From network heterogeneities to familiarity detection and hippocampal memory management

Jane X. Wang

Applied Physics Program, University of Michigan, Ann Arbor, Michigan 48109, USA

Gina Poe

Department of Anesthesiology and Department of Molecular and Integrative Physiology,
University of Michigan Medical School, Ann Arbor, Michigan 48109-5615, USA

Michal Zochowski*

Department of Physics, Biophysics Program, Michigan Center for Theoretical Physics,
University of Michigan, Ann Arbor, Michigan 48109, USA

Hippocampal-neocortical interactions are key to the rapid formation of novel associative memories in the hippocampus and consolidation to long term storage sites in the neocortex. We investigated the role of network correlates during information processing in hippocampal-cortical networks. We found that changes in the intrinsic network dynamics due to the formation of structural network heterogeneities alone act as a dynamical and regulatory mechanism for stimulus novelty and familiarity detection, thereby controlling memory management in the context of memory consolidation. This network dynamic, coupled with an anatomically established feedback between the hippocampus and the neocortex, recovered heretofore unexplained properties of neural activity patterns during memory management tasks which we observed during sleep in multiunit recordings from behaving animals. Our simple dynamical mechanism shows an experimentally matched progressive shift of memory activation from the hippocampus to the neocortex and thus provides the means to achieve an autonomous off-line progression of memory consolidation.

PACS number(s): 87.18.Sn, 87.19.lj, 87.19.lv, 87.19.lp

I. INTRODUCTION

The memory formation process is founded upon synaptic reorganization and modification regulated by neural activity. When associative memories are first formed, cortical sensory areas which project to the hippocampal associative network activate the hippocampus and rapidly (within seconds) form a new network of synaptic weights encoding that memory. Over the span of days and weeks, rapidly formed novel memory networks in the hippocampus are consolidated to the cortex in a time- and activity-dependent fashion [1–3], eventually allowing memories to be independent of the hippocampus altogether [4]. Recent studies [5] have shown that storage and recall of spatial memory can occur independently of the hippocampus once schemas have been formed. Moreover, studies investigating brain metabolism and activity-related genes in mice suggest the decreasing importance of the hippocampus as time passes after learning and the increasing importance of several cortical regions [6]. These and other findings [7] suggest that the hippocampus is a general-purpose learner of new facts and events, both spatial and nonspatial [8], but that the cortex handles long-term storage of memory. Electrophysiological [2,9–11] and genetic [12] studies have combined with behavioral and neurological case studies [13,14] to build a coherent cellular and behavioral theory of how the consolidation process occurs offline (e.g., during sleep) through the reactivation of patterns of neuronal activity observed during awake learning [3,15–17].

From a dynamical perspective it is generally assumed that an enhanced spiking activity in the form of persistent reverberation for several seconds is the neural correlate of working memory [18–20]. The formation of these persistent activity patterns has been studied extensively [21–23]. Some of this work concentrated on investigating which intrinsic neuronal properties can support such activity patterns [24,25], while others focused on defining the exact activity matrix that would support attractors exhibiting localized, memory-specific, persistent activity [26,27]. We have shown recently that selective persistent activity during reactivation is an intrinsic property of an inhomogeneous dynamic memory structure [28] and is due to recurrent excitation supported by the networks with small-world (SW) topology [29]. Biologically, such heterogeneities are shown to exist [30]. Moreover, we showed that the network can regulate the stability of the persistent activity regime through change of global parameter, namely excitation. This allows the networks to undergo a seamless transition between activity regimes.

It remains unclear, however, what the dynamical underpinnings of time-dependent memory transfer from the hippocampus to the cortex are and how this dynamics is modulated by stimulus novelty. Experimental work has shown that the reactivation of a given experience during sleep is greatest when the experience is novel and diminishes with increased exposure [2,31]. Moreover, hippocampal recordings indicate that there is a significant phase shift of neural activity with respect to the hippocampal theta rhythm during the consolidation process [2], which could indicate a difference in input drives through the two hippocampal excitatory input pathways as consolidation progresses [3], as the firing of neurons in the hippocampal subfield CA1 switches from being

---

*michalz@umich.edu

aligned with the peaks of hippocampal theta oscillation to being aligned with the peaks of cortical theta rhythm. However, basic questions remain concerning (i) how the stimulus novelty is assessed from changes in localized activity patterns, (ii) how these changes are related to structural network modifications, (iii) how the hippocampal-cortical interaction regulates memory storage and erasure within hippocampus, and finally (iv) how all these processes come together to generate the experimentally observed, complex, and novelty dependent memory management scheme.

Here we show that this phenomenon can be easily explained through generic modifications of network structure which in turn evokes dynamical changes in network response. Namely, our results indicate that the dynamic formation of localized network inhomogeneities, coupled with basic anatomy of hippocampal-cortical structure, can underlie both novelty detection within hippocampal and cortical networks, as well as memory management processes based on this novelty assessment. To be able to concentrate solely on the structural network underpinnings of the observed dynamics, we use integrate-and-fire neurons; however, the results apply to biologically detailed neuronal models.

In order to more closely examine the network structural and dynamical underpinnings of these phenomena, we present each component of our model separately and discuss their implications on the novelty detection and the memory management. Both the hippocampus and the cortex were each modeled as a reduced assembly of excitatory and inhibitory networks [Fig. 1(b)] having periodic small-world topology per the Watts-Strogatz formulation [32]. This general topology was found to be present in local and global brain networks [33,34]. Dynamic small-world topology allows for simultaneous local propagation of activity as well as long-range re-injection of current, promoting formation of "on" states of persistent activity [35].

First, we show that a relatively small increase of connectivity in a discrete (i.e., well-defined) network region can play two distinctly different roles, depending on the network dynamical regime. When the network is in the low excitation regime, the changes of local network response to incoming sensory stimuli can act as a familiarity or novelty detection mechanism. However, when the global network excitation is increased, the same region will exhibit a persistent self-activation in the absence of external input. Our results indicate that the evolution of these two dynamical states correspond to observed neurobiological responses to a presentation of increasingly familiar stimulus during animal wake state and to memory reactivation experienced during sleep, respectively.

Further, we show that structural network inhomogeneities provide at the same time a dynamical mechanism of intranetwork novelty detection and internetwork signaling of the level of discrete memory consolidation within the cortical network. This last mechanism subsequently provides a self-regulated means for the hippocampus to clear already consolidated memory traces. When implemented in conjunction with a simple learning rule, as well as the assumption of fast plasticity in the hippocampus coupled with slow plasticity in the cortex, we can reproduce complex memory management processes similar to that observed in behavioral data.
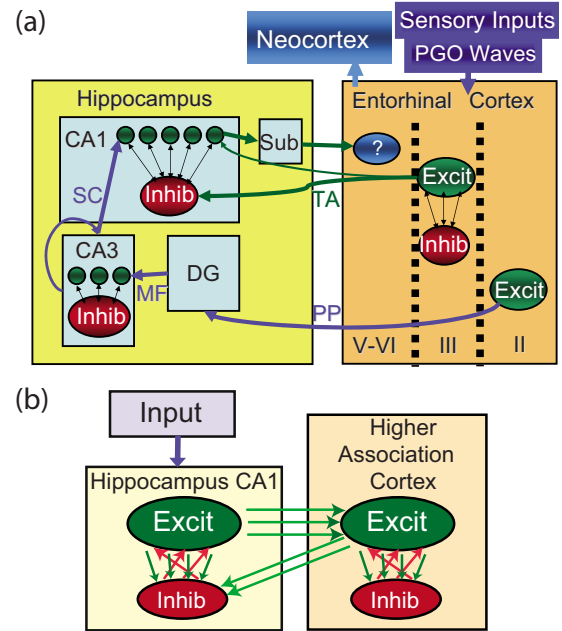

FIG. 1. (Color online) Diagrams of network structure. (a) Circuit diagram of anatomical connectivity between hippocampal and cortical structures. Entorhinal cortex layers II, III, and IV–VI project through the perforant path (PP) to the dentate gyrus (DG) and CA3, through the temporoammonic (TA) path to the subiculum (sub) and CA1, and from the CA1 and sub to the deeper layers of the entorhinal cortex, respectively. MF=mossy fibers and SC =shaffer's collaterals. (b) Diagram of model used in simulations. Single network (hippocampus or cortex): the network is composed of a larger population of excitatory neurons and a smaller population of inhibitory neurons. Both inhibitory and excitatory networks are small-world networks having periodic boundary conditions. Feedback between hippocampus and cortex: the excitatory hippocampal neurons locally innervate the excitatory cortical network (e.g., the entorhinal cortex). The cortical excitatory network suppresses the hippocampal excitatory network through random inhibitory pathways.

II. MODEL STRUCTURE AND METHODS

A. Intrahippocampal-cortical network

The two brain structures were composed of a population of 500 excitatory neurons coupled with a smaller population of 100 inhibitory neurons. The network size ratios and connection densities used were chosen to grossly reflect biological distributions and connectivity patterns in the hippocampus [Fig. 1(a)]; however, these parameters are easily modifiable without loss of observed dynamical response.

We used leaky-integrate-and-fire neurons given by

$$\tau_m \frac{dV_{i/e}^j}{dt} = -\alpha_j V_{i/e}^j + I_{i/e} + \sum_k w^{jk} I_{syn}^k \qquad (1)$$

to represent the reduced dynamics of the network elements. The i/e denotes either an inhibitory or excitatory neuron; $V_{i/e}^j$ is the membrane voltage of the jth 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 jth neuron from the kth neuron; and w^{jk} is the connection strength between neurons j and k. For the global excitatory network the local connections are established between cells such that the relative distance from one to another lies within the radius $R_e=5$, $p_g^e=0.15$ is the rewiring parameter defining the fraction of the number of local connections to the number of random, long-range ones, and the connections are of strength $w_{ex}=2$. Similarly, the global inhibitory interneuron subnetwork has $R_i=1$, $p_g^i=1$, and $w_{in}=10$, forming a random graph network. Every inhibitory cell receives input from $n_{ei}=5$ neighboring excitatory neurons with strength $w_{ei}=4$, and every excitatory neuron receives input from $n_{ie}=10$ random inhibitory ones with strength $w_{ie}=2$. Locality and relative distance were determined by considering a one-dimensional lattice with periodic boundary conditions, done for graph visualization purposes. Synaptic strengths were chosen to balance number of incoming connections so that total possible input to all cells remains the same. The external current $I_{i/e}$ is uniform over the entire inhibitory or excitatory network and functions as a global modulatory mechanism (control parameter) that mediates response transitions from low-frequency random activity, to spontaneous activation of discrete network regions, and finally to global bursting. This network architecture promotes global inhibition driven by focal excitation that creates selective, persistent reactivation patterns. For a detailed description, refer to [28].

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

$$I_{syn}^k(t) = \exp\left(\frac{-(t-t_{spike}^k)}{\tau_s}\right) - \exp\left(\frac{-(t-t_{spike}^k)}{\tau_f}\right), \quad (2)$$

where $(t-t_{spike}^k)$ 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. Aside from the deterministic input drive received from other cells, all neurons have a $p_{fire}=10^{-3}$ probability of firing spontaneously at any time step, defined as 0.5 ms.

In this reduced model, the network inhomogeneities are built into the excitatory subnetworks of both the hippocampal network and the cortical network by adding random connections to distinct nonoverlapping subgroups of excitatory neurons, i.e., neuron IDs 1–100, 101–200, 201–300, 301–400, and 401–500. The additional connections increase the density of interconnectivity within these regions beyond the average global connectivity level, allowing subgroups of neurons to recurrently innervate and effectively increasing regional excitability. These subgroups can be thought of as memory structures formed through long term potentiation (LTP) processes which are known to occur readily during exploration of a novel environment [36–39].

B. Interhippocampal-cortical feedback

In the brain, the cortex and hippocampus are connected via two main input pathways: (i) the perforant path (PP)

from layer II of the entorhinal cortex to the dentate gyrus, to CA3, and then to CA1 [Fig. 1(a) and modeled as "input" in Fig. 1(b)]; and (ii) the TA pathway directly from layer III of the entorhinal cortex to the inhibitory interneurons in the lacunosum-moleculare layer of the CA1 region and on to the subiculum [represented as higher association cortex excitatory to hippocampus inhibitory cell connections in Fig. 1(b)] [40]. These two PP and TA input pathways function separately to encode novel memories and serve as a consolidation index for familiar memories, respectively [41]. It is the slowly building familiarity index of the TA pathway that is the first step in memory consolidation which is modeled herein. To model this neurophysiology, the model network hippocampus and cortex were coupled through localized excitatory connections from the hippocampus to the cortex, and also with diffuse feedback inhibition from the cortex to the hippocampus [Fig. 1(b)]. This connectivity grossly reproduces the anatomic connectivity [Fig. 1(a)] between the two structures [40]. The one-to-one excitatory mapping from the hippocampus to the cortex is instituted for visualization purposes only; the qualitative results of this model would remain the same as long as the cortical structures, representing the long-term consolidated memories, can effectively and selectively affect hippocampal memory activation or reactivation.

C. Activity-dependent synaptic modifications

In the last stage of our simulations, we introduce self-regulated formation of new connections within the excitatory networks to show the progression of sequential memory management: rapid memory formation in the hippocampus, its reactivation in hippocampus and consolidation in the cortex, and subsequent erasure in the hippocampus. Hippocampal and cortical excitatory subnetworks are allowed to undergo synaptic modification based on spiking activity of these cells. Subnetworks are of small-world topology, with a local radius of $R_e=10$ and a rewiring parameter $p_g^e=0.15$, and are composed of 50% nonmodifiable, homogeneous, active connections with constant weight $w_{ex}=2$ as well as 50% modifiable, initially silent synapses, which are connections initially with weight 0 but can modulate their strength between 0 and $w_{ex}=2$ as a function of neuronal activity [42].

The changes in synaptic strength are implemented based on a simplified neurobiological rule of spike-timing dependent plasticity [43–46]. The synapse strength is incrementally increased when the pre- and postsynaptic neurons fire together within a set interspike interval (ISI) of $T_L=7.5$ ms, and, conversely, synaptic efficacy in the modifiable group is decreased when the two cells do not activate congruously, i.e., their ISI is above the set threshold $T_F=15$ ms:

$$\Delta w_{jk}^* = \begin{cases} \dfrac{w_{ex}}{\tau_{learn}^{h/c}} & \text{if } t_j-t_k < T_L; \\[2mm] -\dfrac{w_{ex}}{\tau_{forget}^{h/c}} & \text{if } t_j-t_k > T_F; \\[2mm] 0 & \text{if } T_L < t_j-t_k < T_F. \end{cases} \quad (3)$$

The w_{jk}^* indicates the weight of modifiable synapses between neurons j and k, $w_{ex}=2$ is the strength of nonmodifiable syn-

apses in the excitatory network, $t_j - t_k$ is the ISI between neurons j and k, and $\tau_{learn}^{h/c}$ and $\tau_{forget}^{h/c}$ are the time constants of learning and forgetting in the networks, where h/c denotes the hippocampal and cortical network, respectively. The time constants of learning and forgetting are much larger in the cortical network, reflecting slower learning (LTP) in the cortex [47–49]. We have used $\tau_{learn}^h = 7.5$ ms, $\tau_{forget}^h = 10$ ms, $\tau_{learn}^c = 25$ ms, and $\tau_{forget}^c = 200$ ms.

In this simplified model we concentrate on memory formation only within hippocampal and cortical structures. In the brain, LTP occurs both within the hippocampus, within the cortex, and between the two structures during learning. LTP occurs readily in the trisynaptic pathway from layer II of the entorhinal cortex (EC) to the dentate gyrus (DG). LTP is also easily produced in the Schaffer collateral (SC) fibers from CA3 to CA1 as noted *in vitro* and *in vivo* [41,50]. LTP in the direct temporoammonic (TA) inputs to CA1 have not been well described; indeed it is only recently that attention has been paid to this input pathway in models of hippocampal function, mostly in reference to memory consolidation as we are considering here. As was noted earlier, LTP in the TA pathway is more difficult to induce and would therefore probably occur more slowly than LTP in the trisynaptic pathway [47,49].

D. Experimental procedures for biological recording and data analysis

The experimental procedures are thoroughly described in [51]. Briefly, rats were anesthetized and implanted with a 14-tetrode drive above the hippocampus CA1 region in the brain. After surgical recovery, rats were food restricted to maintain 80–95 % of their free feeding weight, and were trained to run on a raised rectangular track for food morsel rewards placed in food cups around the edges of the track. Rats ran laps on this same track for 45 min each day to familiarize them with the environment, procedure, and recording setup.

REM sleep was characterized by lack of movement and sustained large theta (5–10 Hz) frequencies in the field potential following at least 3 min of non-REM sleep. Cell spike, field potential, and position data were recorded while the rat traversed the familiar training track for 20 min, then traversed a similar track located in a previously hidden area of the room for another 20 min, then returned for a final 20 min run on the familiar track. The same procedure of familiar-novel-familiar maze running followed by sleep recording was followed every day for a week while the initially novel maze became familiar to the animal.

The relative amplitude of the spike peak and trough, and other wave-form characteristics were used to identify nearly 100 recorded pyramidal cells and interneurons from the CA1 cell body layer. The spike times of each cell were then listed and compared with the position of the animal at the time of firing, the state of the animal, and the phase of the field potential filtered for theta. Thirty-one of the recorded CA1 pyramidal cells were selected for further analysis because they showed consistent place-specific firing (place fields) on either the familiar maze only ($n=12$), or formed a stable new place field on the novel maze ($n=19$). The firing rate of the familiar and novel place cells during the exploration phase and during the subsequent 4-h sleep period was calculated. The reactivation rate and theta pattern of cell firing during REM sleep was compared with the activity rates and patterns of the same cells during the prior exploration period. Theta phase and firing rate changes during running and REM sleep were first reported in Poe *et al.* [52].

III. RESULTS

We show below that formation of structural network heterogeneities defined as local variations of synaptic density can lead to dramatic changes in network dynamics which may underlie stimulus novelty detection and regulate memory management between the hippocampus and the neocortex. We ultimately show that this simple mechanism modulating hippocampal activation through cortical feedback reproduces the experimental data presented and, further, replicates the full process of hippocampal memory management (i.e., hippocampal storage→hippocampal reactivation →cortical storage→hippocampal deactivation). For clarity, in the sections below, we discuss each dynamical component of the phenomena separately.

A. Single network mechanism can underlie novelty detection and memory reactivation

We have shown earlier [28] that network heterogeneity may underlie selective network reactivation. Here we want to show that the structural network modifications may play a twofold role during network dynamics. Random addition of relatively few synapses (1–2 % of total possible connections) to a selected network region can dramatically change activity response of this region to stimulation when the network is in its low global excitation state (i.e., low I_e), and at the same time it can lead to formation of persistent activity state within the same region when the network is in its high global excitation state (i.e., high I_e).

To illustrate these effects we first measured network responses to a focal external drive [Figs. 2(a)–2(c)]. The network shows preferential activation of the region with added connections directly related to the magnitude of the structural network heterogeneity. Since the formation of the heterogeneity is the outcome of LTP processes incurred during learning [36–39], the changes in the intrinsic response of the network to the stimulation can be directly linked to the novelty or familiarity of the presented stimulus.

Furthermore, as we have shown before, these regions of network inhomogeneities can be spontaneously activated when network's global excitation level (I_e) is increased. Figures 2(d)–2(f) depict examples of spontaneous reactivation as a function of connectivity density within the heterogeneous region. One can observe clear reactivation exemplified in the persistent activation of the neurons within the heterogeneous network region. The reactivation itself is due to reciprocal feedback activity which is mediated by the fact that SW topology provides a structurally random yet stable reinjection mechanism supporting prolonged activation of neu-
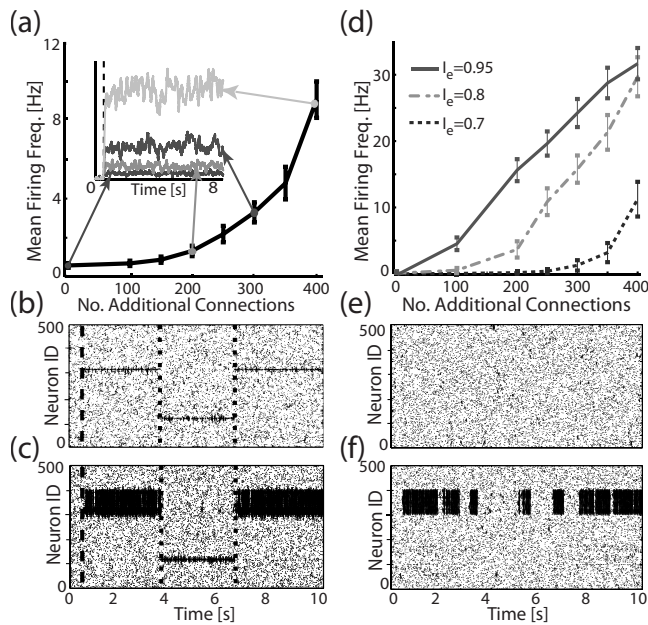
FIG. 2. Local addition of new connections changes the local response properties in the network during input stimulation and, further, produces self-activation in the absence of input when global network excitation is raised. (a) Activation of a network region, measured as a mean firing frequency of neurons in the subnetwork (neuron IDs 300–400), in response to stimulation of six cells (neuron IDs 315–320; global excitation I_e=0.6; stimulation current I_{stim}=0.7) as a function of number of added connections to the subnetwork. Activation is averaged over 20 runs and over time. Inset: sample time course of activation for four different connectivity densities (dashed line denotes onset of the stimulation). (b), (c) Sample raster plots of the network response during alternating stimulation to illustrate locality of response; neurons 315–320 are stimulated between the dashed line and first dotted line, neurons 115–120 between the first dotted and second dotted lines, and neurons 315–320 are finally stimulated between second dotted line and end of run. (b) No heterogeneities are present. (c) N=400 connections are added to the neuron IDs 300–400 region of the network. (d)–(f) Reactivation as a function of local connectivity density. (d) Mean firing frequency as a function of added connections for different values of global excitation. (e), (f) Sample raster plots depicting network reactivation when (e) no heterogeneities are present, and (f) N=400 connections are added to the region encompassing neuron IDs 300–400.

rons in spite of their refractory time [29]. The discrete localization of the reactivating region is, on the other hand, due to lowered threshold within the inhomogeneity for such dynamics to occur as well as increased inhibition spreading randomly to other network regions.

Thus we show that network structural inhomogeneity provides a dynamical mechanism mediating and modulating local, discrete network responses to stimulation, while also allowing self-reactivation under conditions of increased global excitation of the network. Here, the network dynamics can be viewed during wake behavior as an unstable attractor that becomes activated only by stimuli of appropriate characteristics, and yet during offline consolidation becomes a stable attractor which can activate spontaneously.
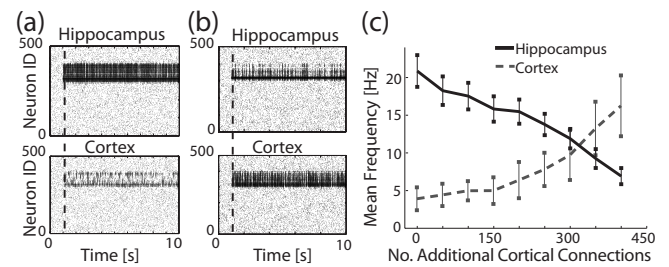

FIG. 3. Changes of the hippocampal response as a function of structural properties of cortical network. The hippocampus has one stored memory (neuron IDs 300–400), in the form of 400 additional network connections. (a) Cortex is a homogenous network (i.e., no stored memory). Six neurons in the hippocampus, IDs 315–320, are stimulated with input I_{stim}=4 and show strong activation, whereas the activation of the homogeneous cortex due to input from hippocampus is limited. (b) The cortex has a single network heterogeneity (memory) stored in the form of 400 additional connections. The hippocampus, with the same memory stored, is stimulated by the same input current, and subsequently triggers the cortical memory, which activates strongly and depresses activation of the hippocampus. (c) Mean activation of hippocampal and cortical networks as a function of added connections. As the memory is progressively stored in the cortex (i.e., becomes more familiar), cortical activation increases while hippocampal activation is depressed.

B. Modulation of hippocampal activation and reactivation by cortical feedback

Having established a common mechanism modulating both network response to stimulus based on its novelty as well as spontaneous reactivation during offline processing, we will proceed to apply this concept within an experimentally established framework of hippocampal and cortical interactions. The underlying assumption that we are making is that progressive storage (i.e., memory formation) of the presented stimulus is achieved by the formation of network inhomogeneity first in the hippocampus (i.e., fast, short-term storage) and then in the cortex (i.e., slower, long-term storage). In order to highlight the effects that cortical storage has on hippocampal activation and eliminate transient effects, we disallow synaptic modifications (i.e., learning) and examine the network dynamics at various static points of cortical memory storage.

We investigated the changes in cortical and hippocampal activation patterns as a function of the degree of regional inhomogeneity in the cortex (representing long-term memory storage) when the external stimulus is present. The hippocampal network (Fig. 3) had a single structural heterogeneity, located at neuron IDs 300–400 and created by the addition of 400 random connections within this region, and was driven by focal external stimulation applied as an additional input current (I_{stim}=4) driving six cells (IDs 315–320). One can observe that when the cortex was homogeneous, with no added connections, the stimulated region in the hippocampus was highly activated [Fig. 3(a)]. However, in the presence of cortical structural inhomogeneity, hippocampal activation was attenuated through diffuse inhibitory feedback stemming from cortical feedback excitation of hippocampal interneurons [Figs. 1(a), 1(b), and 3(b)]. In general, we see that hip-

pocampal activation systematically decreased as additional connections were added to the cortex, while cortical activity increased at the same time [Fig. 3(c)]. Therefore, the level of long-term memory consolidation in the cortex is able to control activation of the same memory in the hippocampus, serving as a novelty detection mechanism which can be utilized by the hippocampus in the consolidation process.

C. Cortical modulation of hippocampal memory reactivation

As noted before, it is thought that memory reactivation observed during sleep plays an important role in long-term memory storage as a possible memory replay mechanism mediating memory consolidation into the cortex. In such a system, it is important that consolidation, and thus reactivation, is regulated by stimulus novelty (i.e., overconsolidation of a given memory may lead to disruption of other memories, while lack of consolidation of novel memory will inhibit its storage). We postulate that, toward this end, the cortex has a novelty-dependent and memory-specific regulation of memory reactivation. We will show below that this mechanism becomes an intrinsic property within the modeled cortical-hippocampal interactions.

1. Simulation results

We demonstrate this mechanism in our hippocampal-cortical network, again in the absence of learning in order to eliminate transient, time-dependent effects. Three network regional inhomogeneities (neuron IDs 0–100, 200–300, 400–500) representing memory structures were created in the hippocampal network and kept unchanged during the simulation. At the same time, the cortical network was initially set to be homogeneous, and then new connections were progressively added to a region matching one of the hippocampal network heterogeneities (neuron IDs 200–300), to represent the progressive consolidation of that cortical memory. Figure 4(a) depicts the regional hippocampal activity in the three network regions of interest, normalized to their activity when there are no additional connections present in the cortex. One can observe a significant decrease of reactivation of the hippocampal network region [Fig. 4(a); "familiar" line], linked to the cortical region where structural inhomogeneity was progressively formed. The reactivation ratios of the other two hippocampal regions remained virtually unchanged [Fig. 4(a); "novel 1" and "novel 2"]. This indicates that the cortex can selectively deactivate reactivation of a particular network region, representing a single familiar memory, within the hippocampus while keeping the reactivation of others virtually unchanged. Figures 4(b) and 4(c) depict an example of localized hippocampal deactivation by the cortex. As soon as the familiar hippocampal region (IDs 200–300) started to reactivate, the linked cortical region immediately activated, during which activity of the whole hippocampus was inhibited. After the reactivation in the familiar memory region was abolished in the hippocampus, the cortex subsequently deactivated and other hippocampal regions (representing novel, as yet cortically unconsolidated memory) were able to again reactivate.
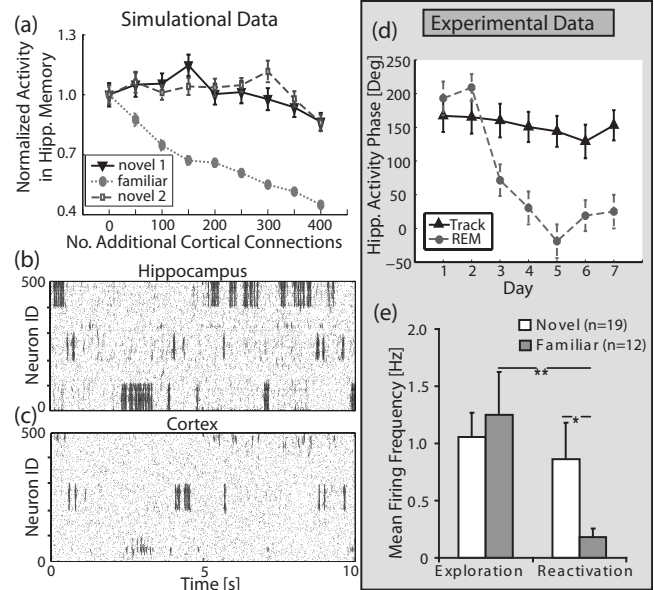

FIG. 4. Selective autonomous memory reactivation in the hippocampal-cortical structure—simulations and experimental data. The hippocampus has three regions of network inhomogeneities imbedded (IDs 1–100, 200–300, 400–500). One memory structure (IDs 200–300) is stored in the corresponding region in the cortex. (a) Average reactivation activity of familiar (neuron IDs 200–300) vs. novel (neuron IDs 1–100 and 400–500) memories as a function of additional cortical connections. The hippocampal heterogeneity, which has progressively stronger representation in the cortex, reactivates significantly less. Activity was measured by normalizing total spike counts within a memory region for the total duration of the run to total spike counts for the homogenous cortex run. Sample raster plots for (b) hippocampus and (c) cortex. (d), (e) Experimental data: (d) Phase locking of hippocampal neurons activity to peak of cortical theta oscillations as a function of days of exposure to the stimulus (i.e., stimulus novelty). Solid black line ("track") denotes phase locking to CA1 layer theta peaks during active exploration; dashed gray line denotes the same phase relation observed during REM sleep reactivation ("REM"). The neural firing becomes progressively in-phase with peak theta oscillation observed at distal cortical TA inputs. (e) frequency of hippocampal activity in novel and familiar environments during exploration (left) and sleep reactivation (right).

2. Experimental confirmation

To validate our results, we compared them with experimental findings [51]. Here we concentrate on two basic aspects: the progressive cortical involvement in hippocampal processing during memory consolidation, and changes in the offline, autonomous (i.e., not stimulus driven) hippocampal processing (i.e., reactivation).

We measure the progressive cortical involvement in hippocampal processing by monitoring the phase shift in firing of hippocampal neurons, in relation to field potential theta oscillation phase. The phases of theta oscillations in the hippocampus and in the cortex are shifted with respect to each other by 180° [3,53]. In addition, according to previous research [3,51,52], the firing pattern will be aligned with the field potential theta oscillation phase of the dominant input structure (i.e., hippocampal or cortical). The activity of hip-

pocampal neurons in relation to field potential theta oscillation phase over the time course of memory consolidation is shown in Fig. 4(d). The strong progressive shift in the activity of hippocampal neurons to fire in the trough of the hippocampal theta cycle, i.e., in phase with theta at the site of direct cortical inputs through the TA pathway, indicates that as the reactivated memory becomes increasing familiar, the cortex plays a progressively larger role in the hippocampal reactivation pattern. This supports directly our results which show that as familiarity is increased, the cortical involvement in hippocampal firing dynamics also increases.

We measure the progressive change in hippocampal offline processing by monitoring the spiking frequency of the reactivating place cells. Once the consolidated cortical TA pathway began to directly drive hippocampal reactivation [Fig. 4(d)], the spiking frequency of neurons encoding the (cortically) familiar environment decreased significantly [Fig. 4(e), right], just as predicted by the simulations. The switch in both theta phase and frequency of firing during reactivation can be explained by the consolidated cortical memory network effectively suppressing hippocampal CA1 reactivation, possibly through projections to the opioid-sensitive inhibitory neurons [47,54,55], just as we observed in our simulations.

D. Cortical modulation of hippocampal memory management sequence

The hippocampus, being a short-term memory storage location [1], is thought to perform three primary memory management tasks: store novel memory traces, reactivate these traces during quiet waking and sleep for consolidation to the cortex, and lastly erase them from itself to prevent saturation. We posit that these complex memory management processes are autonomously controlled on the basis of their familiarity. As the last part of this paper, we present the full model, with synaptic plasticity (i.e., learning dynamics) and show that localized cortical activation together with the modeled hippocampal-cortical feedback can act as a dynamical, autonomous hippocampal memory management mechanism. Self-regulation of this process within the hippocampal-cortical structure has the required and experimentally established phases (i.e., initial hippocampal learning during stimulus exposure, reactivation when the stimulus is not present, inhibition of reactivation when cortical heterogeneity is formed, and subsequent deactivation of the memory through deconstruction of the hippocampal heterogeneity). To do so, we introduce self-regulating synaptic modifications. As described in the Methods section, in this set of simulations the network is composed of both fixed synapses and modifiable ones. The modifiable synapses are initially silent [42] and become selectively active, driven by an activity dependent synaptic modification process [Eq. (3)].

During the simulation, a subset of neurons (IDs 200–300) in the hippocampal excitatory network were injected with external current at times denoted by the shaded time segments on the hippocampus raster plot in Fig. 5(a) to simulate a sensory experience, and both hippocampal and cortical networks were allowed to modify their silent synapses starting
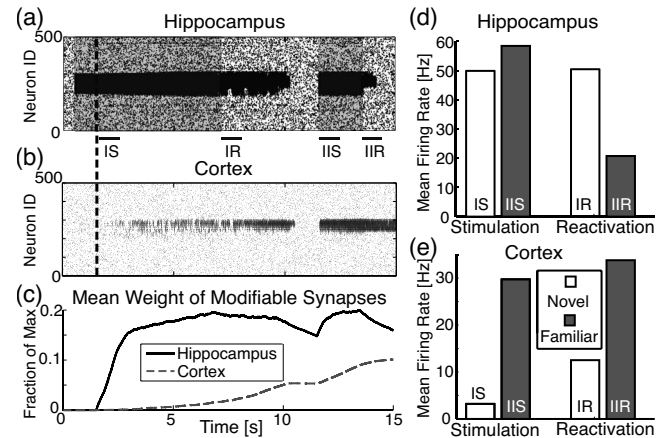

FIG. 5. Memory management through hippocampal-cortical feedback. Raster plots of activity in (a) hippocampus and (b) cortex. The hippocampus is presented with the stimulus (represented by shaded region) in neurons 200–300, immediately undergoing fast, local synaptic formation in the stimulated region. Concurrently, cortical activity driven by the hippocampus induces slow synaptic modifications in the cortex. Dashed vertical line denotes start of learning for both networks. Formation of local connections in the hippocampus allows spontaneous reactivation of the network even when the external stimulation is terminated. Activation of the cortex eventually inhibits reactivation in the hippocampus and the hippocampal network begins the process of forgetting the stored memory. Upon subsequent stimulation of this same region in the hippocampus (second shaded portion, starting at 11.5 s), the cortex immediately activates and further learns. When the stimulus is again stopped, reactivation of hippocampal region is much shorter due to stronger activation in and inhibitory feedback from the cortex. (c) Average synaptic weight of modifiable synapses is normalized to maximum possible value, w_{ex}=2. (d), (e) Histograms of spiking frequency obtained from different time regions of the simulation (denoted by the labeled black bars IS, IIS, IR, IIR) for (d) hippocampus and (e) cortex. IS labels the first stimulation time window, corresponding to novel exploration [Fig. 4(e)]; IIS labels the second stimulation (familiar exploration); IR labels the first reactivation time window (novel reactivation); IIR labels the second (familiar reactivation). All time windows are 1 s. The changes in frequency response in the hippocampus reproduce accurately the experimental data [Fig. 4(e)].

at 1.5 s (dashed vertical line). The external stimulation coupled with synaptic plasticity allowed for rapid formation of network inhomogeneity in the hippocampus, while synaptic modifications happened on a much slower time scale in the cortex [Fig. 5(c)]. When the external stimulation was stopped, the local structural changes created in the hippocampus drove its continued reactivation, allowing for further activation and structural modifications in the cortex. At a critical point, the cortical heterogeneity became large enough that its activity blocked the reactivation in the hippocampus through the interneuronal feedback [Figs. 5(a) and 5(b)]. As the hippocampus shut down, its inhomogeneity started to clear out due to its ability to quickly depotentiate the synapses, while the cortex maintained its memory structure even in the absence of stimulation or activation.

Anatomically, this depotentiation occurs through TA inputs to interneurons which have spike-blocking activity [47]

and release depotentiation-enhancing peptides [55]. Concurrently, the hippocampus increases its sensitivity to the direct TA cortical inputs [54]. Upon repeated stimulation to the same hippocampal area [second shaded region in Fig. 5(a)], we see that reactivation in the hippocampal network became shorter as the memory became progressively more familiar to the cortex and direct cortical inputs to the hippocampus through the TA pathway became more active. Figures 5(d) and 5(e) compare observed spike frequencies from the behavioral time points corresponding to the experimental data [Fig. 4(d)]. Here, one can observe a slight increase of the hippocampal spike frequency during the second stimulation period (data calculated from time period labeled "IIS"). This is due to the fact that the inhomogeneity was not completely cleared from the hippocampus, and therefore the hippocampal network activation due to the formation of network heterogeneity corresponding to this memory was stronger than the simultaneous inhibition received from the cortical network. At the same time the cortical feedback mediated dramatic shortening or abolition of reactivation when the external stimulation ceased.

IV. DISCUSSION

In this paper we show that distributed network dynamics modulated by local modifications in network structure can play a pivotal role in complex processes of memory management. Specifically, we have demonstrated that structural network inhomogeneities created through local modifications of network connectivity can act in two ways depending on the dynamical regime: they effect differential activation of the network in response to the external stimulus, and they can mediate autonomous reactivation of selective network regions. The former phase represents associative memory processes during active exploration of the environment. We have shown that differential activation during stimulation may serve as a novelty or familiarity assessment of the incoming stimulus in the cortical network, which in turn may facilitate self-controlled memory management in the hippocampal-cortical interaction. The latter phase indicates memory reactivation observed during various stages of quiet waking and sleep [1–3]. The transition from one phase to the other can be self-regulated through adjustment of the global network excitation. In the brain, such regulation is known to exist and is controlled through neuromodulatory processes [56–58].

By utilizing the structural network underpinnings of the dynamical network response together with cortical feedback we can reproduce the sequential memory management stages that have been observed in the hippocampus. This mechanism is built on several phenomena supported by experimental findings, which we have explored in this paper: discrete activation and reactivation of heterogeneous structures within hippocampal and cortical networks, cortical regulation of linked hippocampal memory structures based on familiarity level, which acts as the basis for novelty discrimination among parallel and concurrent memories, and finally spike-timing dependent plasticity of the hippocampus and cortex which occur on different time scales. We simulate hippocampal and cortical response to both novel and increas-

ingly familiar stimuli and show that, upon repeated exposure, hippocampal reactivation of the memory is lessened due to increased feedback from the cortical memory region.

We have compared the obtained results with the available experimental data. Experimental findings show that the frequency of reactivating neurons in the hippocampus coding familiar stimuli is significantly lower than the reactivation frequency while encoding novel stimuli. Furthermore, the observed progressive theta phase shift in activation of hippocampal CA1 neurons as a function of memory novelty (from in-phase with the hippocampal theta rhythm to in-phase with the theta peaks at the cortical input pathway) indicates a progressive increase in cortical driving, which is observed in our model. This is also consistent with recent research which shows a temporal correlation between cortical and hippocampal replay of consolidated memories, indicating a strong interaction between the two structures during sleep [59]. Neocortical up-down states have been shown to be phase-locked to hippocampal interneurons [60], indicating that this temporal correlation is at least partly due to their excitatory cortical drive, also in support of our model.

The reduced model presented here is not meant to faithfully reproduce every structural and dynamical aspect observed experimentally but to act as a tool to elucidate the link between structural network modifications and its dynamics during associative network storage processes—in essence, to highlight the role of network processes and dynamics in memory formation. For the sake of visual simplicity, and as a test of proof of principle, we implemented several artificial aspects into our model, such as nonoverlapping, localized memory structures. Further preliminary work shows, however, that the qualitative results presented in this paper do not change by implementing distributed, overlapping memories.

We have also not implemented any underlying oscillatory rhythms within our model cortex and hippocampus. The comparison of our results with the phase locking observed in the experimental data is only to highlight the increased role of cortical input during progressing memory consolidation. Furthermore, it is important to note that the cortical feedback itself is excitatory but that, in our model, it targets only inhibitory interneurons of the hippocampal structure. Anatomically it is known that this excitatory feedback also targets the pyramidal cells [40] and could consequently mediate the phase locking observed in the data.

Finally, the increase in cortical firing rate after consolidation that is predicted in our model [Fig. 5(e)] could also be manifested as an increase in the functional connectivity between the cortex and hippocampus via a strengthening of the TA inputs to CA1. This synaptic weight modification would have additional effect on the dynamics of hippocampal-cortical interactions and the increase in firing rates during reverberation. Thus whether slow increases in cortical firing increases the TA inputs or LTP of the TA inputs occurs slowly, the net effect on the network activity pattern is the same; increased input strength to the hippocampal inhibitory cells would effectively shut down hippocampal activity after consolidation, probably due to increased activation of the opioid sensitive interneurons in the distal stratum lacunosum-moleculare (SLM) layers [47,54,55]. Indeed, the firing rate of hippocampal reactivation is reduced in familiar

memory networks [Fig. 4(e)], and that shutdown occurs primarily at the phase of theta when the CA1 cells are most depolarized and CA3 inputs should be most capable of causing CA1 cells to fire.

Clearly these simplifying assumptions do not fully capture the complexity of hippocampal-cortical processing, but nevertheless are useful in illuminating possible network dynamical mechanisms mediating memory management as well as opening interesting avenues of future research.

ACKNOWLEDGMENTS

[1] J. J. Kim and M. S. Fanselow, Science **256**, 675 (1992).
[2] G. R. Poe, D. A. Nitz, B. L. McNaughton, and C. A. Barnes, Brain Res. **855**, 176 (2000).
[3] V. Booth and G. R. Poe, Hippocampus **16**, 161 (2006).
[4] J. L. McClelland, B. L. McNaughton, and R. C. O'Reilly, Psychol. Rev. **102**, 419 (1995).
[5] D. Tse, R. F. Langston, M. Kakeyama, I. Bethus, P. A. Spooner, E. R. Wood, M. P. Witter, and R. G. M. Morris, Science **316**, 76 (2007).
[6] P. W. Frankland and B. Bontempi, Nat. Rev. Neurosci. **6**, 119 (2005).
[7] L. R. Squire and P. J. Bayley, Curr. Opin. Neurobiol. **17**, 185 (2007).
[8] L. R. Squire, Science **316**, 57 (2007).
[9] M. A. Wilson and B. L. McNaughton, Science **265**, 676 (1994).
[10] Y. L. Qin, B. L. McNaughton, W. E. Skaggs, and C. A. Barnes, Philos. Trans. R. Soc. London, Ser. B **352**, 1525 (1997).
[11] M. Remondes and E. M. Schuman, Nature (London) **431**, 699 (2004).
[12] S. Ribeiro, C. V. Mello, T. Velho, T. J. Gardner, E. D. Jarvis, and C. Pavlides, J. Neurosci. **22**, 10914 (2002).
[13] L. R. Squire, Learn. Memory **13**, 699 (2006).
[14] P. E. Wais, J. T. Wixted, R. O. Hopkins, and L. R. Squire, Neuron **49**, 459 (2006).
[15] K. Louie and M. A. Wilson, Neuron **29**, 145 (2001).
[16] N. J. Fortin, K. L. Agster, and H. B. Eichenbaum, Nat. Neurosci. **5**, 458 (2002).
[17] S. Ribeiro, V. Goyal, C. V. Mello, and C. Pavlides, Learn. Memory **6**, 500 (1999).
[18] S. Funahashi, C. J. Bruce, and P. S. Goldman-Rakic, J. Neurophysiol. **61**, 331 (1989).
[19] J. M. Fuster, Prog. Brain Res. **85**, 313 (1990).
[20] P. S. Goldman-Rakic, Neuron **14**, 477 (1995).
[21] E. M. Galloway, N. H. Woo, and B. Lu, Brain Res. **169**, 251 (2008).
[22] B. Tahvildari, A. A. Alonso, and C. W. Bourque, J. Neurophysiol. **99**, 2006 (2008).
[23] B. Tahvildari, E. Fransn, A. A. Alonso, and M. E. Hasselmo, Hippocampus **17**, 257 (2007).
[24] G. Mongillo, O. Barak, and M. Tsodyks, Science **319**, 1543 (2008).
[25] V. Volman, R. Gerkin, P.-M. Lau, E. Ben-Jacob, and G.-Q. Bi, Phys. Biol. **4**, 91 (2007).
[26] S. Romani, D. J. Amit, and G. Mongillo, J. Comput. Neurosci. **20**, 201 (2006).

[27] A. Compte, Neuroscience **139**, 135 (2006).
[28] P. Jablonski, G. R. Poe, and M. Zochowski, Phys. Rev. E **75**, 011912 (2007).
[29] A. Roxin, H. Riecke, and S. A. Solla, Phys. Rev. Lett. **92**, 198101 (2004).
[30] Y. Wang, H. Markram, P. H. Goodman, T. K. Berger, J. Ma, and P. S. Goldman-Rakic, Nat. Neurosci. **9**, 534 (2006).
[31] C. Pavlides and J. Winson, J. Neurosci. **9**, 2907 (1989).
[32] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[33] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, J. Neurosci. **26**, 63 (2006).
[34] O. Sporns and J. D. Zwi, Neuroinformatics **2**, 145 (2004).
[35] R. C. Malenka and R. A. Nicoll, Neuron **19**, 473 (1997).
[36] M. R. Mehta, C. A. Barnes, and B. L. McNaughton, Proc. Natl. Acad. Sci. U.S.A. **94**, 8918 (1997).
[37] A. D. Ekstrom, J. Meltzer, B. L. McNaughton, and C. A. Barnes, Neuron **31**, 631 (2001).
[38] C. D. Davis, F. L. Jones, and B. E. Derrick, J. Neurosci. **24**, 6497 (2004).
[39] D. A. Nitz and B. L. McNaughton, J. Neurophysiol. **91**, 863 (2004).
[40] P. Andersen, R. Morris, D. Amaral, T. Bliss, and J. O'Keefe, *The Hippocampus Book* (Oxford University Press, Oxford, 2007).
[41] O. S. Vinogradova, Hippocampus **11**, 578 (2001).
[42] J. T. R. Isaac, R. A. Nicoll, and R. C. Malenka, Neuron **15**, 427 (1995).
[43] C. C. Bell, V. Z. Han, Y. Sugawara, and K. Grant, Nature (London) **387**, 278 (1997).
[44] G.-Q. Bi and M.-M. Poo, J. Neurosci. **18**, 10464 (1998).
[45] H. Markram, J. Lbke, M. Frotscher, and B. Sakmann, Science **275**, 213 (1997).
[46] M. Lengyel, J. Kwang, O. Paulsen, and P. Dayan, Nat. Neurosci. **8**, 1677 (2005).
[47] H. Dvorak-Carbone and E. M. Schuman, J. Neurophysiol. **81**, 1036 (1999).
[48] M. Remondes and E. M. Schuman, Nature (London) **416**, 736 (2002).
[49] M. Remondes and E. M. Schuman, Learn. Memory **10**, 247 (2003).
[50] R. J. Hussain and D. O. Carpenter, Cell Mol. Neurobiol. **21**, 357 (2001).
[51] P. Alvarez and L. R. Squire, Proc. Natl. Acad. Sci. U.S.A. **91**, 7041 (1994).
[52] G. R. Poe, B. L. McNaughton, C. A. Barnes, M. S. Suster, K.

L. Weaver, and J. L. Gerrard, Abstr. Soc. Neurosci. **22**, 1871 (1996).

[53] A. Kamondi, L. Acsdy, X.-J. Wang, and G. Buzski, Hippocampus **8**, 244 (2007).

[54] G. Maccaferri and C. J. McBain, Neuron **15**, 137 (1995).

[55] J. J. Wagner, L. R. Etemad, and A. M. Thompson, J. Pharmacol. Exp. Ther. **296**, 776 (2001).

[56] S. Gais and J. Born, Proc. Natl. Acad. Sci. U.S.A. **101**, 2140 (2004).

[57] M. E. Hasselmo, B. P. Wyble, and G. V. Wallenstein, Hippocampus **6**, 693 (1996).

[58] R. Stickgold, Trends Cogn. Sci. **2**, 484 (1998).

[59] D. Ji and M. A. Wilson, Nat. Neurosci. **10**, 100 (2007).

[60] T. T. G. Hahn, B. Sakmann, and M. R. Mehta, Nat. Neurosci. **9**, 1359 (2006).