Memory formation: from network structure to neural dynamics

BY SARAH FELDT¹, JANE X. WANG², VAUGHN L. HETRICK³,
JOSHUA D. BERKE³ AND MICHAŁ ŻOCHOWSKI^{1,4,*}

¹*Physics Department*, ²*Applied Physics Program*, ³*Psychology Department*, and
⁴*Biophysics Program, University of Michigan, Ann Arbor, MI 48109, USA*

Understanding the neural correlates of brain function is an extremely challenging task, since any cognitive process is distributed over a complex and evolving network of neurons that comprise the brain. In order to quantify observed changes in neuronal dynamics during hippocampal memory formation, we present metrics designed to detect directional interactions and the formation of functional neuronal ensembles. We apply these metrics to both experimental and model-derived data in an attempt to link anatomical network changes with observed changes in neuronal dynamics during hippocampal memory formation processes. We show that the developed model provides a consistent explanation of the anatomical network modifications that underlie the activity changes observed in the experimental data.

Keywords: neuronal networks; memory formation; functional clustering;
network structure; network dynamics

1. Introduction

The problem of understanding neuronal correlates of brain function has been addressed extensively over many decades but still remains unresolved. While a vast amount is known about the basic anatomy and physiology of the brain, the dynamics and interactions of neuronal ensembles that underlie various cognitive tasks are yet to be understood. This is largely because of the anatomical complexity of the neuronal networks that comprise the brain. The cortex alone contains 10^{10} neurons and 1.5×10^{14} synapses, making it impossible to derive any detailed properties of its connectivity. It is not even clear that having a detailed knowledge of the connectivity would be sufficient to understand brain function, as it significantly evolves on time scales ranging from tens of milliseconds to years, through processes such as constant rewiring (Song & Abbott 2001) (i.e. creation, annihilation and modulation of synapses), neuronal loss and/or adult neurogenesis (Abrous *et al.* 2005). Additionally, even if the anatomical connectivity were known, this would not necessarily lead to an understanding of the spatio-temporal patterning of neural activity, which is the basis of function (Fujisawa *et al.* 2008). It, therefore, becomes imperative to define other approaches that rely more on functional commonality (i.e. coding the same

*Author for correspondence (michalz@umich.edu).

One contribution of 10 to a Theme Issue ‘Experiments in complex and excitable dynamical systems’.

aspects of cognitive tasks) as opposed to anatomical connectivity. Moreover, since it is known that brain function is distributed over large neuronal ensembles, or, even more globally, between different brain modalities, it is important to understand how these ensembles self-organize to generate desired functions (movement, memory storage/recall, etc.) (Hebb 1949; Gerstein *et al.* 1978; Singer 1999; Zhou *et al.* 2007).

From the experimental perspective, the emergence of new multiunit electrophysiological and/or optical imaging techniques has been crucial as they provide information, albeit sparsely, on distributed neural activity during various cognitive tasks. Thus, the research task has been partially redefined, first to understand the functional (dynamical) network correlates that underlie the given cognitive phenomena, and then, based on these, to understand the anatomical structures and physiological processes that lead to them. Thus, in short, we are asking two questions: what macroscopically observed neural interactions are the hallmark of a given cognitive process, and what anatomical or physiological state underlies these interactions? Exploring these questions requires the formulation of new metrics that will allow the identification of emerging dynamical patterns during brain function. However, since it is quite difficult to experimentally link the observed dynamical changes to the underlying structural changes, extensive modelling efforts must also be done, where one can directly observe how known structural changes induce differences in functional relationships between neurons. While the knowledge gained from this modelling does not give direct evidence linking the experimentally observed changes in functional behaviour with underlying structural changes, it can provide confirmation that the experimental data are consistent with certain hypotheses.

In order to define these new metrics, we must turn back to cognitive sciences to identify which dynamical neuronal patterns are important. It is assumed that functional neural ensembles form and disintegrate dynamically (Milner 1974; von der Malsburg 1995; Engel & Singer 2001; Singer 2001), through spatio-temporal patterning of spiking activity comprising many individual neurons. The temporal correlation hypothesis (von der Malsburg 1981; Engel *et al.* 1991; Singer 1993; Gray 1999) postulates that correlated neuronal activity mediates rapid feature binding and thus the formation of intermittent functional ensembles in the brain. Therefore, the problem of identifying these functional neural ensembles can potentially be reduced to the identification of temporally correlated groups of neurons. However, it is also clear that the formation of these ensembles is mediated through rapid anatomical/physiological changes. It has been established that the temporally ordered co-activation of neural populations leads to rapid synaptic changes via spike-timing-dependent synaptic modification (spike-timing-dependent plasticity, STDP) processes (Bi & Poo 1998; Abbott & Nelson 2000; Song *et al.* 2000; Song & Abbott 2001). Since these synaptic modifications are directional, one would also expect changes in directional relationships between the firing patterns of neurons.

Here, we focus on the formulation of quantitative links between the anatomical and dynamical macroscopic network processes that underlie initial memory formation in the brain. We analyse hippocampal tetrode recordings obtained from freely moving mice that are exposed to a novel environment and look for two effects: (i) the enhancement of directional timing relationships between neuronal pairs; and (ii) a decrease in the overall temporal distance between firings of

subpopulations of cells during memory formation. Based on our previous work (Zochowski & Dzakpasu 2004; Waddell *et al.* 2007), we hypothesize that an increase in the number of pairs showing significant directional interdependences is indicative of the strengthening of direct connections between interacting neurons, whereas a decrease in the temporal distance indicates an overall increased fidelity in neuronal representation of the new environment (O'Neill *et al.* 2008). Finally, we link these two effects through computational modelling and show that the formation of heterogeneities in the anatomical connectivity of the network, because of the addition/strengthening of a relatively small number of neural connections, can lead to the rapid formation of co-activating discrete neuronal ensembles, as well as their re-activation during sleep. We hypothesize that this subsequently leads to the formation of a distributed memory representation as is observed experimentally.

2. Methods: metrics and analysis

(a) Directional interactions: causal entropies

We monitor the directional interactions between neurons over time using causal entropies (CEs) (Zochowski & Dzakpasu 2004; Dzakpasu & Zochowski 2005; Waddell & Zochowski 2007). CEs are an asymmetric, time-adaptive metric constructed to detect asymmetric locking between two spike trains based on the intervals between spiking events (Waddell *et al.* 2007). They are computed by first constructing two time-adaptive histograms of the inter-event intervals between the spike trains (P_{ij} and, separately, P_{ji}), and then calculating the Shannon entropy.

Briefly, let t_n^i be the time of the n th spike for neuron i and $\tau_j(n)$ be the last time that neuron j fired before t_n^i . We then calculate $\Delta t_{ij}(n) = t_n^i - \tau_j(n)$. Note that, if neuron i fires multiple times, $\Delta t_{ij}(n)$ is only calculated between the closest spike of neuron i and the previous spike of neuron j . Please see figure 1a for a schematic. We separately calculate Δt_{ji} (from spikes in j to spikes in i) in the same manner as shown in figure 1b.

At t_n^i , we update the histogram P_{ij} by adding Δp to the bin corresponding to $\Delta t_{ij}(n)$ and then renormalizing the histogram. The parameter Δp is a free parameter that controls the effective length of history of the time-adaptive measure (Waddell *et al.* 2007). This process establishes an exponential attenuation to the memory of the histogram, allowing it to adapt to changes in synchrony over time. The causal entropy $CE_{ij}(n) = -\sum_k P_{ij}(k) \log(P_{ij}(k))$ is then computed, where k indexes the bins of the histogram. We then separately calculate CE_{ji} through the same procedure as above, interchanging the i and j in the prior description.

The advantage of CEs is the ability to detect asymmetric locking between pairs of neurons. If neuron i regularly fires shortly after neuron j , but j does not regularly follow i , then CE_{ij} will become small, while CE_{ji} will remain relatively large. Therefore, one may take the causal entropy difference $CED_{ij}(n) = CE_{ij}(n) - CE_{ji}(n)$ to measure the degree and direction of locking between the two neurons. (See figure 1c for an example of histograms depicting this behaviour.) For a detailed description of CE, see Waddell *et al.* (2007).

Figure 1. (a) Sample calculation of Δt_{ij} from spikes in train S_i to train S_j . (b) Sample calculation of Δt_{ji} from spikes in train S_j to train S_i . (c) Sample histograms for spike trains in which spikes in S_j lead spikes in S_i but there is no consistent relationship from spikes in S_i to spikes in S_j .

We track the number of significant CE pairs over time by computing $\text{CED}_{ij}(n)$ and comparing this with the same measure computed for surrogate datasets. Please see §2c for a description of the surrogate data. Significance is determined as being two standard deviations away from the mean. Note that the mean and standard deviation of the surrogate data CED distribution change over time as the CE histograms are updated. For the experimental data analysed in the paper, we calculated histograms for Δt within a window of 500 ms, using a bin size of 10 ms and $\Delta p = 0.05$. The model data were analysed using a window of 400 time steps, a bin size of 1 time step and $\Delta p = 0.005$.

(b) *Correlated activity: functional clustering algorithm*

In order to detect functionally correlated groups of neurons, we implement the functional clustering algorithm (FCA) (Feldt et al. 2009). This algorithm groups neurons based on pairwise similarity of spiking dynamics. In this paper, to define similarity between neurons, we compute the average minimum distance (AMD), which detects co-firing events between pairs of neurons (Feldt et al. 2009). To compute the AMD between two spike trains S_i and S_j , we calculate the distance δt_n^i from each spike in S_i to the closest spike in S_j as shown in figure 2a. We then define

$$D_{ij} = \frac{1}{N_i} \sum_n \delta t_n^i \quad (2.1)$$

and

$$D_{ji} = \frac{1}{N_j} \sum_n \delta t_n^j, \quad (2.2)$$

Figure 2. (a) Example of $\delta t_n^{i/j}$ used to calculate the AMD between spike trains. In this case, the distance is calculated for each spike in one train to the closest spike in the other train. (b) Example of the FCA applied to four spike trains. Spike trains with the highest scaled significance are merged (here S_1 and S_2 are joined to create S_5) and the procedure is repeated until no significant joining steps remain. The resulting clusters are indicated by the circled groups in the obtained dendrogram.

where $N_{i/j}$ is the total number of spikes in S_i or S_j , respectively. Finally, we define the AMD between spike trains S_i and S_j to be

$$\text{AMD}_{ij} = \frac{D_{ij} + D_{ji}}{2}. \quad (2.3)$$

The FCA then proceeds as follows. A schematic of the algorithm can be seen in [figure 2b](#).

1. We first create a matrix of AMD values between all spike trains.
2. We then use surrogate datasets (see §2c) to calculate 95% confidence intervals for each pairwise AMD. These significance levels are used to calculate the scaled significance between each pair of AMD values. The scaled significance is measured in units defined as the distance from the midpoint of the cumulative distribution function derived from surrogate data to the 95 per cent significance cut-off. Thus, a scaled significance value equal to 1 denotes the 95 per cent significance level, and values higher than 1 are significant, while values lower than 1 are deemed insignificant.
3. The pair of trains with the highest significance is then chosen to be grouped together, and the scaled significance of this pair is recorded. A unique element of the FCA is that the two spike trains that are grouped together are then merged by joining the spikes into a single new train ([figure 2b](#)).
4. The trains that are being joined are then removed, the AMD matrix is recalculated for the new set of trains, new surrogate datasets are created, and a new scaled significance matrix is calculated.
5. We repeat the joining steps (3 and 4), recording the scaled significance value used in each step of the algorithm until the point at which no pairwise similarity is statistically significant. The significance cut-off is indicated by the dashed line in [figure 2b](#). At this point, the functional groupings are determined by observing which spike trains have been combined during the clustering algorithm (e.g. circled groups in the dendrogram of [figure 2b](#)).

(c) Creation of surrogate data

In order to determine the significance of the above measures, we create surrogate datasets by adding a jitter to each spike in the train. This jitter is drawn from a uniform distribution ([Date *et al.* 1998](#)) within a given window around each spike. This destroys correlations between firing events of neurons while preserving the number of spikes and average properties of the interspike interval distribution. For the experimental data, a jitter window of 2 s was used for the CE analysis, and a window of 20 s was used for the FCA analysis. For the model data, a window of 400 time steps was used for both calculations.

(d) Experimental protocol

The experimental data presented in this paper were recorded using tetrode methods from the hippocampus of freely moving mice as they explored a novel environment. The datasets analysed here are from those previously published in [Berke *et al.* \(2008\)](#), in which a complete description of the experimental methods can be found. All animal experiments were approved by the University of Michigan Committee on the Use and Care of Animals. Six tetrodes were implanted into the dorsal hippocampus of each animal. The mice were placed in a novel rectangular track in which they underwent periods of exploration and sleep, which were confirmed through the inspection of the EEG signal and locomotor activity. Here, we analyse data from 77 pyramidal neurons in mouse 1 (42 CA1; 21 CA2; 14 CA3) and 28 pyramidal neurons in mouse 2 (14 CA1, 14 CA3). The

early exploration window was taken to be 100 s of exploration when the animal was initially placed in the environment, whereas late exploration was defined to be 100 s of data taken from the end of the session after the animal had slept in the environment.

3. Modelling: understanding the structural basis of the observed temporal patterning

It has been shown that the hippocampus can rapidly form new memory representation and, in a very short time period, is able to generate experience-dependent reactivation during various stages of sleep (Wilson & McNaughton 1994; Buzsaki 1998; Kudrimoti *et al.* 1999; Booth & Poe 2006) and quiet waking periods (Foster & Wilson 2006). During this reactivation, the spatio-temporal patterning of neuronal activity is correlated with the patterning of the preceding awake activity (Louie & Wilson 2001). Furthermore, the correlation between cells co-active during waking is higher during sleep (Pavlidis & Winson 1989; Wilson & McNaughton 1994; Poe *et al.* 2002). We used a simple integrate-and-fire based model to show that local strengthening of network connectivity underlies the formation of localized network heterogeneity, which, in turn, leads to rapidly increasing neuronal co-activation and potential formation of distributed memory representation. Furthermore, changes of global network excitability, driven, for example, through changes in the neurochemical environment during sleep (Marrosu *et al.* 1995; Booth & Poe 2006), can produce spontaneous reactivation of the previously co-active network regions. Thus, the structure of the activating regions in the network are inherently determined by the heterogeneities in network topology. Here, we show that this simple model coupled with an STDP-based learning rule phenomenologically captures the processes observed experimentally.

The model uses leaky integrate-and-fire neurons given by

$$\tau_m \frac{dV_{e/i}^j}{dt} = -\alpha_j V_{e/i}^j + I_{e/i} + \sum_k w^{jk} I_{\text{syn}}^k \quad (3.1)$$

to represent the dynamics of the network elements, with e/i denoting either an excitatory or inhibitory neuron. $V_{e/i}^j$ is the membrane voltage of the j th neuron; α_j is the membrane leak rate constant randomly distributed, such that $\alpha_j \in [1, 1.3]$; $\tau_m = 30$ ms is the membrane time constant; I_{syn}^k is the incoming current to the j th neuron from the k th neuron; and w^{jk} is the connection strength between neurons j and k and is discussed below. The excitatory subnetwork of 500 cells is in a one-dimensional small-world formulation with periodic boundaries and has connectivity radius $R_e = 3$. Additionally, $p_g^e = 0.15$ is the rewiring parameter defining the fraction of the number of local connections to the number of random, long-range ones. The inhibitory interneuron subnetwork of 100 cells has $R_i = 1$, $p_g^i = 1$ and $w_{\text{in}} = 10$, forming a random graph network. Every inhibitory cell receives input from $n_{\text{ei}} = 5$ neighbouring excitatory neurons with strength $w_{\text{ei}} = 4$, and every excitatory neuron receives input from $n_{\text{ie}} = 10$ random inhibitory ones with strength $w_{\text{ie}} = 2$. Synaptic strengths were chosen to balance the number of

incoming connections so that the total possible input to all cells remains the same. The external current $I_{e/i}$ is uniform over the entire excitatory/inhibitory network and functions as a global control parameter that controls response transitions, from low-frequency random activity (low I_e), to spontaneous activation of localized network regions (for intermediate I_e values) and finally to global bursting (high I_e) (Jablonski *et al.* 2007). We have shown that the intermediate values of global excitation (I_e) provide a mechanism for selective activation of network regions having slightly higher connectivity (i.e. potentially the ones experiencing synaptic formation and strengthening), while rest of the network remains quiescent. Additionally, the transition from low to intermediate levels of excitation in our model may represent the transition in the hippocampus between the active awake state and the sleep state that exhibits memory reactivation (Wang *et al.* 2008). For the simulations presented in this paper, $I_e = 0.65$ during stimulation phase and $I_e = 1.05$ during reactivation (sleep phase); $I_i = 1.2$.

When the membrane potential of a given cell reaches a maximum value of $V_{\text{reset}} = 1$, the neuron emits an action potential, its membrane potential is reset to $V_{\text{rest}} = 0$ and the neuron enters a refractory period for $\tau_{\text{refr}} = 10$ ms. The synaptic current emitted by spiking neuron (k) is of the form

$$I_{\text{syn}}^k(t) = \exp\left(\frac{-(t - t_{\text{spike}}^k)}{\tau_s}\right) - \exp\left(\frac{-(t - t_{\text{spike}}^k)}{\tau_f}\right), \quad (3.2)$$

where $(t - t_{\text{spike}}^k)$ is the time since neuron k last spiked, $\tau_s = 1.5$ ms is the slow time constant, and $\tau_f = 0.15$ ms is the fast time constant. In addition to the synaptic current from other cells, all neurons have a $p_{\text{fire}} = 10^{-3}$ probability of firing spontaneously per millisecond.

The excitatory subnetwork undergoes synaptic modification based on spiking activity of these cells. This subnetwork is composed of 25 per cent non-modifiable, homogeneous, active connections with strength $w_{\text{ex}} = 2$, in addition to 75 per cent modifiable synapses, which are connections initially with weight 0 but can modulate their strength between $w_{\text{ex}} = 0$ and $w_{\text{ex}} = 2$ based on neuronal activity (Isaac *et al.* 1995). The changes in synaptic strength of modifiable synapses are implemented based on a simplified neurobiological rule of STDP (Bell *et al.* 1997; Markram *et al.* 1997; Bi & Poo 1998; Lengyel *et al.* 2005). Upon firing of neuron k , synaptic strength from neuron j to neuron k is incrementally increased if neuron j (presynaptic cell) fires before neuron k (postsynaptic cell) within a set interspike interval (ISI) of $T_L = 20$ ms. At the same time, synaptic strength from neuron k to neuron j is decreased by the same amount. Additionally, synaptic efficacy from the presynaptic to the postsynaptic cell is decreased by a smaller amount when the two cells do not activate congruently, i.e. their ISI is above the set threshold $T_F = 200$ ms. Formally,

$$\Delta w_{jk}^* = \begin{cases} \frac{w_{\text{ex}}}{\tau_{\text{learn}}} & \text{if } t_j - t_k < T_L, \\ -\frac{w_{\text{ex}}}{\tau_{\text{forget}}} & \text{if } t_j - t_k > T_F, \\ 0 & \text{if } T_L < t_j - t_k < T_F \end{cases} \quad (3.3)$$

and

$$\Delta w_{kj}^* = \begin{cases} -\frac{w_{\text{ex}}}{\tau_{\text{learn}}} & \text{if } t_j - t_k < T_L, \\ 0 & \text{otherwise.} \end{cases} \quad (3.4)$$

The w_{jk}^* indicates the weight of the modifiable synapse from neuron j to neuron k , $w_{\text{ex}} = 2$ is the strength of non-modifiable synapses in the excitatory network, $t_j - t_k$ is the ISI between neurons j and k , and $\tau_{\text{learn}} = 20$ and $\tau_{\text{forget}} = 70$ are the rates of learning and forgetting in the network. For comparison purposes, learning is not turned on until 3 s into the simulation. Calculations are performed using Euler's method with time steps of 0.05 ms.

To simulate external sensory input, a constant current of value 0.9 is fed into a localized subgroup of 100 neurons in the excitatory neuronal network (neuron IDs 201–300) at times 2–6 s and 9–12 s. A period of sleep is implemented between 6.5–8.5 s by raising I_e to 1.05 to represent biological neuromodulatory mechanisms during sleep.

To be able to compare the results of our simulation to experimental data, we analyse only the activity patterns of the network of excitatory neurons. The early plasticity window analysed here was defined to be 3–4 s, just after plasticity had been turned on in the simulation. This was done in order to mimic the early exploration window for the mice when they were initially placed in the environment. Late plasticity was defined as 10–11 s (after the sleep period), again for comparison with the late exploration window in the experimental data.

4. Results

We analysed both experimental and model-derived data to explore changes in network structure and dynamics as a result of synaptic modifications during exposure to a stimulus. In the experimental data, we analyse spike train data obtained from tetrode recordings of freely moving mice (see §2*d*) as they learned a novel track environment. An example raster plot of the obtained recordings along with the mouse's behavioural state can be seen in figure 3*a*.

To better understand the structural network changes that underlie the observed changes in dynamics, we also analyse data from the excitatory network of a model of hippocampal memory formation (see §3). An example raster plot of the model data can be seen in figure 4*a*. In this model, memories are formed by the stimulation of a selected group of neurons with modifiable synapses that can be strengthened/weakened through a learning rule implemented to simulate STDP processes. The modifiable synapses were initially silent (Isaac *et al.* 1995) and became selectively active, driven by an activity-dependent synaptic modification process. After stimulation starts at 2 s, the external input given to neurons 201–300 induced these neurons to fire with spatio-temporal patterning that induced rapid strengthening of synapses when synaptic plasticity is turned on at 3 s. The synaptic plasticity continues until the stimulation is ceased at 6 s.

Below we compare and quantify the changes in directional neural interactions and evolution of functional clustering for the experimental data and model data.

Figure 3. (a) Example raster plot of recordings from mouse 2 as it explores and sleeps in a novel environment. (b) Number of significant CED pairs as a function of the mouse's time in the environment. The rise in the number of significant pairs indicates the formation of directional lead-lag relationships between neurons as the mouse sleeps and continues to explore the environment.

Figure 4. (a) Example raster plot of model data. Neurons 201–300 receive external input to simulate the presence of a stimulus as indicated by the light grey bar. An STDP rule is applied for the period marked by the dark grey bar, and a global external input is applied to simulate sleep as indicated by the black bar, causing re-activation in the previously stimulated neurons. (b) Number of significant CED pairs and average value of simulated modifiable synapses. Once plasticity is turned on, we see a rise in the number of significant directional relationships between neurons, which continues to increase during the sleep period and persists during further stimulation.

(a) Directional interactions

In order to quantify the changes in directional interactions between neurons in both experimental and model data, we analysed pairwise interactions between neurons using causal entropies (see §2a). This measure detects directional lead-lag patterning between spiking of neurons as a function of time. In figures 3b and 4b, we show the number of significant CED pairs as calculated for the experimental and model data, respectively.

In the experimental data, we see that the number of significant CED pairs begins to grow during the first period of sleep and continues to increase as the mouse further explores and learns the new environment. This corresponds to an increase in the number of significant directional (lead–lag) relationships between neurons, which is consistent with the development of enhanced connectivity between cells during memory consolidation (Waddell *et al.* 2007).

We also see an increase in the number of directional relationships between neurons in the model data once plasticity has been turned on. Here, we calculate patterning only between the pairs of stimulated neurons. When the external stimulus current is first applied at 2 s, the stimulated neurons begin to fire more rapidly and we see the development of a few significant CED pairs (figure 4*b*, black line). Once plasticity is added 1 s later into the simulation, the standard STDP rule strengthens the modifiable directional connections between presynaptic and postsynaptic neurons, leading to more reliable lead–lag patterning between the firing of the neurons. The grey line in figure 4*b* represents the evolution of the mean weight of modifiable synapses within the stimulated region. This leads to a further increase in the number of significant CED pairs due to these known structural changes in the network topology. Thus, the rise in average synaptic strength is correlated with the rise in significant CED pairs. It should be noted that the delay between the initial occurrence of significant CED pairs and rise in synaptic strength is because of the fact that the stimulation period begins before plasticity is turned on, and some directional relationships arise from the initial network connectivity. The levelling out of the synaptic weight curve observed at the end of the first stimulation period and during the second stimulation is due to the fact that the weight of the modifiable synapses is bounded at $w_{\text{ex}} = 2$ and thus the synapses are limited in their ability to continuously increase in strength. The increase in significant CED pairs during that time can therefore be attributed to reorganization of the structural connectivity patterns rather than their mere strengthening.

During the simulated sleep environment, the structural heterogeneity formed during the initial stimulus presentation is more excitable when compared with other network regions and thus mediates its reactivation. This reactivation exemplifies itself as an occurrence of synchronized bursts that in turn lead to further modifications of network structure. These directional relationships persist through the second period of stimulation after sleep.

Although we have no way of determining network connectivity in the experimental data, we can study the changes in network topology in our model as the connection strengths evolve throughout the simulation. To further understand how the emergence of directional firing relationships corresponds to structural changes in the underlying network connectivity, we compared the correlation between the significance of lead–lag patterning (i.e. the number of significant CED pairs) and the absolute value of the difference in directional synaptic weights (i.e. $w_{jk} - w_{kj}$). In figure 5, we show a histogram of this relationship towards the end of the second period of external stimulation (time = 11.5 s for the data from figure 4). At this point in time, the modifiable synapses have evolved to form directional relationships between neurons. A large structural directional relationship will thus be given by synaptic pairwise weight differences close to 2, while synaptic pairwise weight differences near 0 indicate the lack of a structural directional connectivity (i.e. the two neurons are not connected or have bidirectional connections with

Figure 5. Histogram of the number of significant CED pairs as a function of the absolute value of the difference in directional weight between pairs of neurons for time = 11.5 s in the simulation presented in figure 4. There is a correlation between a large synaptic pairwise weight difference and the occurrence of a significant CED relationship. At the same time, there is also an increase of significant CED pairs for the lowest values of the synaptic pairwise weight difference, indicative of the formation of polysynaptic patterns.

Table 1. Percentage of significant CED pairs and change in FCA significance.

	CED: early exploration/ plasticity (%)	CED: late exploration/ plasticity (%)	FCA: significance change
experiment	1.8 ± 0.6	7.5 ± 1.5	0.54 ± 0.46
model	8.3 ± 1.5	56.4 ± 9.1	0.57 ± 0.03

similar weight). Note that pairs of neurons with a synaptic pairwise weight difference of 0 have been excluded from the plot because of the large number of non-connected neurons. We observe that a large synaptic weight difference corresponds to a high number of significant CED pairs, indicating that indeed the direct directional connection between neuronal pairs leads to the formation of a lead-lag pattern between the two cells. The smaller peak in the histogram for low synaptic pairwise weight differences is probably due to complex indirect interactions between groups of neurons (i.e. the lead-lag pattern forms through polysynaptic interactions rather than monosynaptic ones).

In table 1, we further quantify the increase in directional relationships by calculating the percentage of significant CED pairs during early exploration/plasticity and during late exploration/plasticity (after sleep). Data were averaged over $n=2$ experimental trials and $n=4$ simulated trials. The number of significant CED pairs was determined by the point at the end of the window designated as the early/late exploration/plasticity period described in §§2d and 3. In both the experimental and model data, we see substantial increases in the percentage of significant CED pairs during the late exploration/plasticity stages, quantifying the increase of directional relationships between neurons as a result of memory consolidation and learning.

Figure 6. (a) The scaled significance used to join trains plotted as a function of joining steps in the FCA for the experimental data presented in figure 3 during the early (black line) and late (grey line) exploration (after sleep) periods. (b) The scaled significance used in the FCA for the model data presented in figure 4 during early (black line) and late (grey line) plasticity (after sleep). The dashed grey line denotes the clustering cut-off. Values above this line are significant steps while values below this line are insignificant. The shaded regions denote the area between the significance curve and the clustering cut-off used to quantify changes in the amount of significance as described in the text.

(b) Functional groupings

In order to study the changes in functional groupings of neurons before and after learning, we implement the FCA with a similarity metric designed to detect co-firing events in neuronal activity. This algorithm not only parses the data into functional groupings, but also assigns values of significance to each joining step that combines groups of neurons (see §2b). It is therefore relevant to compare how significant the joining steps in the algorithm are between early and late exploration/plasticity. To do so, we compare the scaled significance used in the clustering during the initial stimulus presentation and a presentation that occurs later in the same trial, after a number of quiet waking and sleep periods. As functional groupings become increasingly coordinated in their activity, we expect to observe an increase in the significance of the joining steps, indicating the decrease in temporal distance between firing times. At the same time, neurons that become increasingly decorrelated in their activity will be indicated by a decrease in the amount of significance. This will lead to steeper decline of the scaled joining significance curve as a function of clustered steps.

In figure 6, we plot the scaled significance used in each joining step of the FCA when applied to examples from the experimental and model data. Examination of the scaled significance for the experimental data shows that we indeed see a change in the slopes of the curve between early and late exploration. The slope of the scaled significance is steeper during late exploration. This is because of the fact that the joining significance is higher during late exploration for the initial steps of the algorithm, which are indicative of the decreased temporal distance between the most correlated neurons. However, later in the algorithm (during the joining steps that are deemed insignificant, representing neurons that are less correlated), we see a decrease (from early to late exploration) in the amount of significance. This indicates the expected loosening of functional interactions between neurons not involved in the coding of the environment.

We also observe an increase in the number of significant steps between early and late exploration. During early exploration, we detect no significant functional clusters, while during late exploration, the first two steps of the algorithm are significant (above the dashed line) indicating the formation of functional clusters. Although we observe very few significant groupings in the experimental data, this is expected because of the low probability of detecting synchronous neurons due to the sparseness of sampling of the network and the distributed nature of neural coding. Even though the probability of detecting co-firing neurons is increased due to the fact that the recorded cells are involved in the same local network as captured by a single tetrode, it remains unclear how sparsely and distributed a given memory is coded. Clearly, we do not have this limitation in our model data.

We quantify the effect of the increase in functional correlations by measuring the area between the significance curve and the clustering cut-off line for the significant joining steps of the algorithm. This area is indicated by the shaded regions of figure 6b. In order to normalize between data with different numbers of neurons, we quantify the change by calculating $(A_{\text{late}} - A_{\text{early}})/(A_{\text{late}} + A_{\text{early}})$. This measure is contained within $[-1, 1]$ and positive values indicate an increase in the scaled significance used to cluster the neurons. We show the results of this calculation for both experimental and model data in table 1. In both cases, we are able to quantify the increase in significance used in the clustering steps of the algorithm.

5. Discussion

The observed changes in the neural patterning during the experiments support the underlying hypothesis that, as the mice explored the track, they learned a new spatial representation of the novel environment (as indicated by the formation of ‘place fields’ (Berke *et al.* 2008)). While in the track, the mice underwent periods of sleep, followed by further exploration of the environment. These epochs of slow-wave sleep have been hypothesized to be a period of memory consolidation (Buzsaki 1998; Kudrimoti *et al.* 1999), which is presumed to involve further alterations in structural as well as functional network connectivity. These structural alterations involve both the strengthening of existing monosynaptic connections between the neurons as well as the development of new connections. As a result of these changes, we observed an increase in directional lead-lag patterning between cells as the mice learned the new environment. Since sleep is thought to play a focal role during memory consolidation, this directional patterning could appear during sleep phases and continue beyond them, as was observed in the data.

Additionally, recent experimental findings have shown that two changes in neural firing patterns correlate with the memory consolidation of neural representations of novel stimuli: neurons that are correlated during initial exposure progressively increase their co-firing, while the neurons that are initially less correlated in their activity become further decorrelated (O’Neill *et al.* 2008). As predicted by this result, we observe a tightening of functional interactions between initially correlated neurons as the neurons that code for the environment increase their co-firing activity.

In the model data, we show how the addition of an STDP-motivated learning rule leads to synaptic modifications, which, in turn, give rise to changes in neuronal dynamics. Here, a subset of the population is stimulated and the plastic synapses are allowed to evolve under the learning rule. The strengthening/weakening of synapses within this discrete region effectively creates a structural inhomogeneity of the network connectivity. Sleep is modelled as a global increase of network excitability, and, during this period, the previously stimulated region is able to reactivate while suppressing the surrounding areas due to the topological nature of the network, which allows for focal excitation as well as global, random inhibition. The observed reactivation also allows synapses to strengthen further. When this external stimulation is repeated, we observe even further strengthening of synapses. These known changes in the network structure give rise to changes in neuronal dynamics, which match the observed changes in the experimental data.

6. Conclusions

As the amount of experimental data depicting neural interactions during various cognitive tasks increases, it is becoming essential to develop metrics that quantify neural relationships during different behavioural states. Equally important is the ability to link these changes in neural relationships to the underlying structural and/or neurobiological changes. This task of linking observed changes in dynamical behaviour to structural changes implies a need to combine experimental data with extensive modelling, where structural changes can be directly linked to dynamical changes. In this paper, we have presented new methods that, when applied to both experimental and model data, depict changes in neural dynamics and allow for explanation of the underlying structural changes that give rise to the observed dynamical changes.

Specifically, we tried to link the progressive functional clustering observed during exposure to a novel environment with the underlying structural network changes. To do so, we used two measures developed in our laboratory (CED and FCA) in order (i) to quantify the emergence of directional dynamical interdependences in the network, which are indicative of the enhancement of network connectivity; and (ii) to monitor the emergence of functional clusters based on activity patterning of neurons. We applied these two measures to experimental data, where they detected increases in directional relationships between neurons and a tightening of functional clusters as the mice explored, slept and learned a novel environment. We then implemented the same measures in a simple model, which implemented STDP processes during exposure to an external stimulus and periods of sleep. We observe very similar changes in the number of significant directional pairs as well as the emergence of functional clusters during both stimulus presentation and sleep. These results are consistent with the hypothesis that the observed dynamical changes are a result of underlying structural changes induced through STDP processes as a function of learning.

This work was supported by the NSF Graduate Research Fellowship (S.F.), NIBIB grant EB008163 (M.Ž.), the Molecular Biophysics Training Program at the University of Michigan (J.X.W.), the Whitehall Foundation (J.D.B.), and the National Institute of Drug Abuse R01 DA14318 (J.D.B.).

References

- Abbott, L. F. & Nelson, S. B. 2000 Synaptic plasticity: taming the beast. *Nat. Neurosci.* **3**, 1178–1183. (doi:10.1038/81453)
- Abrous, D. N., Koehl, M. & Le Moal, M. 2005 Adult neurogenesis: from precursors to network and physiology. *Physiol. Rev.* **85**, 523–569. (doi:10.1152/physrev.00055.2003)
- Bell, C. C., Han, V. Z., Sugawara, Y. & Grant, K. 1997 Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature* **378**, 278–281. (doi:10.1038/387278a0)
- Berke, J. D., Hetrick, V., Breck, J. & Greene, R. W. 2008 Transient 23–30 Hz oscillations in mouse hippocampus during exploration of novel environments. *Hippocampus* **18**, 519–529. (doi:10.1002/hipo.20435)
- Bi, G. Q. & Poo, M. M. 1998 Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18**, 10 464–10 472.
- Booth, V. & Poe, G. R. 2006 Input source and strength influences overall firing phase of model hippocampal CA1 pyramidal cells during theta: relevance to REM sleep reactivation and memory consolidation. *Hippocampus* **16**, 161–173. (doi:10.1002/hipo.20143)
- Buzsaki, G. 1998 Memory consolidation during sleep: a neurophysiological perspective. *J. Sleep Res.* **7**(Suppl. 1), 17–23. (doi:10.1046/j.1365-2869.7.s1.3.x)
- Date, A., Bienenstock, E. & Geman, S. 1998 On the temporal resolution of neural activity. Technical Report, Division of Applied Mathematics, Brown University.
- Dzakupasu, R. & Zochowski, M. 2005 Discriminating differing types of synchrony in neural systems. *Physica D* **208**, 115–122. (doi:10.1016/j.physd.2005.06.005)
- Engel, A. K. & Singer, W. 2001 Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* **5**, 16–25. (doi:10.1016/S1364-6613(00)01568-0)
- Engel, A. K., Konig, P. & Singer, W. 1991 Direct physiological evidence for scene segmentation by temporal coding. *Proc. Natl Acad. Sci.* **88**, 9136–9140. (doi:10.1073/pnas.88.20.9136)
- Feldt, S., Waddell, J., Hetrick, V. L., Berke, J. D. & Zochowski, M. 2009 Functional clustering algorithm for the analysis of dynamic network data. *Phys. Rev. E* **79**, 056104. (doi:10.1103/PhysRevE.79.056104)
- Foster, D. J. & Wilson, M. A. 2006 Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* **440**, 680–683. (doi:10.1038/nature04587)
- Fujisawa, S., Amarasingham, A., Harrison, M. T. & Buzsaki, G. 2008 Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat. Neurosci.* **11**, 823–833. (doi:10.1038/nn.2134)
- Gerstein, G. L., Perkel, D. H. & Subramanian, K. N. 1978 Identification of functionally related neural assemblies. *Brain Res.* **140**, 43–62. (doi:10.1016/0006-8993(78)90237-8)
- Gray, C. M. 1999 The temporal correlation hypothesis of visual feature integration: still alive and well. *Neuron* **24**, 31–47. (doi:10.1016/S0896-6273(00)80820-X)
- Hebb, D. O. 1949 *The organization of behavior*. New York, NY: Wiley.
- Isaac, J. T., Nicoll, R. A. & Malenka, R. C. 1995 Evidence for silent synapses: implications for the expression of LTP. *Neuron* **15**, 427–434. (doi:10.1016/0896-6273(95)90046-2)
- Jablonski, P., Poe, G. R. & Zochowski, M. 2007 Structural network heterogeneities and network dynamics: a possible dynamical mechanism for hippocampal memory reactivation. *Phys. Rev. E* **75**, 011912. (doi:10.1103/PhysRevE.75.011912)
- Kudrimoti, H. S., Barnes, C. A. & McNaughton, B. L. 1999 Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. *J. Neurosci.* **19**, 4090–4101.
- Lengyel, M., Kwag, J., Paulsen, O. & Dayan, P. 2005 Matching storage and recall: hippocampal spike timing-dependent plasticity and phase response curves. *Nat. Neurosci.* **8**, 1677–1683. (doi:10.1038/nm1561)
- Louie, K. & Wilson, M. A. 2001 Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron* **29**, 145–156. (doi:10.1016/S0896-6273(01)00186-6)

- Markram, H., Lubke, J., Frotscher, M. & Sakmann, B. 1997 Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275**, 213–215. (doi:10.1126/science.275.5297.213)
- Marrosu, F., Portas, C., Mascia, M. S., Casu, M. A., Fa, M., Giagheddu, M., Imperato, A. & Gessa, G. L. 1995 Microdialysis measurement of cortical and hippocampal acetylcholine release during sleep–wake cycle in freely moving cats. *Brain Res.* **671**, 329–332.
- Milner, P. M. 1974 A model for visual shape recognition. *Psychol. Rev.* **81**, 521–535. (doi:10.1037/h0037149)
- O'Neill, J., Senior, T. J., Allen, K., Huxter, J. R. & Csicsvari, J. 2008 Reactivation of experience-dependent cell assembly patterns in the hippocampus. *Nat. Neurosci.* **11**, 209–215. (doi:10.1038/nn2037)
- Pavlides, C. & Winson, J. 1989 Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *J. Neurosci.* **9**, 2907–2918.
- Poe, G. R., Thompson, C. M., Riley, B. T., Tysor, M. K., Bjorness, T. E., Steinhoff, B. P. & Ferluga, E. D. 2002 A spatial memory task appropriate for electrophysiological recordings. *J. Neurosci. Methods* **121**, 65–74. (doi:10.1016/S0165-0270(02)00233-9)
- Singer, W. 1993 Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* **55**, 349–374. (doi:10.1146/annurev.ph.55.030193.002025)
- Singer, W. 1999 Neuronal synchrony: a versatile code for the definition of relations? *Neuron* **24**, 49–65, 111–125. (doi:10.1016/S0896-6273(00)80821-1)
- Singer, W. 2001 Consciousness and the binding problem. *Ann. NY Acad. Sci.* **929**, 123–146.
- Song, S. & Abbott, L. F. 2001 Cortical development and remapping through spike timing-dependent plasticity. *Neuron* **32**, 339–350. (doi:10.1016/S0896-6273(01)00451-2)
- Song, S., Miller, K. D. & Abbott, L. F. 2000 Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat. Neurosci.* **3**, 919–926. (doi:10.1038/78829)
- von der Malsburg, C. 1981 The correlation of brain function. Internal Report, MPI Biophysical Chemistry.
- von der Malsburg, C. 1995 Binding in models of perception and brain function *Curr. Opin. Neurobiol.* **5**, 520–526. (doi:10.1016/0959-4388(95)80014-X)
- Waddell, J. & Zochowski, M. 2007 Intraburst versus interburst locking in networks of driven non-identical oscillators. *Phys. Rev. E* **76**, 056216. (doi:10.1103/PhysRevE.76.056216)
- Waddell, J., Dzakpasu, R., Booth, V., Riley, B., Reasor, J., Poe, G. & Zochowski, M. 2007 Causal entropies—a measure for determining changes in the temporal organization of neural systems. *J. Neurosci. Methods* **162**, 320–332. (doi:10.1016/j.jneumeth.2006.12.008)
- Wang, J., Poe, G. R. & Zochowski, M. 2008 From network heterogeneities to familiarity detection and hippocampal memory management. *Phys. Rev. E* **78**, 041905. (doi:10.1103/PhysRevE.78.041905)
- Wilson, M. A. & McNaughton, B. L. 1994 Reactivation of hippocampal ensemble memories during sleep. *Science* **265**, 676–679. (doi:10.1126/science.8036517)
- Zhou, C. S., Zemanova, L., Zamora-Lopez, G., Hilgetag, C. C. & Kurths, J. 2007 Structure–function relationship in complex brain networks expressed by hierarchical synchronization. *New J. Phys.* **9**, 178–198. (doi:10.1088/1367-2630/9/6/178)
- Zochowski, M. & Dzakpasu, R. 2004 Conditional entropies, phase synchronization and changes in the directionality of information flow in neural systems. *J. Phys. A* **37**, 3823–3834. (doi:10.1088/0305-4470/37/12/007)