of coactivation patterns which can be generated in the inferred neural network. These cell assemblies can be specific to an epoch, or related to task learning, i.e. shared by the Task and Sleep Post epochs only. Our method can detect replay events, even when 'templates' (Johnson and Redish 2007; Pfeiffer and Foster 2013; Singer et al. 2013), as provided for example by the sequential activation of hippocampal place cells, are not available. Indeed our approach is based on the notion of simultaneous coactivation (Wilson and McNaughton 1994), irrespectively of temporal ordering aspects. As the Ising model gives the distribution of snapshots of the activity, couplings are symmetric and no ordering can be predicted by the model. Yet, the analysis illustrated in Sections 3.5 and 3.6, based on the inference and sampling of a Bernoulli-GLM with potentially asymmetric couplings, shows that asymmetries in the activation orders of the replay neurons are weak. It is plausible that information encoded in prefrontal cortex has aspects that are not inherently sequential, like the current rule used to solve the task. Even in hippocampal CA1, repeated synchronous activations of neurons in very narrow time windows (∼30 ms) have been observed (Harris et al. 2003). It is interesting that ∼30 ms is also the typical time-scale of coactivation that we observe for replay cell assemblies in prefrontal cortex, and roughly corresponds to a gamma cycle (Battaglia et al. 2011; Cardin et al. 2009). We believe that the inference of functional connectivity-based models for the distribution of the neural activity offers a more accurate approach than simpler statistical methods for cell–assembly detection (Appendix A).

The idea of exploiting maximum entropy models to study the neural code is not new. For instance, a similar exploration of the energy landscape and search for self-sustaining patterns of k-pairwise models was done in previous studies of the retina (Tkačik et al. 2014), to uncover the features of the neural code for visual signals. In that study, the parameter H was inferred from the data, together with the $N(N+1)/2$ parameters h_i and J_{ij}, and captured the 'global synchrony' or 'population firing rate' present in the observed activity patterns of the retina in response to a real visual stimulus. In the present study, however, we do not know what is the real stimulus (if any) feeding the system. The global drive H is a control parameter in our model, which we can tune at our convenience; its biological meaning and relevance were discussed in Section 4. In view of the complex nature of the neural activity of the rat and of the largely unknown inputs from other brain areas, inferring a stationary Ising model from the data may be considered as an oversimplification. It is therefore remarkable that, even if the observed cell assemblies derive from the presence of a particular and transient excitation of the recorded network, the Ising model is able to capture and reproduce those cell assemblies independently of their specific origin, and to identify the replay events likely to be at the basis of the memory consolidation process (Battaglia et al. 2011). In a forthcoming study, we will show how the use of detailed inferred models allows us, in addition, to quantify in an accurate way the reactivitation of Task-related activity over time in the Sleep epochs (Peyrache et al. 2009; Benchenane et al. 2010; Singh et al. 2015).

Acknowledgments This work is a follow-up of a previous study in collaboration with F.P. Battaglia and U. Ferrari (Tavoni et al. 2015), to whom we are very grateful.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

Appendix A: Comparison with Principal Component Analysis–based methods

In this section we compare our method to identify coactivated groups of neurons in data with techniques based on Principal Component Analysis (PCA). We consider again session 1, for which (Peyrache et al. 2009) have shown the presence of reactivation during Sleep Post of the first principal component of the Pearson Correlation matrix of the activity in Task. Despite this result, we show below that the identification of cell assemblies in single epochs is in general difficult with PCA.

We bin the neuron spike trains into time windows of 10 ms (as in our model inference) and 100 ms (as used by Peyrache et al. in their analysis) and compute the Pearson correlation matrix of the activity for these two choices of the time bin. For $\Delta t = 10$ ms we identify six signal (principal) components, and five for $\Delta t = 100$ ms. Our criterion for identification of the signal eigenvectors is as follows: we select the modes whose corresponding eigenvalues are larger than the upper bound of the Marcenko-Pastur eigenvalue distribution, $\lambda^+ = (1+\sqrt{N/B})^2$, where $B =$ number of time bins in the recording, $N =$ number of recorded neurons. The largest entries of the first component in the Task epoch correspond to the replay group 1-9-20-21-26, which is also represented (at least partially) in the two principal components of Sleep Post. This result explains the agreement between the replay group identified in (Peyrache et al. 2009) and in our analysis (see also (Tavoni et al. 2015)).

We have then tried to use clustering procedures to identify neural groups. We represent each neuron as a point in the space of the signal components, with coordinates $v_m(i)\sqrt{\lambda_m}$ (where $v_m(i)$ is the entry corresponding to neuron i in the m*th* signal eigenvector, and λ_m is the m*th* eigenvalue). We then apply the classical *k-means* clustering algorithm to these N points, where the number k of clusters is arbitrarily chosen, as we expect groups of closely correlated neurons to be represented by points far from the origin and close to each other in this dimensionally-reduced space. Unfortunately, this method applied to our data does not seem to be able to identify significant clusters, well separated from noisy clusters, as shown in Fig. 19, where the identified clusters are projected onto the bi-dimensional space of the first two signal components. Each panel shows the clustering (optimized over 10^4 random initial conditions) in each epoch, for a particular choice of $k = 2, 3, 4$. Neurons in the same cluster are represented

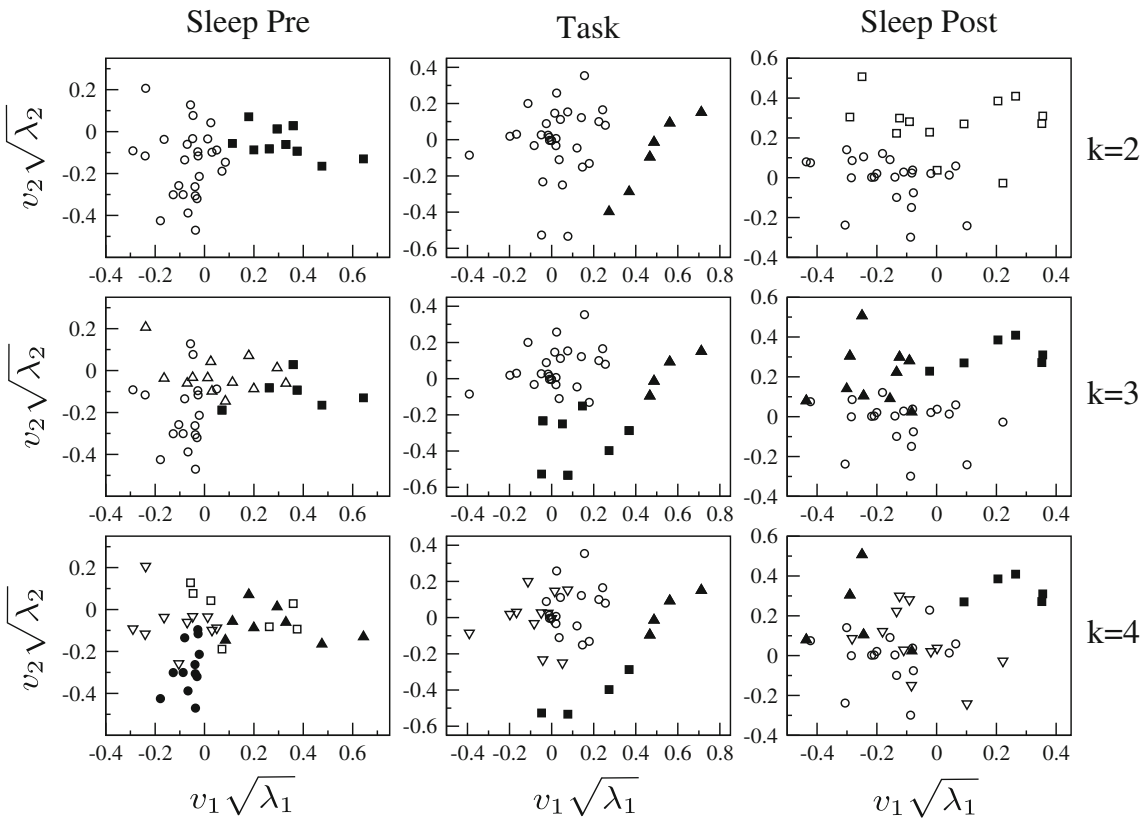

Fig. 19 Clustering of neurons in the space of principal components; projection onto the first two principal components is shown in each panel, for a particular choice of the number of clusters k, from $k = 2$ (*first row*) to $k = 4$ (*third row*). Different clusters are identified by different symbols, with full symbols for signal clusters ($d_c > 0.3$) and empty symbols for fully noisy clusters ($d_c < 0.3$)

by the same symbol; full symbols show the farthest clusters from the origin (distance $d_c > 0.3$), empty symbols correspond to the closest ones ($d_c < 0.3$). With a few exceptions, e.g. the cluster of upward full triangles in Task, which is rather robust with respect to the choice of k, signal clusters are in general not clearly separated from each other and from the noisy clusters. In addition no obvious choice for the value of k seems to be optimal to extract the groups of maximally coactivated neurons in each epoch. Finally, this method assigns each neuron to one cluster, and does not allow us to identify overlapping cell assemblies.

Appendix B: Analysis of sessions 2, 3 and 4

The analysis described in Sections 3.1 and 3.2 for session 1 is here illustrated for other three representative sessions (called 2, 3 and 4), with the purpose of validating the method and showing the phenomenology of coactivated groups that can be found in the data. We first compute the neuron susceptibilities as a function of drive H (Methods). For all sessions we find bell-shaped susceptibilities similar to those of session 1, with several susceptibilities showing maxima or minima comprised between 0 and 0.25 (indicating neurons which are substantially independent); some susceptibilities have maxima or minima at very close values of H, and correspond to neurons that are coactivated or coinhibited. For each epoch we rank the susceptibility maxima χ_i^+ and show, through computation of the coactivation ratio (17), that neurons having large χ_i^+ are strongly coactivated in the real data, thus validating our method to identify cell assemblies.

Session 2 has highly similar coactivated neurons in the Task and Sleep Post Task, and these neurons are not coactivated in the Sleep Pre Task (Fig. 20), similarly to the replay group found for session 1. This is in agreement with the identification of a replay group for this session from comparison of the inferred interaction networks in (Tavoni et al. 2015).

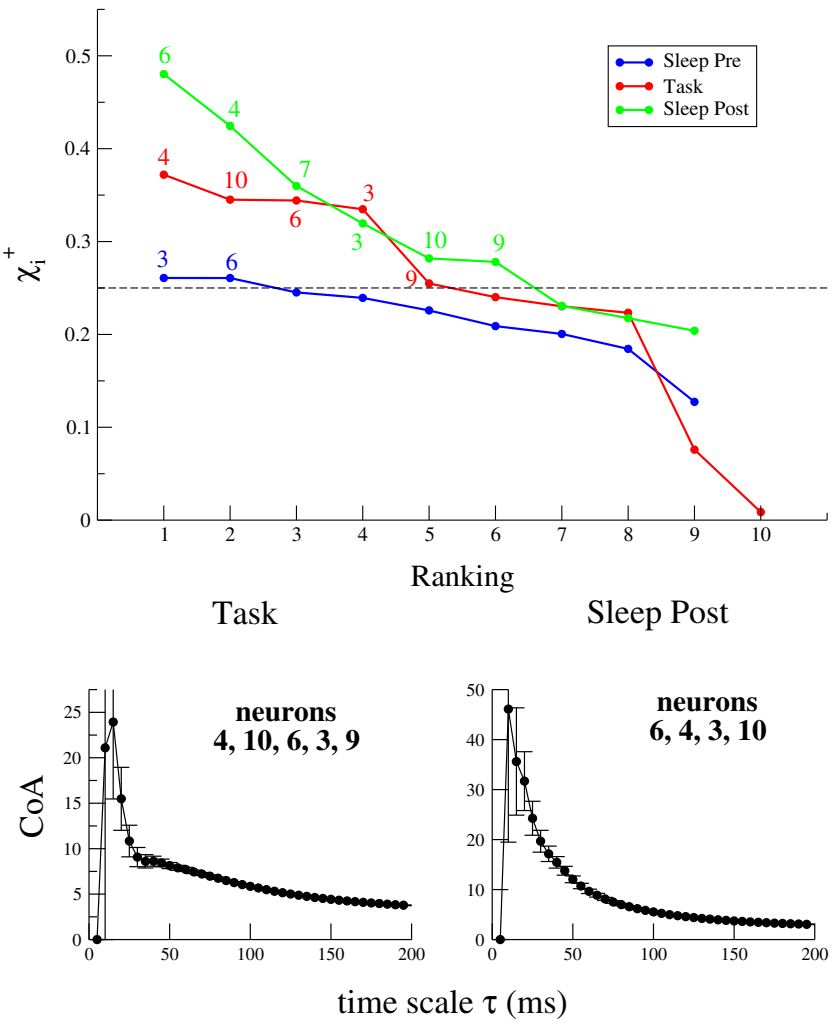

Fig. 20 Analysis of session 2. Top: Ranking of susceptibility maxima χ_i^+ in the Sleep Pre (*blue*), Task (*red*) and Sleep Post (*green*) epochs. Black dashed line indicates the threshold ($\chi_i^+ = 0.25$) for independent neurons and labels indicate neurons with $\chi_i^+ > 0.25$. Bottom: Coactivation ratio (Eq. (17)) for groups of neurons with top susceptibilities in Task and Sleep Post. Neurons 3-4-6-10 constitute a replay group

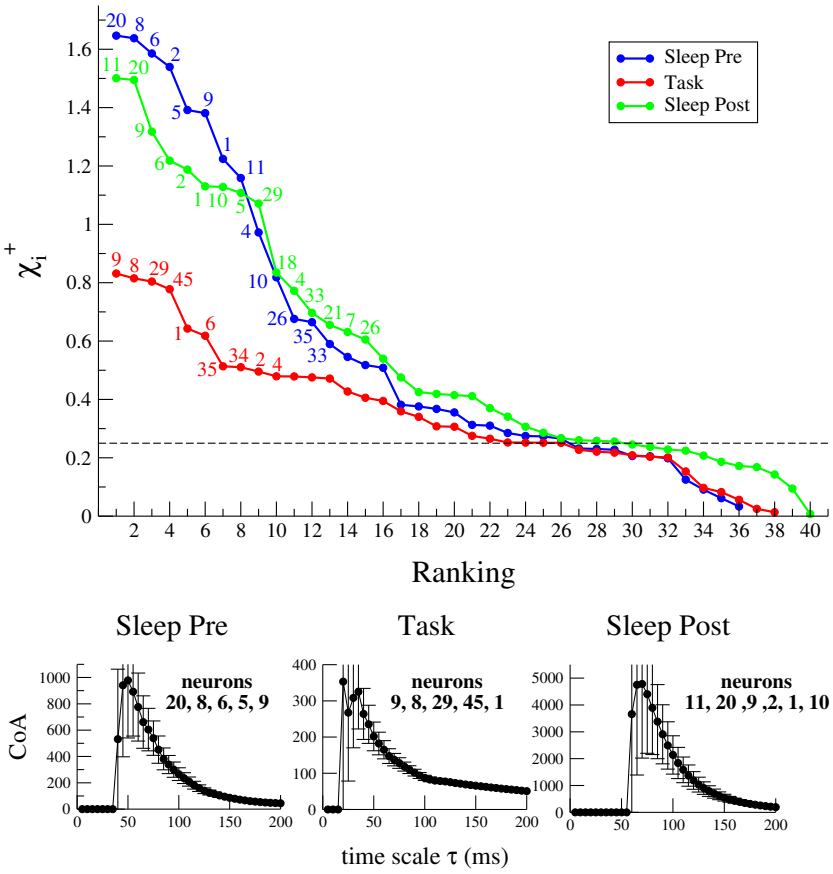

Fig. 21 Analysis of session 3, see caption of Fig. 20 for description

Differently from session 2, session 3 shows large susceptibility maxima also in the Sleep Pre Task (Fig. 21) indicating that cell assemblies have been sampled in all epochs, as confirmed by the large coactivation ratios of the identified groups. These groups are substantially different across the three epochs and no large replay group has been recorded in this session.

Session 4 is characterized by the presence of a group of neurons with high χ_i^+, large and positive couplings and a large coactivation ratio in all experimental epochs (Fig. 22).

Appendix C: Minimal number of coactivation events required for detection

The resolving power of our approach in detecting rare coactivation events can be estimated analytically. To fix notations we consider a set of N neurons, with uncorrelated spiking activities and average firing rate f. On top of this activity a simple firing pattern (spikes of neurons 1 and 2 separated by less than ΔT) is repeated K times; note that the choice of a pattern involving two neurons only can be considered as a worst case, as patterns involving a higher number of neurons would produce more correlations and would be easier to detect. The total recording time is T and the time-bin duration is ΔT. Within this framework we can compute (for simplicity we assume that K is small compared to the expected number of spikes, $f\,T$):

– the expected largest apparent correlation between any two neurons resulting from limited sampling in the absence of any pattern repetition ($K = 0$). This value can be estimated from extreme value theory to be (for $f\,\Delta T \ll 1$):

$$C_{noise-max} \simeq 2\,f\,\frac{(\Delta T)^{3/2}}{T^{1/2}}\sqrt{\log N} \tag{29}$$

– the minimal number of repetitions, K_{min}, such that the correlation between neurons 1 and 2 produced by the K repetitions exceeds $C_{noise-max}$:

$$K_{min} \simeq \frac{T}{\Delta T}\,C_{noise-max} \tag{30}$$

– the susceptibility curve of neurons 1 and 2 (identical) for K repetitions. To compute this curve, we first infer the Ising model for neurons 1 and 2 (all neurons being independent), which involves two local inputs $h_1 = h_2$ and one coupling J_{12}. The parameters can be analytically inferred as functions of K. We then compute

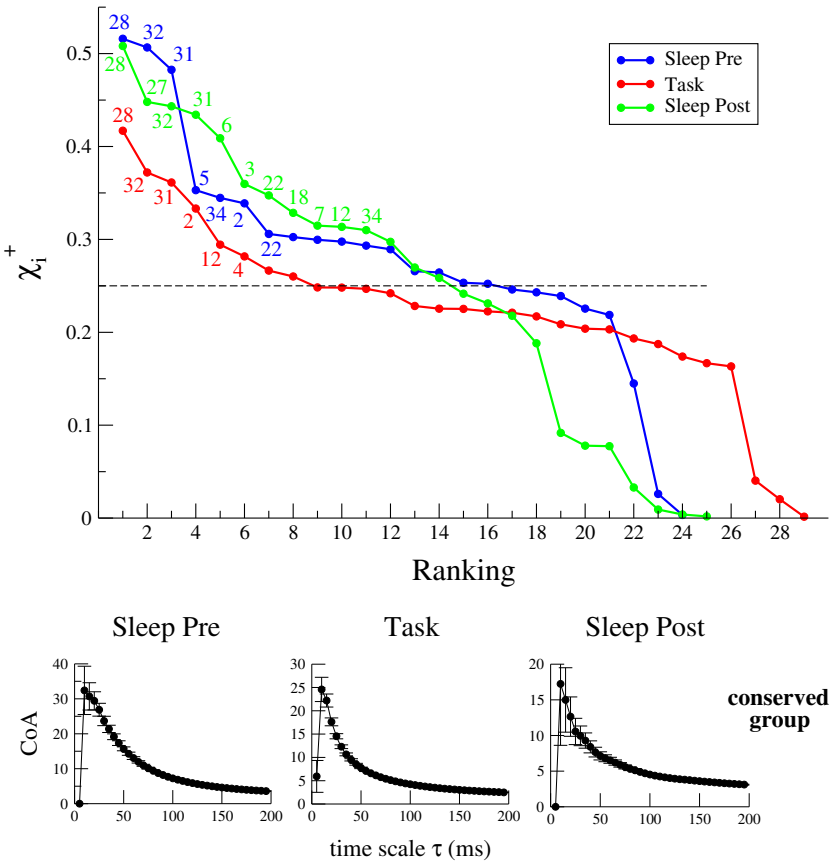

Fig. 22 Analysis of session 4, see caption of Fig. 20 for description. The bottom panel shows the coactivation ratio (Eq. (17)) for the 'conserved' group of neurons (28-31-32) with large susceptibility maxima in all three epochs

the average activity $\langle \sigma_1 \rangle(H)$ as a function of the drive H and its derivative with respect to H, that is, the susceptibility $\chi_1(H)$.

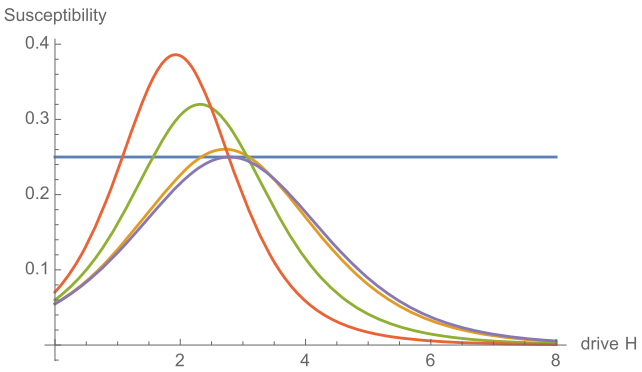

Fig. 23 Analytical model for computation of the minimum frequency of detectable coactivation events. For the parameters chosen here (T, f match the experimental data values) the minimum number of repetitions required for the correlation to exceed noise is $K_{min} \simeq 70$. The plot shows the susceptibilities $\chi(H)$ vs. H, for $K = K_{min}$ (*orange curve*), $K = 10\,K_{min}$ (*green curve*), $K = 30\,K_{min}$ (*red curve*). The dark purple susceptibility curve corresponds to the case of an independent neuron ($K = 0$ repetition), and the dark purple horizontal line shows the upper value of this susceptibility (0.25)

Results are reported in Fig. 23. For the Task epoch of session 1, we have $T \simeq 1400$ sec, $f \simeq 5.8$ Hz; with these parameters the number of repetitions is $K_{min} \simeq 70$. This corresponds to a minimum frequency of repetition of the pattern equal to $f_{min} = K_{min}/T \simeq 0.05$ Hz, less than 1 % of the average firing frequency of neurons. As K exceeds K_{min} the maximum of the susceptibility curve increases beyond 0.25 and moves to the left, in agreement with results (for experimental data) shown in Section 4 (see in particular Fig. 18, left).

References

Abbeel, P., Koller, D., & Ng, A.Y. (2006). Learning factor graphs in polynomial time and sample complexity. *The Journal of Machine Learning Research*, *7*, 1743–1788.

Barton, J., & Cocco, S. (2013). Ising models for neural activity inferred via selective cluster expansion: structural and coding properties. *Journal of Statistical Mechanics: Theory and Experiment*, *2013*(03), P03002.

Barton, J.P., De Leonardis, E., Coucke, A., & Cocco, S. (2016). Ace: adaptive cluster expansion for maximum entropy graphical model inference. *Bioinformatics*. doi:10.1093/bioinformatics/btw328.

Battaglia, F.P., Benchenane, K., Sirota, A., Pennartz, C.M.A., & Wiener, S.I. (2011). The hippocampus: hub of brain network communication for memory. *Trends in Cognitive Sciences*, *15*(7), 310–318.

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P.L., Gioanni, Y., Battaglia, F.P., & Wiener, S.I. (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal-prefrontal network upon learning. *Neuron*, *66*(6), 921–936.

Billeh, Y.N., Schaub, M.T., Anastassiou, C.A., Barahona, M., & Koch, C. (2014). Revealing cell assemblies at multiple levels of granularity. *Journal of Neuroscience Methods*, *236*, 92–106.

Brown, E.N., Frank, L.M., Tang, D., Quirk, M.C., & Wilson, M.A. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *The Journal of Neuroscience*, *18*(18), 7411–7425.

Cardin, J.A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.H., & Moore, C.I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature*, *459*(7247), 663–667.

Carr, M.F., Jadhav, S.P., & Frank, L.M. (2011). Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, *14*(2), 147–153.

Chandler, D. (1987). *Introduction to Modern Statistical Mechanics*. Oxford University Press.

Chang, S.W.C., Gariépy, J., & Platt, M.L. (2013). Neuronal reference frames for social decisions in primate frontal cortex. *Nature Neuroscience*, *16*(2), 243–250.

Chapin, J.K., & Nicolelis, M.A.L. (1999). Principal component analysis of neuronal ensemble activity reveals multidimensional somatosensory representations. *Journal of Neuroscience Methods*, *94*(1), 121–140.

Cocco, S., & Monasson, R. (2011). Adaptive cluster expansion for inferring boltzmann machines with noisy data. *Physical Review Letters*, *106*(9), 090601.

Cocco, S., & Monasson, R. (2012). Adaptive cluster expansion for the inverse ising problem: convergence, algorithm and tests. *Journal of Statistical Physics*, *147*(2), 252–314.

Diba, K., & Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, *10*(10), 1241–1242.

Euston, D.R., Tatsuno, M., & McNaughton, B.L. (2007). Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science*, *318*(5853), 1147–1150.

Foster, D.J., & Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, *440*(7084), 680–683.

Ganguli, S., & Sompolinsky, H. (2012). Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. *Annual Review of Neuroscience*, *35*, 485–508.

Ganmor, E., Segev, R., & Schneidman, E. (2009). How fast can we learn maximum entropy models of neural populations *In Journal of Physics: Conference Series, volume 197, page*, *012020*.

Ganmor, E., Segev, R., & Schneidman, E. (2011a). The architecture of functional interaction networks in the retina. *The Journal of Neuroscience*, *31*(8), 3044–3054.

Ganmor, E., Segev, R., & Schneidman, E. (2011b). Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *Proceedings of the National Academy of Sciences*, *108*(23), 9679–9684.

Gerwinn, S., Macke, J., & Bethge, M. (2010). Bayesian inference for generalized linear models for spiking neurons. *Frontiers in Computational Neuroscience*, 4:12:1–17.

Harris, K.D., Csicsvari, J., Hirase, H., Dragoi, G., & Buzsáki, G. (2003). Organization of cell assemblies in the hippocampus. *Nature*, *424*(6948), 552–556.

Hebb, D.O. (1949). *The organization of behavior: A neurophysiological theory*: Wiley.

Hoffman, K.L., & McNaughton, B.L. (2002). Coordinated reactivation of distributed memory traces in primate neocortex. *Science*, *297*(5589), 2070–2073.

Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., & Yuste, R. (2004). Synfire chains and cortical songs: temporal modules of cortical activity. *Science*, *304*(5670), 559–564.

Ji, D., & Wilson, M.A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*(1), 100–107.

Johnson, A., & Redish, A.D. (2007). Neural ensembles in ca3 transiently encode paths forward of the animal at a decision point. *The Journal of Neuroscience*, *27*(45), 12176–12189.

Lee, A.K., & Wilson, M.A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron*, *36*(6), 1183–1194.

Litwin-Kumar, A., & Doiron, B. (2012). Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature Neuroscience*, *15*(11), 1498–1505.

Lopes-dos Santos, V., Conde-Ocazionez, S., Nicolelis, M.A.L., Ribeiro, S.T., & Tort, A.B.L. (2011). Neuronal assembly detection and cell membership specification by principal component analysis. *Plos One*, *6*(6), e20996.

Lopes-dos Santos, V., Ribeiro, S., & Tort, A.B.L. (2013). Detecting cell assemblies in large neuronal populations. *Journal of Neuroscience Methods*, *220*(2), 149–166.

Luczak, A., Barthó, P., Marguet, S.L., Buzsáki, G., & Harris, K.D. (2007). Sequential structure of neocortical spontaneous activity in vivo. *Proceedings of the National Academy of Sciences*, *104*(1), 347–352.

McCormick, D.A., Connors, B.W., Lighthall, J.W., & Prince, D.A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *Journal of Neurophysiology*, *54*(4), 782–806.

Peyrache, A., Benchenane, K., Khamassi, M., Wiener, S., & Battaglia, F. (2010a). Sequential reinstatement of neocortical activity during slow oscillations depends on cells' global activity. *Frontiers in Systems Neuroscience*, *3*, 18.

Peyrache, A., Benchenane, K., Khamassi, M., Wiener, S.I., & Battaglia, F.P. (2010b). Principal component analysis of ensemble recordings reveals cell assemblies at high temporal resolution. *Journal of Computational Neuroscience*, *29*(1-2), 309–325.

Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S.I., & Battaglia, F.P. (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nature Neuroscience*, *12*(7), 919–926.

Pfeiffer, B.E., & Foster, D.J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74–79.

Qin, Y.L., Mcnaughton, B.L., Skaggs, W.E., & Barnes, C.A. (1997). Memory reprocessing in corticocortical and hippocampocortical neuronal ensembles. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *352*(1360), 1525–1533.

Roumis, D., & Franck, L. (2015). Hippocampal sahrp-waves ripples in waking and sleeping states. *Current Opinion in Neurobiology*, *35*, 6–12.

Schneidman, E., Berry, M.J., Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, *440*(7087), 1007–1012.

Shlens, J. (2014). Notes on generalized linear models of neurons. arXiv:1404.1999.

Singer, A.C., Carr, M.F., Karlsson, M.P., & Frank, L.M. (2013). Hippocampal swr activity predicts correct decisions during the initial learning of an alternation task. *Neuron*, *77*(6), 1163–1173.

Singh, A., Peyrache, A., & Humphries, M. (2015). Task learning reveals signatures of sample-based internal models in rodent prefrontal cortex. bioRxiv. doi:10.1101/027102.

Tavoni, G., Ferrari, U., Battaglia, F.P., Cocco, S., & Monasson, R. (2015). Functional coupling networks inferred from prefrontal cortex activity show learning-related effective plasticity. bioRxiv. doi:10.1101/028316.

Tkačik, G., Marre, O., Amodei, D., Schneidman, E., Bialek, W., & Berry II, M.J. (2014). *Searching for collective behavior in a large network of sensory neurons*: Plos Computational Biology.

Truccolo, W., Eden, U.T., Fellows, M.R., Donoghue, J.P., & Brown, E.N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of Neurophysiology*, *93*(2), 1074–1089.

Wilson, M.A., & McNaughton, B.L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *265*(5172), 676–679.