Heteroassociative storage of hippocampal pattern sequences in the CA3 subregion

Raphael Yokoingawa de Camargo[1], Renan Schiavolin Recio[1], Marcelo Bussotti Reyes

[1] Center for Mathematics, Computing and Cognition , Universidade Federal do ABC, São Bernardo do Campo, SP, Brazil

Corresponding Author:

Raphael de Camargo[1]

Alameda da Universidade, s/n° - Bairro Anchieta - São Bernardo do Campo, SP, 09606-045, Brazil

Email address: raphael.camargo@ufabc.edu.br

Heteroassociative storage of hippocampal pattern sequences in the CA3 subregion

Raphael Y. de Camargo[1], Renan S. Recio[1], and Marcelo B. Reyes[1]

[1] Center for Mathematics, Computing and Cognition , Universidade Federal do ABC, São Bernardo do Campo, SP, Brazil

ABSTRACT

Background. Recent research suggest that the CA3 sub-region of the hippocampus have properties of both autoassociative memory, due to its ability to complete partial cues, tolerate noise, and store associations between memories, and heteroassociative memory, due to its ability to store and retrieve sequences of patterns. Although there are several computational models of the CA3 as an autoassociative network, more detailed evaluations of its heteroassociative properties are missing. **Methods.** We developed a model of the CA3 subregion containing 10,000 integrate-and-fire neurons and recurrent excitatory and inhibitory connections, and which exhibits coupled oscillations in the gamma and theta ranges. We stored thousands of pattern sequences using a heteroassociative learning rule with competitive synaptic scaling. **Results.** We showed that a purely heteroassociative network model can (i) retrieve pattern sequences from partial cues with external noise and incomplete connectivity, (ii) achieve homeostasis regarding the number of connections per neuron when many patterns are stored when using synaptic scaling, (iii) continuously update the set of retrievable patterns, guaranteeing that the last stored patterns can be retrieved and older ones forgotten. **Discussion.** Heteroassociative networks with synaptic scaling rules seem sufficient to achieve many desirable features regarding connectivity homeostasis, pattern sequence retrieval, noise tolerance and updating of the set of retrievable patterns.

Keywords:

INTRODUCTION

Even though it is well established that hippocampal formation is responsible for temporary storage and retrieval of memories (Squire, 1992; Andersen et al., 2007), memory coding is still not completely understood. The CA3 subregion contains a particularly large number of recurrent connections among the pyramidal neurons. Existing attractor neural network models (Hopfield, 1982; Amit, 1989) show that networks with recurrent connections can use a Hebbian learning rule to store patterns, represented as a set of active neurons, and retrieve them from partial cues. Based on these models, Rolls et al. (1997) proposed that the CA3 subregion could work as an autoassociative memory, where neurons from the same pattern have recurrent excitatory connections between them, which enable associations between features of a memory, retrieval of stored memories from presentation of partial cues, and noise tolerance during memory retrieval. These characteristics of autoassociative networks make them excellent candidates for the storage of complex episodic memories (Rolls, 2010), which contain smaller components organized in a flexible way and that could be used as cues to retrieve a whole episode (Cohen and Eichenbaum, 1993).

Another important characteristic of the CA3 subregion is that, during exploratory behavior in rats, this area shows a local field potential (LFP) signal composed of gamma (40-100Hz) components nested in a theta (5-10Hz) rhythm (Bragin et al., 1995; Colgin, 2016). Moreover, O'Keefe and Recce (1993)

observed that the phase of the theta cycle at which a place cell is activated depends on the distance of the rat to the cell preferred place, an effect called theta phase precession. Jensen and Lisman (1996) proposed that the hippocampus could work as a heteroassociative memory (Sompolinsky and Kanter, 1986), a kind of network that stores sequences of patterns, represented by neuronal ensembles. Similarly, Levy (1996) proposed a simple model composed of McCulloch-Pitts threshold neurons that could learn to store sequences using local associative learning rules. The main difference to autoassociative memory is that neurons from a pattern have recurrent connections to neurons from the next pattern in the sequence. The advantage is that an entire sequence of patterns can then be retrieved in a theta cycle, with one pattern per gamma cycle nested in the theta cycle. But the lack of autoassociative could make these network less reliable in the presence of internal and external noise and unable to perform pattern completion. To combine the properties of auto and heteroassociative networks, models containing both kinds of connections were proposed. In these models, recurrent connections within a single CA3 area would work as autoassociative connections, while heteroassociave connections would be located either as feedback connections from CA3 to dentate gyrus (Lisman et al., 2005) or between different CA3 areas (Samura et al., 2008).

More recently, the concept of theta sequences (Foster and Wilson, 2007) was proposed, based on experimental evidence that sequences of neuronal ensembles are sequentially activated at subsequent gamma cycles and coupled to theta oscillations. Activations occur preferentially at some phases of theta cycles and sequences seem to be bounded within theta cycles. In contrast to phase precession, the appearance of theta sequences appears to require learning (Feng et al., 2015) and was associated with the representation of current goals (Wikenheiser and Redish, 2015) of rodents. The relationship of theta sequences with episodic (Wang et al., 2014) and spatial memories (Dragoi and Buzsáki, 2006) indicates that the CA3 recurrent and CA3-CA1 connections may encode these memories as sequential activation of neuronal ensembles.

Another feature of the hippocampus is that memories appear to be temporarily stored there and later coded as remote memories in the cortex (Frankland and Bontempi, 2005). This property could be due to the limited capacity of auto- and heteroassociative networks, since an increase in the number of stored memories leads to a catastrophic interference effect (Amit, 1989), where the saturation of connection weights makes all stored memories unavailable. Although the replacement of older memories by newer ones also receives the name of catastrophic forgetting in other types of networks (French, 1999), this effect is consistent with one of the hypothesized roles of the CA3 as a temporary memory storage. Proposals for enabling this replacement of older memories include the use of STDP rules (Caporale and Dan, 2008) and memory recall and consolidation during sleep (Born et al., 2006).

Although there are several computational models of the CA3 as an autoassociative network (Rolls et al., 1997; Cutsuridis and Wennekers, 2009), more detailed evaluations of its heteroassociative properties are missing. In this work, we use a heteroassociative network model of the CA3 subregion containing 10,000 neurons to investigate the following properties: (i) retrieval of pattern sequences from partial cues with external noise and incomplete connectivity, (ii) achievement of homeostasis regarding the number of connections per neuron when many patterns are stored using synaptic scaling, (iii) replacement of the set of stored patterns, guaranteeing that the last stored patterns can be retrieved and older ones forgotten.

METHODS

The network model

We implemented a heteroassociative network, that simulates the CA3 subregion, composed of 10,000 integrate-and-fire neurons representing pyramidal cells. Neurons are connected by excitatory recurrent connections, with kinetic models of AMPA synaptic channels and weights defined by a heteroassociative learning rule. The network is shown in Figure 1.
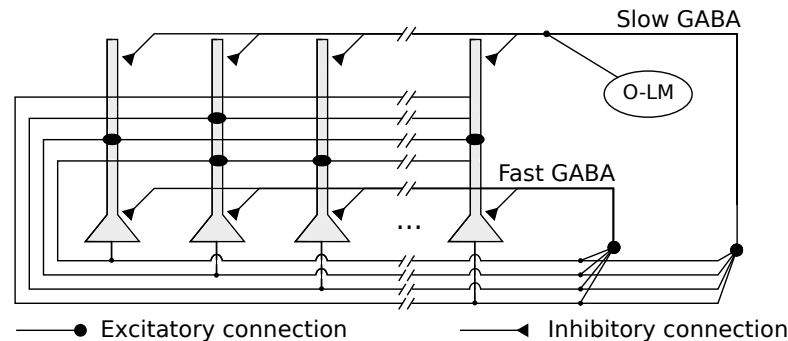
Figure 1. Network architecture, with the recurrent connections among pyramidal neurons, fast and slow GABAergic connections, and O-LM inhibitory population.

The network has two kinds of feedback inhibition, mediated by fast and slow GABA channels (Pearce, 1993), modeled as direct connections between every pair of pyramidal neurons. The fast inhibition is responsible for controlling the level of activity in the network and reducing the interference between stored pattern sequences. The slower inhibition helps to control the level of activity in the network in a slower time scale. We also simulated periodic inhibition from a oriens-lacunosum-moleculare (O-LM) interneuron (Freund and Buzsáki, 1998) population, modeling it as a single spike generator representing the intrinsic firing rate in the theta range (5 Hz) of these cells when depolarized (Gloveli et al., 2005). It sends connections to the slow GABA channel of every pyramidal neuron and generates the theta rhythm in the network.

Following the model used by de Almeida et al. (2007), we model the membrane potential V_n of pyramidal neuron n by:

$$\frac{dV_n}{dt} = \frac{1}{\tau_m}((I_{syn} + I_{rep}) * r_{Input} - V_n + V_{rest})$$

using the values $r_{Input} = 33\,\text{M}\Omega$, $V_{rest} = -60\,\text{mV}$, $\tau_m = 2\,\text{ms}$. When V_n reaches a threshold of $-50\,\text{mV}$, it is reset to V_{rest}, with a refractory period of 13.3 ms. The current I_{rep} is an hyperpolarizing current that provides spike frequency adaptation and is dependent on the time of the last generated spike t_{spk}. It is given by:

$$I_{rep} = -560\,pA * \exp\left(-\frac{t - t_{spk}}{5\,ms}\right)$$

Finally, I_{syn} represents the synaptic current into the neuron, and is given by:

$$I_{syn} = I_{ext} + I_{ampa} + I_{gaba} + I_{gabaS}$$

87   where I_{ext} represents the external input, I_{ampa} the excitatory AMPA synapses, I_{gaba} the inhibitory fast
88   GABA synapses, and I_{gabaS} the slow GABA synapses. We used the Euler method with an integration step
89   of $0.1\,\text{ms}$, which is sufficient for simple integrate-and-fire models.

90   **The synaptic model**
91   We modeled the synaptic channels using a dual exponential model, given by:

$$I_k = A_k \sum_s W_s \frac{[\Delta t_s]_+}{\tau_1 - \tau_2} \left(\exp\left(-\frac{\Delta t_s}{\tau_1} \right) - \exp\left(-\frac{\Delta t_s}{\tau_2} \right) \right)$$

92   where $[\Delta t_s]_+ = t - t_s - t_{delay}, \forall t > t_s - t_{delay}$ is the time since the spike s generated at time t_s was delivered
93   considering the delay t_{delay}, and 0 otherwise. A_k represents the maximum conductance and W_s the synaptic
94   weight, i.e. the connection strength between the neurons. Parameters τ_1 and τ_2 are the time constants, and
95   the double exponential is reduced to the alpha function when $\tau_1 = \tau_2$:

$$I_k = A_k \sum_s W_s \frac{[\Delta t_s]_+}{\tau_1} \exp\left(1 - \frac{\Delta t_s}{\tau_1} \right)$$

96   We used the $\tau_1 = 2\,\text{ms}$, $\tau_2 = 8\,\text{ms}$ and $A_k = 3200\,\text{pA}$ for AMPA channel of recurrent connections
97   (Spruston et al., 1995), $\tau_1 = \tau_2 = 2\,\text{ms}$ and $A_k = 3200\,\text{pA}$ for the external input channels, $\tau_1 = \tau_2 = 5\,\text{ms}$
98   and $A_k = 540\,\text{pA}$ for the fast GABA channels (Pearce, 1993), and $\tau_1 = 7\,\text{ms}$, $\tau_2 = 57\,\text{ms}$ and $A_k = 30\,\text{pA}$
99   for the slow GABA channels (Pearce, 1993). We did not include NMDA channels, which have a main
100   role in LTP and LTD processes, since our learning phase is executed offline, as we discuss in subsection
101   "Learning rule".

102   We placed all neurons in a plane and defined the axonal delay between each pair of pyramidal neurons
103   based on their euclidean distance. We used an axonal propagation velocity of $300\,\mu\text{m/s}$ (Meeks and
104   Mennerick, 2007) and considered a square $2 \times 2\,\text{mm}$ area, based on anatomical data (Amaral and Witter,
105   1989), resulting in a mean axonal delay of about $3.3\,\text{ms}$ (Aaron and Dichter, 2001). Since we are using
106   integrate-and-fire neurons, we also included the excitatory postsynaptic potential (EPSP) propagation
107   time to the soma. We used a delay of $5\,\text{ms}$, estimated by experimentally measured differences between
108   the time to peak of EPSPs in the soma from mossy fiber inputs, near the soma, and recurrent connections,
109   at apical dendrites (Miles and Wong, 1986).

110   For inhibitory feedback connections we used a delay of $2.5\,\text{ms}$, based on studies on interneurons from
111   the CA3 (Diego et al., 2001). For the slow inhibition, we used a delay of $10\,\text{ms}$, to include the fact that
112   these inhibitory cells are located near the basal positions of the pyramidal cells and makes connections to
113   the upper apical positions (Hájos et al., 2004).

114   External noise was modeled as a Poisson process with $1\,\text{Hz}$ rate per neuron, representing independent
115   inputs to the neurons not related to the retrieval of sequences. We used weights that were large enough to
116   cause the network neurons to fire spontaneously at the rate of $0.75\,\text{Hz}$, which is similar to the rate $0.80\,\text{Hz}$
117   produced during pattern retrieval in the simulations.

118   **Pattern sequences**
119   We created the patterns by generating random permutations, without replacement, of the neuron indexes
120   and getting the first r values of the permutation, where r is equal to the number of neurons times the
121   pattern density for the simulation. We define the density as the number of neurons active in a pattern

122 divided by the total number of neurons. For heteroassociative networks, there are estimates by de Almeida
123 et al. (2007) of density of 0.003. For our simulated network with 10,000 neurons this would result in only
124 30 neurons per pattern, which would be insufficient for noise tolerance and sequence completion. We use
125 the term sequence completion to mean that the next pattern in the sequence can be retrieved correctly even
126 when the current pattern was only partially activated. Although using more neurons per pattern decreases
127 the storage capacity of the network, we decided to use a density of 0.01, resulting in 100 neurons per
128 pattern.

129 We define a pattern p^μ as a binary vector with size equal to number of pyramidal neurons in the
130 network. Each position p_i^μ has value 1 if neuron i is part of p^μ and 0 otherwise. We considered
131 sequences of 7 patterns, which appears to be the typical number of patterns that can be coded inside theta
132 cycles (Lisman and Idiart, 1995; Lisman and Jensen, 2013), and stored them using a heteroassociative
133 learning rule (Sompolinsky and Kanter, 1986) that associates each pattern in the sequence with the
134 subsequent one. For example, in a sequence of patterns s1-s2-s3-s4-s5-s6-s7, neurons from pattern s1 will
135 have excitatory connections to neurons from pattern s2, which in turn connect to neurons from s3, and so
136 on. We included additional connections from the last to the first pattern, permitting sequence retrieval
137 using a cue from any of its patterns.

Learning rule

139 The model uses a heteroassociative learning rule (Sompolinsky and Kanter, 1986) for pattern sequence
140 storage. We simulated an offline learning procedure, performed before the simulation, by considering that
141 the patterns from each sequence were presented in sequence. This enables the usage of an LTP-like rule
142 to learn associations between subsequent patterns in each sequence, which can be summarized as:

$$\Delta w_{ij} = \gamma^+ p_i^\mu p_j^{mod(\mu+1,k)}$$

143 where w_{ij} represents the weight of the connection from neuron i to neuron j, $\gamma^+ = 1$ the learning rate, p_i^μ
144 the neuron i from pattern μ and $p_j^{mod(\mu+1,k)}$ the neuron j from the next pattern in a sequence of k patterns,
145 $\mu = 0, \ldots, k-1$. The function $mod(x,k)$ represents the modulo operation, with $mod(-1,k)$ defined as
146 $k-1$. This produces an association between the last and the first patterns in the sequence. We also used
147 an LTD-like learning rule that performs a dissociation between a pattern and the previous pattern in the
148 sequence, summarized as:

$$\Delta w_{ij} = -\gamma^- p_i^\mu p_j^{mod(\mu-1,k)}$$

149 We executed simulations with and without LTD, in which case we used $\gamma^- = 1$ and $\gamma^- = 0$, respec-
150 tively.

151 We included synaptic scaling (Abbott and Nelson, 2000) as an heterosynaptic long-term depression
152 (hLTD) mechanism (Chistiakova and Volgushev, 2009), where we update the weights of synapses that
153 were not active during the learning phase. Synaptic scaling models a competition between the synapses of
154 a single neuron. After learning a set of 100 pattern sequences, we apply an additive synaptic scaling rule
155 to all neurons. We decrease all synapses from each neuron by a fixed amount, defined as the ratio between
156 the total weight increase since the last scaling procedure and the number of connections. Synapses that
157 become negative are set to zero and the difference is uniformly subtracted from the remaining synapses.

In biological networks, anatomical constraints can limit the set of neurons to which a neuron can connect. We simulate these constraints using different levels of initial connectivities, where we randomly define for each neuron the set of neurons to which it can connect. For an initial connectivity of 0.6, each neuron initially connects to 60% randomly selected neurons in the network, with random weights drawn from a uniform distribution between 0 and a defined maximum value. Initial connection weights are scaled to maintain the excitatory drive with different initial connectivities comparable. Connections to neurons outside the set of initial connections cannot be created during learning.

The learning rule is then sequentially applied to each pattern sequence to be stored. During learning connections are eliminated due to LTD or synaptic scaling, so that after storing many sequences the network stabilizes at a lower connectivity, effectively simulating a "lifetime" of episodic sequences learning until maturity.

Sequence retrieval

We defined as a pattern cue the stimulation of a subset of the neurons that are active in the pattern. We always provide inputs for same number of neurons, so that with a cue size of 1.0, we stimulate all neurons from the pattern and no other neuron. For instance, with a cue size of 0.7, only 70% of the neurons in the pattern receive the input, while random neurons outside the pattern are selected to keep the input size controlled over different simulation runs. We consider that during retrieval the entorhinal cortex (EC) provides the input to the CA3 via the perforant path, using a coding where an ensemble of neurons, representing an input pattern, fire synchronously inside a single gamma cycle (Chrobak and Buzsáki, 1998). In this case, each neuron stimulated by the cue receives a single spike as input, with a synaptic strength large enough to cause its firing. A second input to the CA3 are the mossy fibers, originating from the dentate gyrus, which appears to deliver sparse patterns for storage in the CA3 subregion (Treves and Rolls, 1992). Since in this work the learning phase was performed offline, we did not include this input in the simulation.

We determine the set of active neurons by checking which neurons fired inside a time window of 10ms. The measurements are performed every 2 ms, with the time window centered at simulation time t. We used a small time window to permit capturing active neurons only from a single gamma cycle at each time t in the middle of the cycles. We define the *overlap* of the active neurons with a stored pattern p^μ as:

$$overlap(t, p^\mu) = \frac{1}{n_{p^\mu}} \sum_{i \in p^\mu} o_i(t) \cdot p_i^\mu$$

where n_{p^μ} is the size of p^μ, o_i is the state of neuron i at the instant t, which can be 0 (inactive) or 1 (active), and p_i^μ is 1 if neuron i is part of p^μ and 0 otherwise. The overlap equals to 1 when all the neurons from a stored pattern are active. Neurons also have spontaneous firings due to external noise and, to correctly evaluate the pattern retrieval, we consider in the metric only the neurons which are part of the pattern. This could be applied because the number of active neurons was always comparable to the number of neurons per pattern. To verify that extra active neurons are randomly activated and not concurrent retrieval of multiple patterns, we also compute the highest overlap among all other patterns.

We evaluated sequence retrieval in the simulated neuronal network using the mean overlap over 10 simulations for each combination of parameter values. We stored 10010 patterns, a multiple of the sequence size of 7 patterns, in the network. Each simulation was executed for 86150 ms, which was sufficient to evaluate the retrieval of the last 3003 stored patterns. The simulation and experimental results analysis was implemented and executed in MATLAB. Simulation source code and raw and

198 processed data are available at a public repository (`https://doi.org/10.6084/m9.figshare.`
199 `5378347.v2`).

RESULTS

The number of retrievable patterns and connections are stable during the network life-time

Our evaluation shows that the number of connections in the recurrent network, initialized with full connectivity and random weights, stabilizes after learning a few hundred patterns, remaining stable as new patterns are added (Figure 2A). Storing a pattern sequence increases the total synaptic weight of each neuron, which is compensated by the synaptic scaling rule. This reduces the weights of all synapses in the neuron, causing some of them to disappear as they approach zero. The number of connections decreases until the number of connections created by learning new patterns is balanced by the removal of old connections. This occurred either when using only synaptic scaling (SS), in which case we set $\gamma^- = 0$, and when using SS combined with LTD (SS-LTD), with $\gamma^- = -1$, showing that synaptic scaling is sufficient to control the number of connections in the network.

The number of retrievable patterns initially increases with the number of stored patterns, followed by a decrease toward a constant value, as shown in Figure 2B. To determine if a pattern p^{mu} is retrievable, we compute the sum of excitatory weights S_i from neurons from the previous pattern p^{mu-1} toward each neuron i in the network. If S_i for each target neuron from p^{mu} is larger than the S_i of each neuron outside p^{mu}, the pattern is considered retrievable. This means that neurons in the pattern receive a larger excitatory drive than the other neurons during pattern retrieval, which permits a global inhibition to be used to suppress the activity of these other neurons. This simple measurement generates an estimate of the number of retrievable patterns directly from the connection weights. The alternative would be performing a simulation where, for each pattern sequence, we would provide a cue of one of its patterns and check if all patterns in the sequence are retrieved correctly. But this would be unfeasible, since we would have to perform simulations to retrieve every pattern for all network configuration combinations, where we would vary the number of stored patterns, pattern density level, initial connectivity and LTD rule.

The initial increase in number of retrievable patterns (Figure 2B) is due to the lack of patterns to retrieve in the beginning and the later decrease occurs due to the interference between stored patterns, called crosstalk. After reaching an equilibrium point, the number of retrievable patterns remains constant. Using a pattern density of 0.02, instead of 0.01, increased the number of connections, causing an increase in the crosstalk between patterns, which resulted in a reduced number of retrievable patterns (Figures 2A-B). Combined with LTD, which caused a larger number of weight decreases, the number of retrievable patterns became almost zero. With density 0.005, the final number of connections was smaller and the number of retrievable patterns higher, showing the benefits of using fewer neurons per pattern. Although a smaller density may reduce the reliability of the network due to neuron firing and synaptic variability, a region containing one million neurons would still have about 5000 active neurons per pattern with a density of 0.005.

Including LTD (SS-LTD) caused only small changes in the number of connections, but the number of retrievable patterns decreased (Figures 2A-B). Although this result might seem to occur due to extra reductions in excitatory weights caused by LTD, we compensate for this by reducing the effect of synaptic scaling, so that the total weights per neuron with or without LTD will be the same. It appears that using LTD causes disruptions in stored pattern sequences when decreasing the connection weights from neurons of the last stored pattern to neurons of the previous pattern.

Anatomical constraints may prevent full connectivity in the CA3 subregion. An initial connectivity of 0.6 means that a neuron can connect to 60% of the neurons. Using a smaller initial connectivity reduces the number of connections, but the number of retrievable patterns is also reduced (Figures 2C-D). But it is interesting that using an initial connectivity of 0.4 and synaptic scaling (SS) alone, there were about 900 retrievable patterns, which is more than the 600 patterns obtained when using LTD (SS-LTD) with full initial connectivity. Also, although 900 patterns may be too few, this value would probably be higher for larger networks from biological systems.
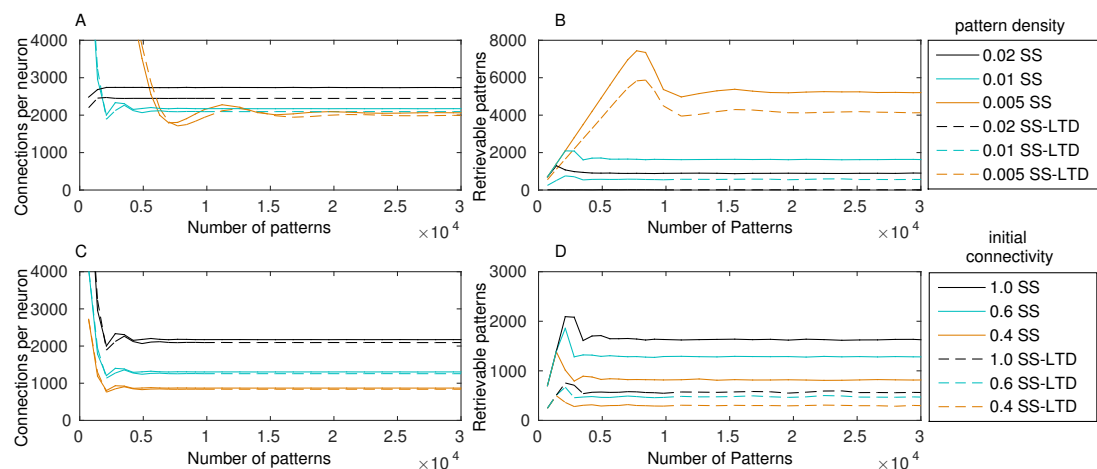
Figure 2. (A) Number of connections in a network with 10,000 neurons after learning a set of pattern sequences when applying only the synaptic scaling (SS) rule, for a pattern density of 0.02, 0.01 and 0.005, and with synaptic scaling and LTD (SS-LTD); (B) Number of retrievable patterns for the same scenarios from (A); (C-D) Same as (A-B), but comparing neuronal initial connectivities of 1.0, 0.6 and 0.4. The graph shows the mean of three executions. Standard deviations values were lower than 10 connections for (A) and (C), 70 patterns for (B) and 35 patterns for (D) and are not shown for clarity.

Increasing the initial total weight per neuron increases the number of connections, caused by less competition in the synaptic scaling rule (Figure 3A). It also increases the number of retrievable patterns, as old patterns take longer to be forgotten, until a transition point where the crosstalk interference becomes larger than the pattern inputs (Figure 3B). After this point, the number of retrievable patterns starts to decrease. The behavior is similar both with and without the usage of LTD, but in the latter the number of retrievable patterns was larger.

Another way to analyze the results is to check how many patterns can be stored in a network with a given number of connections per neuron. For instance, de Almeida et al. (2007) estimated that neurons in the CA3a subarea connect to 20% of the neurons. In our network, it would result in about 1600 retrievable patterns and would be far from the transition point. Including more neurons or using a smaller pattern density, as is the case in biological systems, would result in a much larger number of retrievable patterns.

The final number of connections in the network is dependent on the initial random weights between neurons, on the connection extent, and on the density of the patterns (Figures 3A,C). With larger initial weights, the sum of synaptic weights toward each neuron will be larger and, consequently, fewer connections will be removed by the effect of synaptic scaling and LTD.

It is also interesting that although a smaller initial connectivity would permit the storage of more patterns with the same number of connections, we actually see a reduction in the number of retrievable patterns (Figures 3A,B). This is caused by the smaller signal-to-noise ratio in the system, as there will be

proportionally fewer coding connections. If we consider the estimated connectivity of 20% (de Almeida et al., 2007), which full connectivity we would have to use an initial weight of 2.0, resulting in about 1400 retrievable patterns. But with an initial connectivity of 60%, it is possible to use a larger initial weight of 6.0, while maintaining the total connections to 20% of the neurons, which resulted in about 1800 retrievable patterns, an actual improvement over the full initial connectivity. With an extent of 40%, the number of retrievable patterns started to decrease with more than 1200 connections per neuron.

Using a density of 0.005 caused small changes in the final number of connections, but the number of retrievable patterns increased (Figures 3C-D), which would be more typical for biological larger networks with lower density. The exception is for the initial connectivity of 0.4, due to the small number connections per neuron in the simulated network.
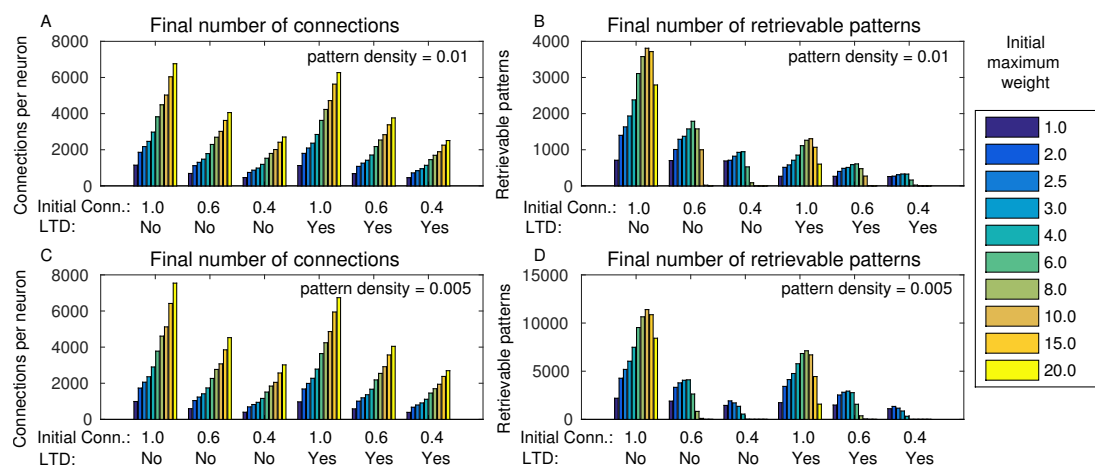
Figure 3. (A) Number of connections after learning 100,000 patterns and applying the synaptic scaling rule for different total weights per neuron, for a pattern density of 0.01, initial connectivities of 1.0, 0.6 and 0.4 with and without LTD; (B) Number of retrievable patterns for the same scenarios from (A); (C-D) Same as (A-B), but with density 0.05. The graph shows the mean of three executions. Standard deviations values were lower than 3 connections for (A) and (C), 32 patterns for (B) and 63 patterns for (D) and are not shown for clarity.

In this section we evaluated the static properties of the connections and the number of retrievable patterns. In the following sections, we perform actual network simulations to evaluate the dynamic operation of the network. In all simulations, we used a density of 0.01, with 100 neurons per pattern, to improve the reliability of the simulations. For the initial random weights, we used the values of 2.5, 3.0, 6.0 and 4.0 for the initial connectivities 1.0, 0.8, 0.6 and 0.4, respectively. These are the values that maximize the number of retrievable patterns when neurons connect to approximately 20% of all neurons.

A heteroassociative network enables the retrieval of stored pattern sequences in the presence of noise

The simulated heteroassociative network could retrieve complete pattern sequences stored on its connection weights after presentation of a partial cue from any pattern in the sequence. Although this retrieval is expected in a noiseless environment, we considered the scenario with a level of noise that caused the neurons to fire spontaneously, without presentation of pattern cues, with a mean spike rate of 0.75Hz. The network oscillates in the theta band at 5 Hz, shown in the approximate LFP signal obtained by sum of the membrane potential of the pyramidal neurons (Figure 4C). This oscillation in 5Hz is due to the periodic inhibition of O-LM cells, with a sharp decrease in the spike rate when the inhibition starts and

291 a gradual recovery of the spike rate when the inhibition fades. Also, no pattern is retrieved during the
292 period (Figure 4A), despite the firing of near 200 neurons at some points (Figure 4B), showing that the
293 firing was random, which is also illustrated in the raster plot (Figure 4D). When a random external input is
294 presented at instants 4200ms and 4400ms, the target neurons are activated, but since they do not represent
295 a stored pattern, this extra activity disappears.
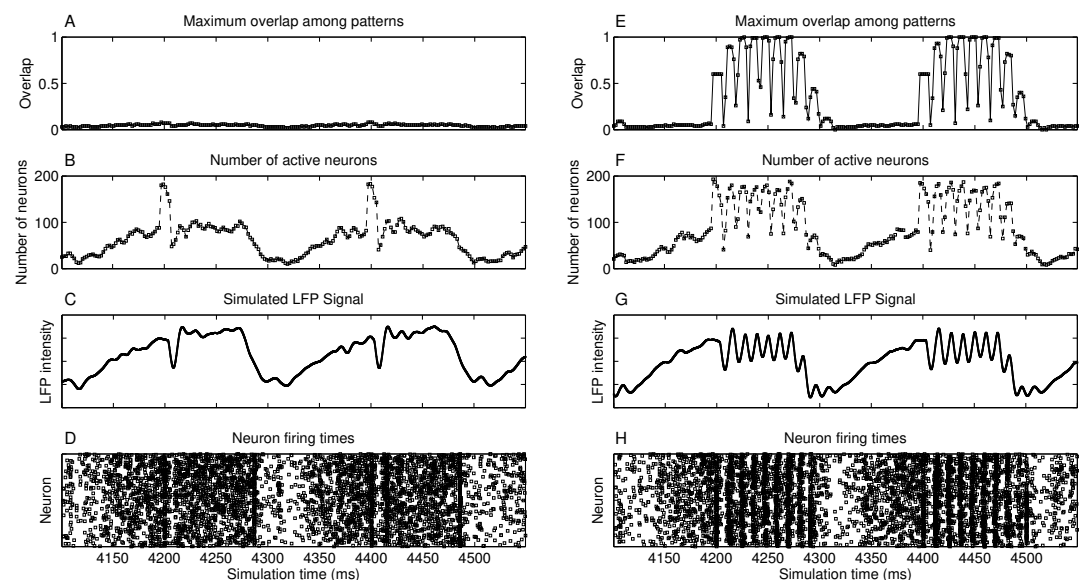
Figure 4. Operation of two heteroassociative networks with 10^4 pyramidal neurons and 10^4 stored patterns with 100 neurons per pattern and coded into sets of 7 patterns. The left-hand side graphs show 450 ms of network operation with presentation of random inputs at instants 4200 ms and 4400 ms, extracted from a 5000 ms simulation. (A) Maximum value of the overlap among all stored patterns; (B) number of active neurons as a function of time; (C) simulated LFP signal from the network; (D) raster plot showing instants of neuronal spikes for all neurons in the network; and (E-H) same as (A-D), but with presentation of partial cues of 0.6 of a pattern at instants 4200 ms and 4400 ms.

296   Presenting a pattern cue caused the retrieval of the complete pattern sequence just after its presen-
297 tation, at 4200 ms and 4400 ms in Figures 4E-H. Each pattern is retrieved during a small time window,
298 corresponding to a gamma cycle inside the theta cycle, as occurs in the theory proposed by Lisman (2005).
299 There are about nine gamma cycles per theta cycle and since each pattern sequence has seven patterns, the
300 complete sequence is retrieved during a single theta cycle. The last pattern in the sequence is connected to
301 the first, which causes the number of gamma cycles to be larger then the sequence size. We can also see
302 that although we present only 60% of the first pattern of the set (cue size of 0.6), the network successfully
303 retrieved and completed the subsequent patterns. Finally, the raster plot of neuronal spikes shows the
304 clustering of neuronal firings during each pattern retrieval, with periods of inactivity between gamma
305 cycles.

306   Longer or smaller pattern sequences can also be stored and retrieved. With a larger sequence, a subset
307 of the pattern sequence is retrieved in a single theta cycle, until the slow inhibitory current starts to win
308 the competition against the excitatory drive. The remaining patterns can then be retrieved in the next
309 theta cycle, by presenting the last pattern retrieved in the previous cycle. With smaller pattern sequences,
310 the theta cycle would just keep looping the patterns in the sequence until the inhibition surpasses the
311 excitatory drive and finishes the theta cycle.

312   In simulations from Figures 4E-H, we presented the pattern cues at the beginning of the theta cycles.

313 Presenting the cues at other phases would cause only part of the sequence to be retrieved, since the
314 periodical inhibition from O-LM cells would cause the pyramidal cells to stop firing before the sequence
315 was finished. Since phase synchronization may be required for memory retrieval in the hippocampus (Fell
316 and Axmacher, 2011), it seems reasonable to provide the inputs at specific theta phases.

317   Sequence completion during retrieval occurs due to the heteroassociative connections from pattern j to
318 next pattern $j+1$. Since many neurons from pattern j must fire to activate pattern $j+1$, randomly firing
319 neurons do not cause random pattern retrievals. Sequence completion also works with initial connectivity
320 of 0.6, as shown in Figure 5A. Also, the fast inhibition guarantees that only a single pattern is retrieved at
321 a time. After evaluating the overlap with all other stored patterns, the second highest overlap was always
322 lower than 0.1 (Figure 5B). Finally, using LTD did not cause perceptible changes in the mean overlap
323 values (Figures 5C-D).
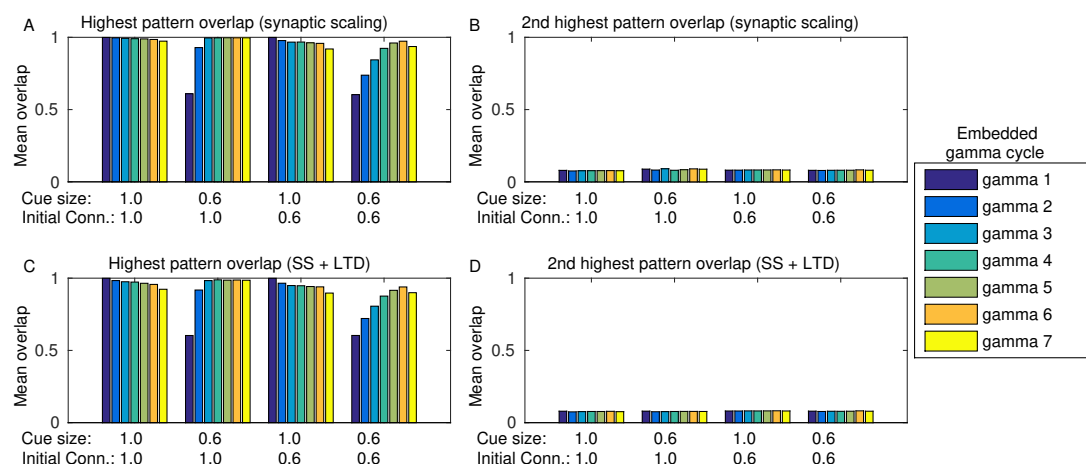
Figure 5. Mean overlap during pattern retrieval within each gamma cycle inside a theta cycle. Gamma 1 refers to the first gamma cycle in each theta cycle, gamma 2 to the second, and so on. (A) The highest overlap among patterns for different cue sizes and initial connectivities and using synaptic scaling only; (B) same as (A), but showing the second highest overlap among all patterns; (C) and (D) same as (A) and (B), but using synaptic scaling and LTD. Graphs show the average mean overlap for 10 simulations. Standard deviations values were lower than 0.015 and are not shown for clarity.

The heteroassociative network forgets older patterns while learning new ones

325 The heteroassociative network model with the synaptic scaling rule replaces old pattern sequences with
326 newer ones. It is a required property for any model of the CA3, since just increasing the number of stored
327 patterns would saturate synaptic weights, causing all stored pattern sequences to be lost, an effect called
328 catastrophic interference (Amit, 1989). It is also in accordance with the theory of temporary storage of
329 memories in the hippocampus, which are latter transferred to the cortex as remote memories (Frankland
330 and Bontempi, 2005).

331   The number of patterns that can be retrieved from the network depends on the cue size, as more
332 complete cues permit the retrieval of patterns with weaker connections. Pattern sequences were stored
333 before the simulation and the sequences were retrieved in ascending order, with one sequence retrieval
334 per theta cycle. Figure 6A shows the retrieval of the last 3000 stored patterns for different cue sizes. The
335 last stored patterns are fully retrievable but, as we go backwards in the set of stored patterns, we report a
336 strong transition point where patterns are no longer retrievable. The transition point depends on the cue
337 size, with larger cues enabling the retrieval of more patterns.

₃₃₈ The minimum cue size required to successfully retrieve the sequences was 0.4, which is related to
₃₃₉ the configuration of excitatory weights between patterns. Although with larger weights retrieval would
₃₄₀ be possible with a smaller cue, these larger weights would also make the network more unstable. Since
₃₄₁ performing fine tuning of the network parameters and evaluating the impact of each are outside the scope
₃₄₂ of this work, we used a value that offered a good trade-off between stability and cue size. Nonetheless,
₃₄₃ the network stored between 200 and 1600 patterns, depending on the cue size. Although this value seems
₃₄₄ small, we used a network of only 10,000 neurons, when compared to 300,000 from the rat CA3 (Rolls,
₃₄₅ 2007) or 2.7 million from the human CA3 (Simic et al., 1997).

₃₄₆ Pattern retrieval depends weakly on the initial connectivity (Figure 6B). Although with smaller
₃₄₇ extents there would be fewer connections from a retrieved pattern j to neurons from pattern $j+1$, these
₃₄₈ connections are stronger. The transition point between retrievable and forgotten patterns is actually the
₃₄₉ same, even for an initial connectivity of 40%. Consequently, even if most neuronal pairs from patterns
₃₅₀ j and $j+1$ cannot create connections due to anatomical or physiological restrictions, the network can
₃₅₁ still operate. This also permits the network to use fewer connections for pattern sequence storage, saving
₃₅₂ scarce metabolic resources in the network.
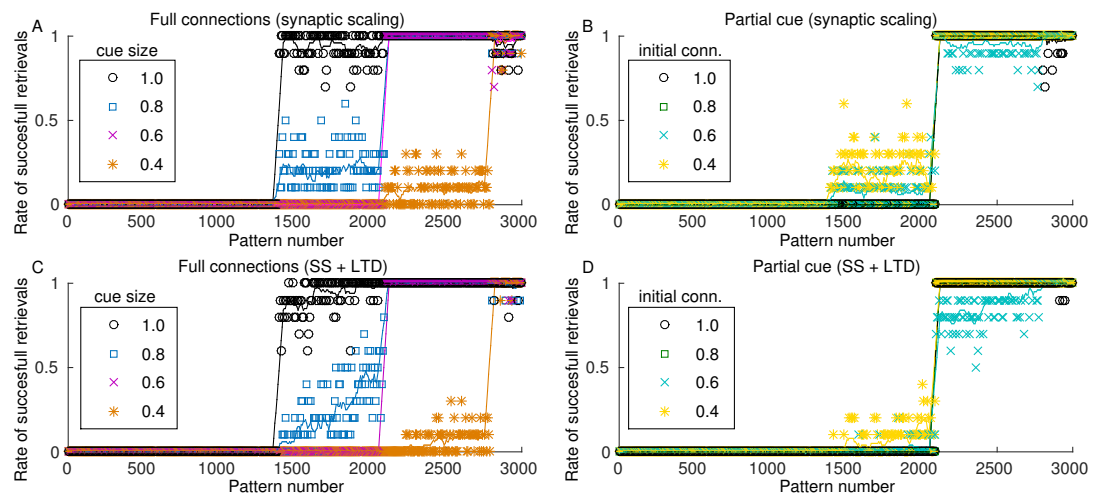
Figure 6. Rate of successful retrieval for the last 3000 stored patterns. Markers show the average retrieval rate for 10 simulations, lines represent the moving average of 10 retrieval rates and standard deviations are not shown for clarity. A retrieval is successful when at least four patterns of the sequence were retrieved with an overlap larger than 0.5. (A) Simulation with all connections, variable cue sizes, and synaptic scaling; (B) simulation with partial cues of 0.6, varying initial connectivity and synaptic scaling; (C) and (D) same as (A) and (B), but with the addition of LTD.

₃₅₃ When including LTD, the capacity of the network is slightly reduced (Figures 6C and 6D), but the
₃₅₄ effect is not as severe as Figure 2 would indicate. Also, although the capacity is reduced, the stored
₃₅₅ sequences are still retrieved correctly, as shown in Figures 5C and 5D. These results indicate that, for the
₃₅₆ proposed heteroassociative network, an additive synaptic scaling is sufficient for replacing old memories
₃₅₇ by new ones, as the addition of LTD does not cause perceptible changes in the retrieval capability of the
₃₅₈ network.

DISCUSSION

Coding of memories as pattern sequences

Episodic memories are characterized by their compositional properties (Cohen and Eichenbaum, 1993; Henke, 2010) and coding as pattern sequences allows different stimuli to be combined in a single composition. The relationship of theta sequences with spatial (Dragoi and Buzsáki, 2006) and episodic memories (Wang et al., 2014), the requirement of learning for their appearance (Feng et al., 2015) and their association with the representation of current goals (Wikenheiser and Redish, 2015) provide important evidence for heteroassociative coding in the hippocampus.

Retrieval of pattern sequences from small cues is possible with both auto- and heteroassociative networks. This retrieval occurs, for example, when mice need to retrieve stored spatial memories from the presentation of a small cue of the original memory (Nakazawa et al., 2002). With heteroassociative coding we can present a small cue from a single pattern in the sequence, and the rest of the sequence would be retrieved sequentially. With autoassociative networks, neurons from the stored pattern would be progressively activated by the self-reinforcing dynamics of the network until the complete pattern is retrieved. Also, both types of network permit the storage of memories of different sizes. With heteroassociative networks, larger complex memories could be coded in longer sequences, while simpler ones could use small sequences.

The hippocampus is considered to mediate some cognitive functions related to memory, such as associative representation of events, sequential organization, and relational networking (Eichenbaum, 2004). Interestingly, heteroassociative coding may address these issues. Associative representation of events can be accomplished by including events in the same pattern sequence, while a sequential organization arises naturally from pattern sequences. Relational networking considers that the same event may be part of multiple memories, which could be accomplished putting the same event in multiple pattern sequences. In this case, an extra mechanism would be required to disambiguate between these sequences during pattern sequence retrieval. For instance, Levy (1996) proposed that a set of context neurons in the CA3 could guide this disambiguation.

One possible problem with heteroassociative coding is that patterns cannot be held as persistent activity for an extended period of time. This activity persistence is normally considered a requirement for memory retrieval and can be accomplished with the usage of autoassociative connections. This persistent activity is normally found in the prefrontal cortex (Fuster and Alexander, 1971; Funahashi et al., 1989), but also in other areas, such as the entorhinal cortex (Egorov et al., 2002). These regions are largely connected to the hippocampus, so we can envision that the hippocampus is responsible for the associations, sequential organization, and relationships (Eichenbaum, 2004) between events. The retrieval of individual patterns in the sequence then causes the persistent activity associated to memories in the different cortex regions.

External noise, sequence completion and limited connectivity

The usage of heteroassociative connections in the recurrent CA3 connections is not a new subject. For instance, Yamaguchi (2003) studied a model that tries to explains theta phase precession using heteroassociative connections from CA3 to CA1 and in the CA3 recurrent connections. Other studies also used simple feedforward network models that associate pairs of patterns (Lytton, 1998; Miyata et al., 2013). But these studies did not evaluate the problem of sequence completion, noise tolerance and operation with limited connectivity.

Other existing heteroassociative models of the hippocampus use autoassociative connections in local

circuits of the CA3, to reduce the effects of noise and incomplete pattern retrieval. But the same set of CA3 recurrent connections cannot store both autoassociative and heteroassociative connections for a given pattern, since this would provide two possible routes for the network dynamics (Rolls and Kesner, 2006). A solution proposed in some theoretical studies is that the autoassociative and heteroassociative connections are present in different parts of the the CA3 (Samura et al., 2008) or in connections from the CA3 to other areas of the hippocampus (Lisman et al., 2005). In both models, the heteroassociative connections perform the storage and retrieval of sequences, leaving to the autoassociative connections the task of removing noise and performing pattern completion during sequence retrieval.

We showed that a heteroassociative network can operate correctly in the presence of external noise and limited synaptic connectivity, since errors in the retrieval of a single pattern do not propagate to the next pattern. Consequently, the presence of both autoassociative and heteroassociative connections in the same network may not be a requirement for preventing these propagating errors. One limitation of the current study is that we did not include noise in the neural responses or in synaptic plasticity. But since we showed that the network operates correctly with limited connectivity, it seems that even though this extra noise would degrade the network performance, it would not prevent the network from operating properly. But if the objective is to obtain extra reliability, a combination of autoassociative and heteroassociative connections may be helpful.

Synfire chains (Abeles, 1991) are defined as the sequential activation of pools or layers of neurons by feedforward connections. It is a broader concept that encompasses many types of network architectures. Although feedback connections may be considered in randomly connect networks (Abeles et al., 2004), the focus is on feedforward networks (Tetzlaff et al., 2002). Nevertheless, one may consider the retrieval of pattern sequences as a synfire chain.

Recent work has shown that 8.2% of the measured connections (12 out of 146) between pyramidal neurons in the CA3 of a rodent were part of reciprocal connection motifs (Guzman et al., 2016). Although reciprocal connections are normally considered as a signature of autoassociative networks, they are also present in heteroassociative networks. For instance, in our simulations of heteroassociative networks, more than 13% of connections were part of reciprocal motifs. Although the percentages of reciprocal motifs found in our simulations and in the rodent CA3 cannot be directly compared, the low number of motifs in the rodent CA3 is a evidence that heteroassociative coding may be used in the CA3 area.

System homeostasis and catastrophic forgetting

It is well known that an increase in the number of stored memories in an auto or heteroassociative network leads to a catastrophic interference effect (Amit, 1989), where the saturation of connection weights makes all stored memories unavailable. It is possible to prevent catastrophic forgetting in Hopfield networks (Hopfield, 1982) changing the learning rule to generate synaptic changes strong enough to guarantee the retrieval of newly stored patterns and bounding the maximum values of synaptic weights (Nadal et al., 1986). This type of memory was called palimpsest memory and permits the retrieval of the last stored memories, but it can store only a small number of patterns. We note that the replacement of old memories for new ones also receives the name of catastrophic forgetting in other types of networks (French, 1999) and is an important concern in procedural learning, since old learned skills may be useful and should not be indiscriminately erased by learning new ones. For the CA3 this effect is not be a real problem due to the temporary nature of memory storage.

Enabling the storage of new memories requires the removal of old memories that are never recalled. According to the remote memories hypothesis, memories are temporarily stored in the hippocampus to be latter coded as remote memories in the cortex (Frankland and Bontempi, 2005). Some form of

unlearning is required, and there are several proposals on how this could be performed, specially during sleep. The two main hypotheses for this unlearning are the active system consolidation (Diekelmann and Born, 2010), where memories in the hippocampus are replayed during sleep and transferred over to the neocortex, causing the potentiation of synapses, and the synaptic homeostasis hypothesis (Tononi and Cirelli, 2014), which states that synapse weights are strengthened by learning during wake periods and decreased during sleeping.

Our model is compatible with both hypotheses. With the synaptic homeostasis hypothesis (Tononi and Cirelli, 2014), our learning of new associations would represent the wake periods and the additive synaptic scaling rule would do the homeostatic process of reducing synaptic weights. We use synaptic scaling to maintain the total sum of weights constant for each neuron, by decreasing all synaptic weights by the same amount. Memories are forgotten as their respective synapses have their weights reduced and weaker synapses are eventually eliminated as their weights vanish.

When considering the active system consolidation (Diekelmann and Born, 2010), stored memories would be recalled and transferred over to the neocortex during sleep. Different mechanisms could be used for maintaining the system homeostasis, including the use of STDP rules (Caporale and Dan, 2008) or activity-dependent synaptic scaling (Abbott and Nelson, 2000). In our model, we used a synaptic scaling based on the sum of the weights, since we did not simulated the learning process itself. Synaptic scaling rules are normally based on the mean activation of the neuron, but over long periods this activity should be proportional to the sum of the input weights, resulting in similar synaptic changes.

The CA3 subregion of the hippocampus seems to be involved in several tasks such as storing temporal sequences and associations between parts of memories, while maintaining a homeostasis over the lifetime of individuals. Moreover, it also needs to operate in the presence of noise, incomplete cues, and limited connectivity. With this study we depicted important properties of the CA3 that can be accomplished using a simple heteroassociative rule combined with competitive synaptic scaling. We believe that both auto- and heteroassociative rules have their roles in memory processes and simulation studies like this allows us to better understand the effects of different learning rules over the operation of these networks.

ACKNOWLEDGMENTS

The authors would like to thank Universidade Federal do ABC (UFABC) for providing the computational infrastructure for the simulations and the members of the Timing and Cognition Lab (`http://neuro.ufabc.edu.br/timing/`), at UFABC, for their suggestions during project development. We also thank the reviewers of the paper for providing insightful contributions.

REFERENCES

Aaron, G. B. and Dichter, M. a. (2001). Excitatory synapses from CA3 pyramidal cells onto neighboring pyramidal cells differ from those onto inhibitory interneurons. *Synapse (New York, N.Y.)*, 42(4):199–202.

Abbott, L. F. and Nelson, S. B. (2000). Synaptic plasticity: taming the beast. *Nature Neuroscience*, 3:1178–83.

Abeles, M. (1991). *Corticonics: Neuronal Circuits of the Cerebral Cortex*. Cambridge University Press, Cambridge, England, 1st edition.

Abeles, M., Hayon, G., and Lehmann, D. (2004). Modeling compositionality by dynamic binding of synfire chains. *Journal of Computational Neuroscience*, 17(2):179–201.

Amaral, D. G. and Witter, M. (1989). The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience*, 31(3):571–591.

Amit, D. J. (1989). *Modeling Brain Function: The World of Attractor Neural Networks*. Cambridge University Press, New York, NY, USA.

Andersen, P., Morris, R., Amaral, D., Bliss, T., and O'Keefe, J. (2007). *The Hippocampus Book*. Oxford University Press, New York, NY.

Born, J., Rasch, B., and Gais, S. (2006). Sleep to remenber. *The Neuroscientist*, 12(5):410–424.

Bragin, A., Jando, G., Nadasdy, Z., Hetke, J., Wise, K., and Buzsáki, G. (1995). Gamma (40-100Hz) Oscillation in the Hippocampus of the Behaving Rat. *Journal of Neuroscience*, 15:47–60.

Caporale, N. and Dan, Y. (2008). Spike timing-dependent plasticity: a Hebbian learning rule. *Annual Review of Neuroscience*, 31:25–46.

Chistiakova, M. and Volgushev, M. (2009). Heterosynaptic plasticity in the neocortex. *Experimental Brain Research*, 199(3-4):377–390.

Chrobak, J. J. and Buzsáki, G. (1998). Gamma oscillations in the entorhinal cortex of the freely behaving rat. *Journal of Neuroscience*, 18(1):388–98.

Cohen, N. J. and Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System*. MIT Press, Cambridge, MA.

Colgin, L. L. (2016). Rhythms of the hippocampal network. *Nature Reviews Neuroscience*, 17(4):239–249.

Cutsuridis, V. and Wennekers, T. (2009). Hippocampus, microcircuits and associative memory. *Neural Networks*, 22(8):1120–8.

de Almeida, L., Idiart, M., and Lisman, J. E. (2007). Memory retrieval time and memory capacity of the CA3 network: role of gamma frequency oscillations. *Learning & Memory*, 14(11):795–806.

Diego, S., Sodroski, J., Hausmann, E. H., Ozel, M., Pauli, G., Koch, M. A., Saphire, E. O., Parren, P. W. H. I., Muir, T. W., Kent, S. B., Minor, W., Main, P., Macarthur, M. W., Moss, D. S., Bacon, D. J., Menendez, A., Olson, A. J., Spehner, J. C., Wormald, M., Vos, A. D., Poignard, P., and Wang, M. (2001). Enforcement of Temporal Fidelity in Pyramidal Cells by Somatic Feed-Forward Inhibition. *Science*, 293(5532):1159–1163.

Diekelmann, S. and Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, 11(2):114–26.

Dragoi, G. and Buzsáki, G. (2006). Temporal encoding of place sequences by hippocampal cell assemblies. *Neuron*, 50(1):145–57.

Egorov, A. V., Hamam, B. N., Fransen, E., Hasselmo, M. E., and Alonso, A. A. (2002). Graded persistent activity in entorhinal cortex neurons. *Nature*, 420(6912):173–178.

Eichenbaum, H. (2004). Hippocampus: Cognitive processes and neural representations that underlie declarative memory. *Neuron*, 44(1):109–120.

Fell, J. and Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature Reviews Neuroscience*, 12(2):105–18.

Feng, T., Silva, D., and Foster, D. J. (2015). Dissociation between the experience-dependent development of hippocampal theta sequences and single-trial phase precession. *The Journal of Neuroscience*, 35(12):4890–902.

Foster, D. J. and Wilson, M. A. (2007). Hippocampal theta sequences. *Hippocampus*, 17(11):1093–1099.

Frankland, P. W. and Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, 6(2):119–30.

French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences*,

531 3(4):128–135.

532 Freund, T. and Buzsáki, G. (1998). Interneurons of the hippocampus. *Hippocampus*, 6(4):347–470.

533 Funahashi, S., Bruce, C., and Goldman-Rakic, P. (1989). Mnemonic Coding of Visual Space in the

534 Monkey's Dorsolateral Prefrontal Cortex. *Journal of Neurophysiology*, 6(2):331–349.

535 Fuster, J. M. and Alexander, G. E. (1971). Neuron Activity Related to Short-Term Memory. *Science*,

536 173(3997):652–654.

537 Gloveli, T., Dugladze, T., Rotstein, H. G., Traub, R. D., Monyer, H., Heinemann, U., Whittington,

538 M. a., and Kopell, N. J. (2005). Orthogonal arrangement of rhythm-generating microcircuits in the

539 hippocampus. *Proceedings of the National Academy of Sciences*, 102(37):13295–300.

540 Guzman, S. J., Schlögl, A., Frotscher, M., and Jonas, P. (2016). Synaptic mechanisms of pattern

541 completion in the hippocampal CA3 network. *Science*, 353(6304):109–120.

542 Hájos, N., Pálhalmi, J., Mann, E. O., Németh, B., Paulsen, O., and Freund, T. F. (2004). Spike timing of

543 distinct types of GABAergic interneuron during hippocampal gamma oscillations in vitro. *Journal of*

544 *Neuroscience*, 24(41):9127–37.

545 Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness.

546 *Nature Reviews Neuroscience*, 11(7):523–32.

547 Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational

548 abilities. *Proceedings of the National Academy of Sciences*, 79(April):2554–2558.

549 Jensen, O. and Lisman, J. E. (1996). Hippocampal CA3 region predicts memory sequences: accounting

550 for the phase precession of place cells. *Learning & Memory*, 3(2-3):279–87.

551 Levy, W. B. (1996). A sequence predicting CA3 is a flexible associator that learns and uses context to

552 solve hippocampal-like tasks. *Hippocampus*, 6:579–590.

553 Lisman, J. (2005). The theta/gamma discrete phase code occuring during the hippocampal phase

554 precession may be a more general brain coding scheme. *Hippocampus*, 15(7):913–22.

555 Lisman, J. and Idiart, M. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science*,

556 267(5203):1512–5.

557 Lisman, J. E. and Jensen, O. (2013). The θ-γ neural code. *Neuron*, 77(6):1002–16.

558 Lisman, J. E., Talamini, L. M., and Raffone, A. (2005). Recall of memory sequences by interaction of the

559 dentate and CA3: a revised model of the phase precession. *Neural Networks*, 18(9):1191–201.

560 Lytton, W. W. (1998). Adapting a feedforward heteroassociative network to Hodgkin-Huxley dynamics.

561 *Journal of Computational Neuroscience*, 5(4):353–64.

562 Meeks, J. P. and Mennerick, S. (2007). Action potential initiation and propagation in CA3 pyramidal

563 axons. *Journal of Neurophysiology*, 97(5):3460–72.

564 Miles, R. and Wong, R. (1986). Excitatory synaptic interactions between CA3 neurones in the guinea-pig

565 hippocampus. *Journal of Physiology*, 373:397–418.

566 Miyata, R., Ota, K., and Aonishi, T. (2013). Optimal design for hetero-associative memory: hippocampal

567 CA1 phase response curve and spike-timing-dependent plasticity. *PLoS One*, 8(10):e77395.

568 Nadal, J. P., Toulouse, G., Changeux, J. P., and Dehaene, S. (1986). Networks of Formal Neurons and

569 Memory Palimpsests. *Europhys. Lett*, 1(10):535–542.

570 Nakazawa, K., Quirk, M. C., Chitwood, R., atanabe, M., Yeckel, M. F., Sun, L., Kato, A., Carr, C.,

571 Johnston, D., Wilson, M., and Tonegawa, S. (2002). Requirement for hippocampal CA3 NMDA

572 receptors in associative memory recall. *Science*, 297(5579):211–8.

573 O'Keefe, J. and Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG

574 theta rhythm. *Hippocampus*, 3(3):317–30.

Pearce, A. (1993). Physiological Evidence for Two Distinct GABA* Responses in Rat Hippocampus. *Cell*, 10:189–200.

Rolls, E. T. (2007). An attractor network in the hippocampus: theory and neurophysiology. *Learning & Memory*, 14(11):714–31.

Rolls, E. T. (2010). A computational theory of episodic memory formation in the hippocampus. *Behavioural Brain Research*, 215(2):180–96.

Rolls, E. T. and Kesner, R. P. (2006). A computational theory of hippocampal function, and empirical tests of the theory. *Progress in Neurobiology*, 79(1):1–48.

Rolls, E. T., Treves, A., Foster, D., and Perez-Vicente, C. (1997). Simulation studies of the CA3 hippocampal subfield modelled as an attractor neural network. *Neural Networks*, 10(9):1559–1569.

Samura, T., Hattori, M., and Ishizaki, S. (2008). Sequence disambiguation and pattern completion by cooperation between autoassociative and heteroassociative memories of functionally divided hippocampal CA3. *Neurocomputing*, 71(16-18):3176–3183.

Simic, G., Kostovic, I., Winblad, B., and Bogdanovic, N. (1997). Volume and number of neurons of the human hippocampal formation in normal aging and Alzheimer's disease. *Journal of Comparative Neurology*, 494:482–494.

Sompolinsky, H. and Kanter, I. (1986). Temporal association in asymmetric neural networks. *Phys Rev Letters*, 57(22).

Spruston, N., Jonas, P., and Sakmann, B. (1995). Dendritic glutamate receptor channels in rat hippocampal ca3 and ca1 pyramidal neurons. *Journal of Physiology (Lond)*, 482.2:325–352.

Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological review*, 99(2):195–231.

Tetzlaff, T., Geisel, T., and Diesmann, M. (2002). The ground state of cortical feed-forward networks. *Neurocomputing*, 44-46:673–678.

Tononi, G. and Cirelli, C. (2014). Sleep and the Price of Plasticity: From Synaptic and Cellular Homeostasis to Memory Consolidation and Integration. *Neuron*, 81(1):12–34.

Treves, a. and Rolls, E. T. (1992). Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus*, 2(2):189–99.

Wang, Y., Romani, S., Lustig, B., Leonardo, A., and Pastalkova, E. (2014). Theta sequences are essential for internally generated hippocampal firing fields. *Nature Neuroscience*, 18(2):282–288.

Wikenheiser, A. M. and Redish, A. D. (2015). Hippocampal theta sequences reflect current goals. *Nature Neuroscience*, 18(2):1–8.

Yamaguchi, Y. (2003). A theory of hippocampal memory based on theta phase precession. *Biological Cybernetics*, 89(1):1–9.