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

Recent research suggest that the CA3 sub-region of the hippocampus operates both as a heteroassociative network, due to the ability to store and retrieve sequences of patterns, and as an autoassociative one, due to its ability to complete partial clues, tolerate noise, and store associations between different parts of episodic memories in flexible ways. Recent evidence regarding hippocampal theta sequences reinforces the hypothesis of coding spatial memories as pattern sequences in the hippocampus. We investigated whether a heteroassociative coding scheme combined with competitive synaptic scaling for homeostasis, would suffice to explain the features that make the hippocampus a reliable network for memory: a) the ability to retrieve stored pattern sequences after the presentation of a partial cue from one of its patterns, in the presence of external noise and incomplete connectivity among neurons, and b) the renewal of the set of stored sequences as new memories are learned, permitting the storage of new sequences while forgetting older ones. We show that these hypotheses hold for a biologically plausible model of the CA3 subregion containing 10,000 integrate-and-fire neurons, with sequence retrieval organized by gamma cycles nested on theta cycles. ... where is the rest of abstract?

theory? (that's not a hypothesis)

this confused me into thinking your model was also autoassociative

? what does renewal mean?

dangerous phrase!

Keywords: Theta sequences, Hippocampus, Computational Neuroscience, Heteroassociative coding

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), the memory coding scheme 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, which enables the association between different memories in flexible ways ~~and the retrieval~~ of stored memories from the presentation of partial cues. These characteristics of autoassociative networks makes 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).

No, between same mems.

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 inside 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, representing neuronal ensembles. A pattern sequence could be retrieved in a theta cycle, with one pattern per gamma cycle nested in the theta cycle. To combine the properties of auto- and heteroassociative networks, models with additional autoassociative connections for each pattern in the sequence were proposed~~, despite the extra required connections~~. Feedback connections from CA3 to Dentate Gyrus (Lisman et al., 2005) or between different CA3 areas (Samura et al., 2008) would store sequences of memories in heteroassociative connections, and recurrent connections within single CA3 areas would work as autoassociative connections.

"at which"?

"represented by" ??

d. g.

10 years!? data, not just concept!

Recently, the concept of theta sequences (Foster and Wilson, 2007) was proposed, representing sequence of neuronal ensembles which are sequentially activated bounded by theta cycles. Differently from 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 code these memories as sequential activation of neuronal ensembles.

"In contrast to..."

What were first two?

temporarily?

A third feature of the hippocampus is that memories appear to be temporally 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 receive the name of catastrophic forgetting in other types of networks (French, 1999), this effect is acceptable in the CA3 due to the temporary nature of memory storage. Proposals for enabling this memory renewal include the use of STDP rules (Caporale and Dan, 2008) and memory recall and consolidation during sleep (Born et al., 2006).

To whom? This is presuming teleological knowledge of CA3 role.

??

Here, we investigated whether a single coding scheme would suffice to explain the features that make the hippocampus a reliable network for memory: the ability to retrieve stored sequences after presentation of a partial cue from one of its patterns, and to ~~reliably~~ retrieve them even in the presence of external noise and incomplete connectivity among neurons. We hypothesize that a neural network, with connections determined exclusively by a heteroassociative competitive learning rule, can reliably a) store and retrieve episodic memories coded as pattern sequences, and b) renew the set of stored sequences as new memories are learned. We also hypothesize that a heteroassociative learning rule, combined with competitive synaptic scaling for homeostasis, would permit the storage of new sequences while forgetting older ones. We show that these hypotheses hold for a biologically plausible model of the CA3 subregion containing 10,000 neurons, with sequence retrieval organized by gamma cycles nested on theta cycles.

What does "renewal" mean? Overwriting?

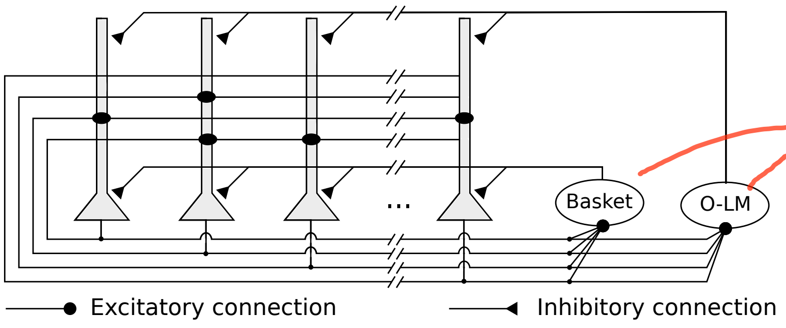
More detail is needed to justify.

But I thought old sequences are 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~~-based~~ models of AMPA synaptic channels and weights defined by a heteroassociative learning rule. The network is shown in Figure 1.



These look like single global interneurons.

Figure 1. Network architecture, with the recurrent connections among pyramidal neurons and connections representing inhibitory basket and O-LM neuronal populations.

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 included a periodic inhibition by O-LM cells (Freund and Buzsáki, 1998), modeled with a spike generator representing the intrinsic firing rate in the theta range (5 Hz) of these

? Poisson? Periodic?

~8 Hz during exploration!

there is no other theta generation, right?

cells when depolarized (Gloveli et al., 2005). It sends connections to the slow GABA channel of every pyramidal neuron, modulating the generation of theta rhythm in the network.

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

pyramidal

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

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

You mean "spike frequency adaptation"?

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

$t_{spk} - t$, right? Otherwise this explodes!

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

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

where I_{ext} represents the external input, I_{ampa} the excitatory AMPA synapses, I_{gaba} the inhibitory fast GABA synapses, and I_{gabaS} the slow GABA synapses. We used the Euler method with an integration step of 0.1 ms.

justification? why not RK4?

The synaptic model

We modeled the synaptic channels using a dual exponential model, given by:

To account for $\Delta t_s < 0$: (I assume it was correct in the code)

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

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

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

We used the $\tau_1 = 2\,\mathrm{ms}$, $\tau_2 = 8\,\mathrm{ms}$ and $A_k = 3200\,\mathrm{pA}$ for AMPA channel of recurrent connections (Spruston et al., 1995), $\tau_1 = \tau_2 = 2\,\mathrm{ms}$ and $A_k = 3200\,\mathrm{pA}$ for the external input channels, $\tau_1 = \tau_2 = 5\,\mathrm{ms}$ and $A_k = 540\,\mathrm{pA}$ for the fast GABA channels (Pearce, 1993), and $\tau_1 = 7\,\mathrm{ms}$, $\tau_2 = 57\,\mathrm{ms}$ and $A_k = 30\,\mathrm{pA}$ for the slow GABA channels (Pearce, 1993).

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

Is spatial pair-wise delay necessary to results? Why not just use this average?

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

External noise was modeled with a Poisson process ~~random number generators~~ with 1 Hz rate per neuron, representing inputs to the neuron not related to the retrieval of sequences. We used weights that were large enough to cause the network neurons to fire spontaneously at the rate of 0.75 Hz, which is similar to the rate 0.80 Hz produced during pattern retrieval in the simulations.

Your definition is wrong. Sparsity is fraction of inactive cells.

With or without replacement?

Patterns should be orthogonal.

Pattern sequences

We created the patterns by generating random permutations with the neuron indexes and getting the first r values of the permutation, where r is equal to the number of neurons times the sparsity for the simulation. We define sparsity as the number of neurons active in a pattern divided by the total number of neurons. ~~For a sparsity of 0.1, 10% of all neurons are active on the pattern.~~ For heteroassociative networks, there are estimates by de Almeida et al. (2007) of sparsity of 0.003. For our simulated network with 10,000 neurons this would result in only 30 neurons per pattern, which would be insufficient for pattern completion and noise tolerance. Although using more neurons per pattern decreases the storage capacity of the network, we decided to use a sparsity of 0.01, resulting in 100 neurons per pattern.

(redundant)

Why not simulate 33,000 neurons? This discrepancy should be explained.

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

Are you assuming Lisman models are correct?

Learning rule

The model uses a heteroassociative learning rule (Sompolinsky and Kanter, 1986) based on asymmetric STDP (Dan and Poo, 2004) for pattern sequence storage. We simulated an offline learning procedure, performed before the simulation, by considering that the patterns from each sequence were presented in sequence, with a separation of about 15 ms, corresponding to one cycle of the gamma rhythm. Consequently, connections weights from the pattern μ to $\mu+1$ are increased, while connections from μ to $\mu-1$ are decreased. The learning rule for each pattern association can be summarized as:

This model does not do STDP.

Does this not presuppose learning of the sequence? I.e. this is not the timescale of experience.

unclear. Do you run training both forward (+, LTP) and backward (−, LTD)?

$$\Delta w_{ij} = \pm\gamma^{\pm} p_i^{\mu} p_j^{mod(\mu\pm1,k)}$$

where w_{ij} represents the weight of the connection from neuron i to neuron j, $\gamma^{\pm}$ the LTP and LTD rate, p_i^{μ} the neuron i from pattern μ and $p_j^{mod(\mu\pm1,k)}$ the neuron j from the next/previous pattern in a sequence of k patterns, $\mu = 0,\ldots,k-1$. The function $mod(x,k)$ represents the modulo operation, with $mod(-1,k)$ defined as $k-1$. This produces an association between the first and last patterns of the sequence. We used $\gamma^{+} = 1$ for LTP and $\gamma^{-} = \{0,1\}$ for simulations without and with LTD.

The network is initialized with connections between every pyramidal neuron pair, with random weights drawn from an uniform distribution between 0 and defined maximum value, simulating a highly connected network from a developing nervous system. ? The learning rule is then sequentially applied for each pattern sequence to be stored. We included synaptic scaling (Abbott and Nelson, 2000) as a~~n~~ heterosynaptic long-term depression (hLTD) mechanism to model a competition between the synapses of a single neuron. After learning a set of 100 pattern sequences, we apply an additive synaptic scaling rule to all neurons. We decrease all synapses from each neuron by a fixed amount, defined as the ratio between the total weight increase since the last scaling procedure and the number of connections. Synapses that become negative are set to zero and the difference is subtracted from the remaining synapses. ✓

Although the usage of a standalone STDP rule is known to generate groups of neurons connected by strong synapses and that fire together (Buonomano, 2005), Fiete et al. (2010) showed that an as~~s~~ym m etric (2 m's) STDP learning rule, combined with heterosynaptic competition within single neurons, allow ed networks to learn long sequences of patterns.

?? In biological networks, each neuron can connect to a limited number of neurons, either by anatomical or physiological reasons. We simulate this constraint using different levels of connectivity ~~extent,~~ where we randomly define for each neuron the set of neurons to which it can connect. For ~~a~~ connectivity ~~extent~~ of 0.6, each neuron initially connects to 60% randomly selected neurons in the network. Connections to neurons outside this initial set cannot be created during learning, but connections can be eliminated due to LTD or synaptic scaling. We also change the connection weights proportionally to the ratio of removed connections, to maintain the excitatory drive with different connectivity ~~extents comparable~~.

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 the entorhinal cortex (EC) provides the input to the CA3, 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.

Actually it's mostly DG

Because of refractory period?

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:

But that would be 15 ms?

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

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 has value 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 measure only the neurons which are part of the pattern. This can be applied because the number of active neurons is always comparable to the number of neurons per pattern. We also compute the second highest overlap, so that we can check if a second pattern is retrieved at the same time.

This is unnecessary because your dot-product* will zero them out anyway because $p_i^u = 0$ for nonpattern neurons. Would prefer a full Hamming distance.

RESULTS

The number of retrievable patterns and connections are stable during the network lifetime

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 including STDP-induced LTD (SS-LTD), with $\gamma^- = -1$, showing that synaptic scaling is sufficient to control the number of connections in the network.

You don't use STDP though?!

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, the network would be initialized close to a pattern and simulated to check if it converges towards the pattern. But this would be infeasible, 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, sparsity level, connectivity ~~extend~~ and LTD rule.

LTD ≠ STDP

ok, but this would be nec. for autoassoc. retrieval.

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 crosstalk interference effect. After reaching an equilibrium point, the number of retrievable patterns remains constant, with new patterns replacing older ones. Using a sparsity 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

Your data (2b) doesn't show this (necessarily)

where is this defined? what is "crosstalk"?

Again, what STDP? The model has LTD, not STDP!

Not opposite, there is one trend, from 0.005→0.01→0.02 w/ decreasing retrievable patterns.

patterns (Figures 2a-b). Combined with STDP-induced LTD, which caused a larger number of weight decreases, the number of retrievable patterns became almost zero. The opposite trend occurred with sparsity 0.005, showing the benefits of using less neurons per pattern.

Including STDP-induced 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 STDP-induced LTD, we compensate for this by reducing the effect of synaptic scaling, so that the total weights per neuron with or without STDP-induced LTD will be the same. It appears that using STDP-induced 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.

Unclear

Full connectivity is not possible in the real CA3 as, due to anatomical and physiological constraints, neurons can form new connections only with part of the neurons in the network. A neuronal connectivity ~~extent~~ of 0.6 means that a neuron can connect to 60% of the neurons. Using ~~smaller~~ lower connectivity ~~extents~~, the number of connections is reduced, but the number of retrievable patterns is also reduced (Figures 2c-d). But it is interesting that using a connectivity ~~extent~~ of 0.4 and synaptic scaling (SS) alone, there were about 900 retrievable patterns, which is more than the 600 patterns obtained when using STDP-induced LTD (SS-LTD) with full connectivity ~~extent~~. Also, although 900 patterns may be too ~~little~~ few, this value would be much higher in the larger networks from biological systems. These results indicate that purely heteroassociative networks with a connectivity ~~extent~~ of 0.4 could be biologically viable.

Could you present estimates based on your results?

Not a strong conclusion based on your small sample of connectivity values.

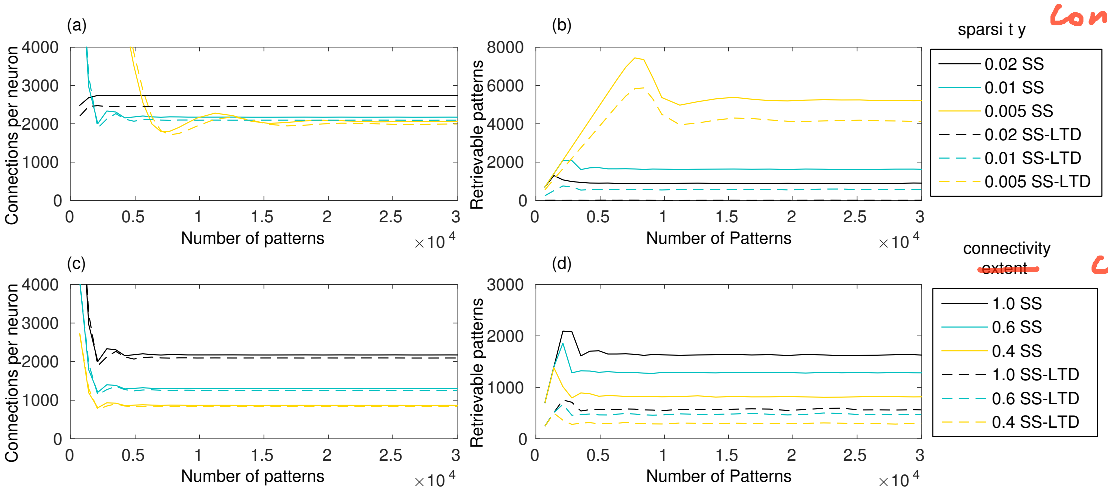


Figure 2. (a) Number of connections in a network with 10,000 neurons after learning a set of pattern sequences when applying only the synaptic scaling (SS) rule, for a sparsity of 0.02, 0.01, and 0.005, and with synaptic scaling and STDP-induced LTD (SS-LTD); (b) Number of retrievable patterns for the same scenarios from (a); (c-d) Same as (a-b), but comparing neuronal connectivity extents of 1.0, 0.6, and 0.4.

Increasing the initial total weight per neuron increases the number of connections, caused by ~~a smaller~~ 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 bifurcation 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 STDP-based LTD, but in the latter the number of retrievable patterns was larger.

No, that's not a bifurcation. It's just a maximum due to a trade-off. You could call it a "transition" point.

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 bifurcation point. Including more neurons or using a smaller sparsity, 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 initial connect~~ion~~ vity ~~extent~~, and on the sparsity of the patterns (Figures 3a,c). With larger initial weights, the sum of synaptic weights toward each neuron will be larger and, consequently, ~~less~~ fewer connections will be removed by the effect of synaptic scaling and STDP-induced LTD.

It is also interesting that although a smaller connectivity extent 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 lower signal-to-noise ratio in the system, as there will be proportionally ~~less~~ coding connections. But with a connectivity extent of 60%, it is possible to use a larger initial weight of 6.5, while maintaining the total connections to 20% of the neurons, which resulted in about 1800 retrievable patterns, an actual improvement over the full connectivity extent. With an extent of 40%, the number of retrievable patterns started to decrease with more than 1200 connections per neuron.

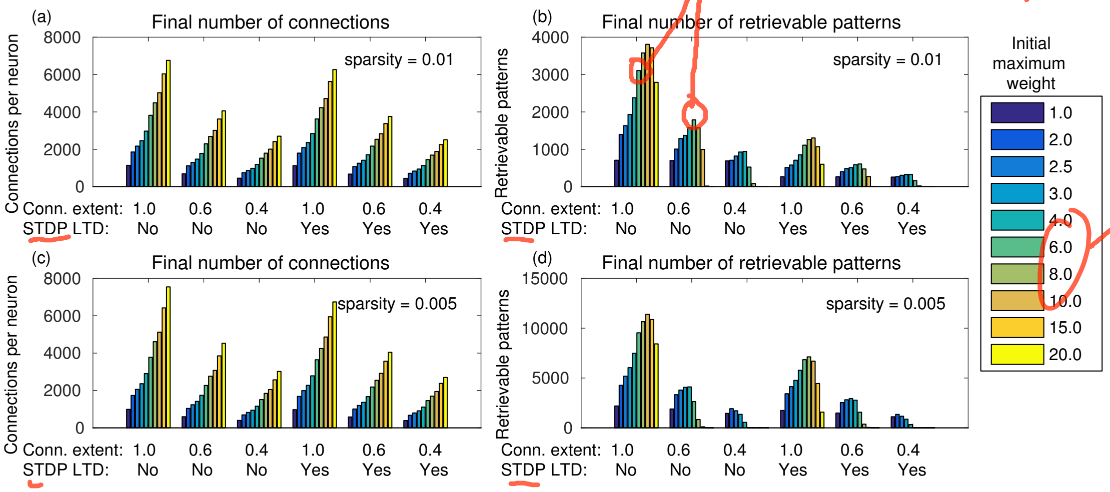
fewer

Using a sparsity of 0.005 caused small changes in the final number of connections, but the number of retrievable patterns increased significantly (Figures 3c-d), which would be more typical for biological larger networks with ~~a smaller~~ sparsity. The exception is for the connectivity extent of 0.4, due to the small number of neurons per pattern in the simulated network.

lower

How is 0.6 an improvement?

Do you mean "conns. per neuron"? Why would pattern size change?



Why not show 6.5 as you describe above?

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 sparsity of 0.01, connectivity extents of 1.0, 0.6 and 0.4 with and without STDP-induced LTD; (b) Number of retrievable patterns for the same scenarios from (a); (c-d) Same as (a-b), but with sparsity 0.05.

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 sparsity 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 connectivity extents 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

Nice!

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 a gradual recovery of the spike rate when the inhibition fades. Also, no pattern is retrieved during the period, 4a despite the firing of near 100 neurons at some points, 4b showing that the firing was random, which is also illustrated in the raster plot (Figure 4d). When a random external input is presented at instants 4200ms and 4400ms, the target neurons are activated, but since they do not represent a stored pattern, this extra activity disappears.

Figure refs

Presenting a pattern cue caused the retrieval of the complete pattern sequence just after its presentation, at 4200 ms and 4400 ms in Figures 4e-h. Each pattern is retrieved during a small time window, What is "connectivity extent" of these simulations?

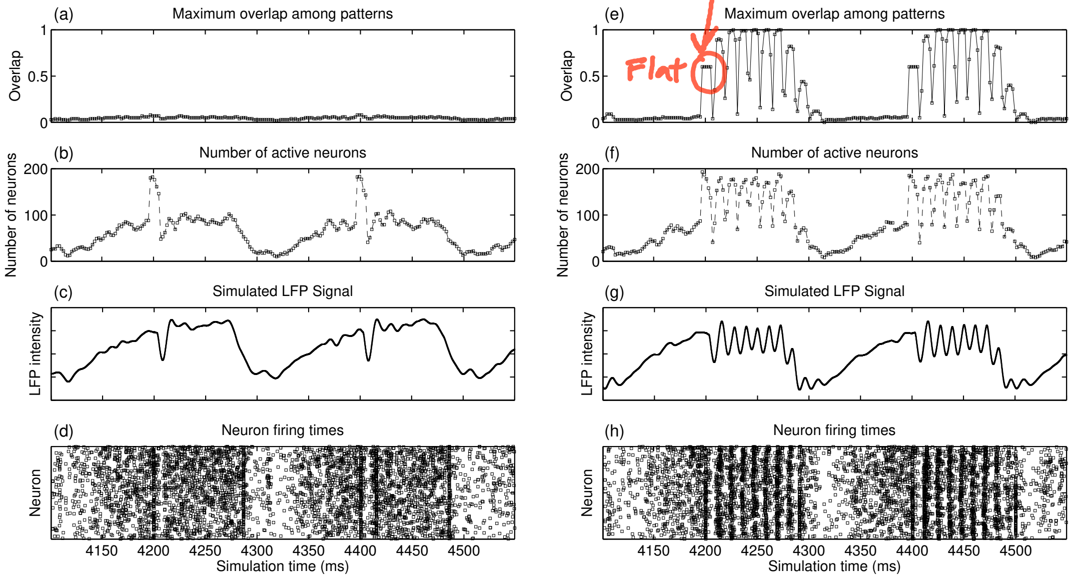


Figure 4. Operation of two heteroassociative networks with 10^4 pyramidal neurons and 10^4 stored patterns with 100 neurons per pattern and coded into sets of 7 patterns. The left-hand side graphs show 450 ms of the network operation with presentation of random inputs at instants 4200 ms and 4400 ms. (a) Maximum value of the overlap among all stored patterns; (b) number of active neurons at each time; (c) simulated LFP signal from the network; (d) raster graph showing the times 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.

corresponding to a gamma cycle inside the theta cycle, as occurs in the theory proposed by Lisman (2005). As the complete sequence has seven patterns, the complete ~~pattern~~ sequence? is retrieved during a single theta cycle. We can also see that although we present only 60% of the first pattern of the set (cue size of 0.6), the network successfully retrieved and completed the subsequent patterns.* Finally, the raster plot of neuronal spikes shows the clustering of neuronal firings during each pattern retrieval, with periods of inactivity between gamma cycles.

* But why not the complete pattern cue? No autoassoc. retrieval? (I was still confused abt. model at this point)

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

In the simulations from Figures 4e-h, we presented the pattern cues at the beginning of the theta cycles. Presenting the cues at other phases would cause only part of the sequence to be retrieved, since the periodical inhibition from O-LM cells would cause the pyramidal cells to stop firing before the sequence was finished. Since phase synchronization may ~~seems to~~ be required for memory retrieval in the hippocampus (Fell and Axmacher, 2011), it seems reasonable to provide the inputs at specific theta phases.

Sequence completion ~~Completion of partial cues~~ during retrieval occurs due to the heteroassociative connections from pattern j to next pattern $j+1$. Since many neurons from pattern j must fire to activate pattern $j+1$, randomly firing neurons do not cause random pattern retrievals. Sequence ~~Pattern~~ completion also works with connectivity extent of 0.6, as shown in Figure 5a. Also, the fast inhibition guarantees that only a single pattern is retrieved at a time. After evaluating the overlap with all other stored patterns, the second highest overlap was always lower than 0.1 (Figure 5b). Finally, using STDP-induced LTD did not change the overlaps significantly (Figures 5c-d).

(Because your partial cue is not completed)

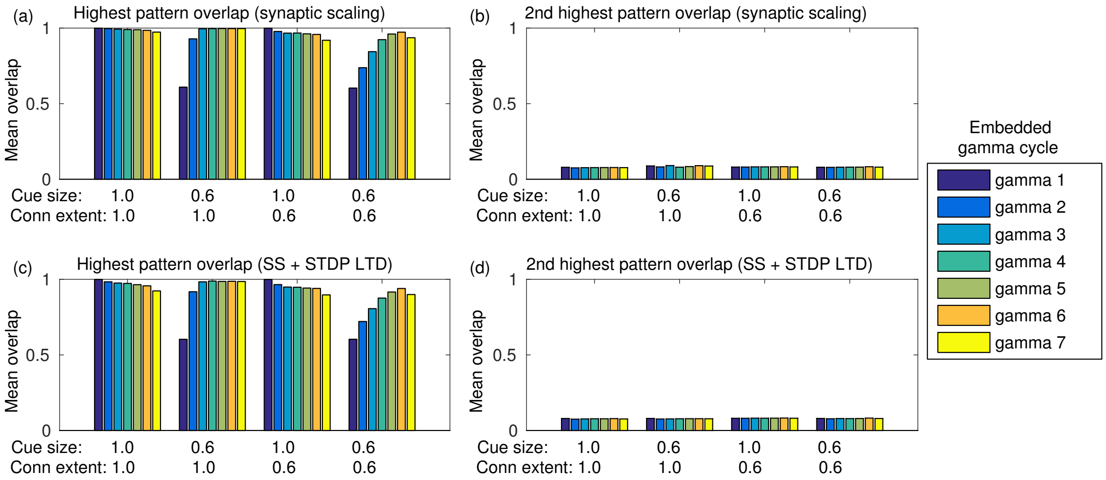


Figure 5. Mean overlap during pattern retrieval within each gamma cycle inside a theta cycle. Gamma 1 refers to the first gamma cycle in each theta cycle, gamma 2 to the second, and so on. (a) The highest overlap among patterns for different cue sizes and connectivity extents and using synaptic scaling 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 STDP-induced LTD.

The heteroassociative network forgets older patterns while learning new ones

The heteroassociative network model with the synaptic scaling rule replaces old pattern sequences with newer ones, working as a palimpsest memory (Nadal et al., 1986). It is a required property for any model of the CA3, since just increasing the number of stored patterns would saturate synaptic weights, causing all stored pattern sequences to be lost, an effect called catastrophic interference (Amit, 1989). It is also in accordance with the theory of temporary storage of memories in the hippocampus, which are later transferred to the cortex as remote memories (Frankland and Bontempi, 2005).

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

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

Pattern retrieval depends weakly on the connectivity extent (Figure 6b). Although with lower extents there would be ~~less~~ fewer connections from a retrieved pattern j to neurons from pattern $j+1$, these connections are stronger. The transition point between retrievable and forgotten patterns is actually the same, even for a connectivity extent of 40%. Consequently, even if most neuronal pairs from patterns j and $j+1$ cannot create connections due to anatomical or physiological restrictions, the network can still operate. This also permits the network to use ~~less~~ fewer connections for pattern sequence storage, saving scarce metabolic resources in the network.

When including STDP-induced LTD, the capacity of the network is slightly reduced (Figures 6c and 6d), but the effect is not as severe as Figure 2 would indicate. Also, although the capacity is reduced, the stored sequences are still retrieved correctly, as shown in Figures 5c and 5d. These results indicate that, for the proposed heteroassociative network, an additive synaptic scaling is sufficient for replacing old

Confusing to use same color scheme for diff. variables

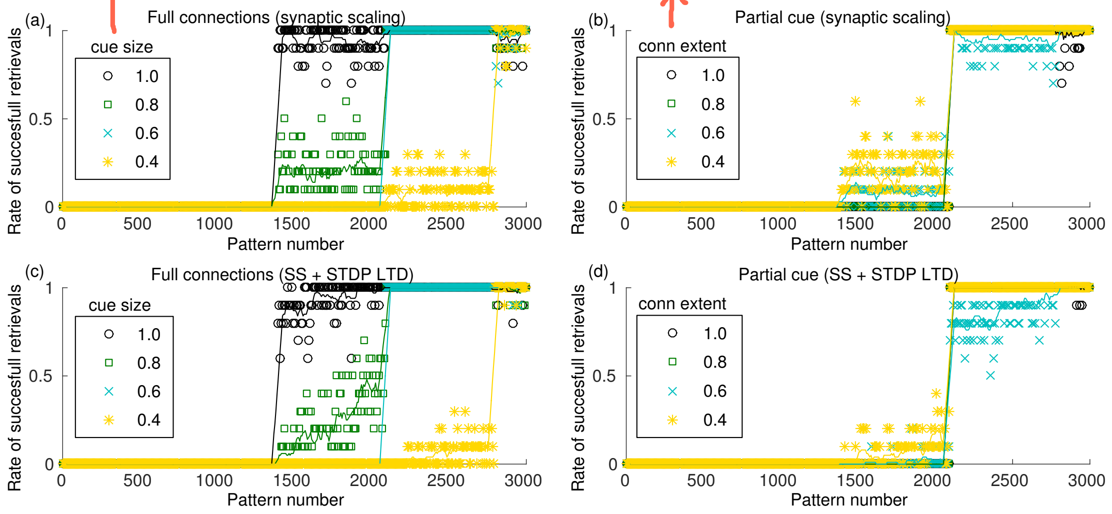


Figure 6. Rate of successful retrieval for the last 3000 stored patterns. Markers show the average retrieval rate for 10 simulations and lines the arithmetic mean of 10 rates around each point. A retrieval is successful when at least four patterns of the sequence were retrieved with an overlap larger than 0.5. (a) Simulation with all connections, variable cue sizes, and synaptic scaling; (b) simulation with partial cues of 0.6, varying connectivity extents and synaptic scaling; (c) and (d) same as (a) and (b), but with the addition of STDP-induced LTD.

memories by new ones. But if one adds STDP-induced LTD, the retrieval capability of the network does not change significantly.

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 ~~and~~ 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.

Under an evolutionary point of view, coding of memories using heteroassociative connections permits their storage using ~~less~~ fewer connections. Coding of a n-neuron pattern using autoassociative learning (?) requires n^2 connections per pattern, as all neurons in the pattern must connect to each other. With heteroassociative coding the larger pattern would be decomposed into a sequence of k smaller patterns of n/k neurons, requiring only $k*(n/k)^2 = n^2/k$ connections for the complete sequence. This results in a k-fold reduction in the number of connections, which is an important biological advantage. Nevertheless, we should consider that autoassociative networks may provide better noise tolerance and retrieval from smaller cues, since the retrieval can be performed iteratively in multiple steps. Consequently, the coding choice is a trade-off between number of connections and reliability.

No, there are non-Ising auto-assoc. networks

Retrieval from small cues seems to be easier with heteroassociative coding. 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 retrieve a complete pattern sequence using a cue of 40% of a single pattern, which represents only 5.7% of the original pattern encoded as a sequence of seven patterns. Heteroassociative coding also permits the storage of memories of different sizes. For example, larger complex memories could be coded in longer sequences, while simpler ones could use small sequences. If autoassociative coding was used, retrieval of a memory using cues of 5.7% would be unlikely to work, since the excitatory weights would need to be large enough to allow the activation of the remaining neurons in the pattern. But these same large weights would compromise the stability of the network during the retrieval, exacerbating the problem of crosstalk interference (Amit,

You're ignoring self-reinforcing dynamics...

1989). The same drawback applies when storing patterns of different sizes, since small patterns would require larger weights, causing the same stability and crosstalk problems caused when performing retrieval from small cues.

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, a mechanism would be required to disambiguate between these sequences during pattern sequence retrieval. (Levy, Hippocampus, 6, 579-90, 1996)

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 ~~the~~ persistent activity associated to memories in the different cortex regions.

External noise, pattern 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 pattern 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 coded 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 during pattern sequence retrieval.

The presence of both autoassociative and heteroassociative connections may use unnecessary resources. Although autoassociative connections can indeed be important for noise and limited synaptic connectivity tolerance, we showed that an exclusively heteroassociative network can operate correctly in these scenarios. Errors in the retrieval of a single pattern do not propagate to the next pattern and the network worked correctly even when we limited each neuron's connectivity to 40% of the neurons in the network. One limitation of the current study is that we did not include noise in the neural responses and 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. Consequently, most of the conclusions obtained from evaluating the effects of synaptic scaling and LTD on memory replacement, network oscillations, connectivity extents, and others, can also be partly applied to synfire chains.

This is vague. "partly" how?

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 can not be directly compared, the low number of motifs in the rodent CA3 is a evidence that heteroassociative may be used in the CA3 area.

a (missing noun)

temporarily (?)

System homeostasis and catastrophic forgetting

Enabling the storage of new memories requires the removal of old memories that are never recalled. According to the remote memories hypothesis, memories are temporally stored in the hippocampus to be later coded as remote memories in the cortex (Frankland and Bontempi, 2005). Although the hippocampal region may be necessary for retrieval of declarative memories for a long period (Winocura and Moscovitch, 2011), it is likely that the areas of the entorhinal cortex (EC) are the most relevant ones.

EC is major input to hippoc. This sentence doesn't make sense, or is very unclear.

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 just by including bounds in the synaptic weights (Nadal et al., 1986), which causes only the last stored memories to be remembered. This kind of memory is called palimpsest, but it can store only a small number of patterns.

are

Some form of unlearning is required, and there 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.

✓

As a last point, we note that the replacement of old memories for new ones also receive 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.

ACKNOWLEDGMENTS

The computational infrastructure for the simulations was provided by Universidade Federal do ABC (UFABC). The authors would like to acknowledge the members of the Timing and Cognition Lab[1], at UFABC, for their insightful suggestions during project development.

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.

---

[1] `http://neuro.ufabc.edu.br/timing/`.

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.

Buonomano, D. V. (2005). A learning rule for the emergence of stable dynamics and timing in recurrent networks. *Journal of Neurophysiology*, 94(4):2275–83.

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

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.

Dan, Y. and Poo, M.-M. (2004). Spike timing-dependent plasticity of neural circuits. *Neuron*, 44(1):23–30.

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.

Fiete, I. R., Senn, W., Wang, C. Z. H., and Hahnloser, R. H. R. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron*, 65(4):563–76.

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*, 3(4):128–135.

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

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

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

Fuster, J. M. and Alexander, G. E. (1971). Neuron Activity Related to Short-Term Memory. *Science*, 173(3997):652–654.

Gloveli, T., Dugladze, T., Rotstein, H. G., Traub, R. D., Monyer, H., Heinemann, U., Whittington, M. a., and Kopell, N. J. (2005). Orthogonal arrangement of rhythm-generating microcircuits in the hippocampus. *Proceedings of the National Academy of Sciences*, 102(37):13295–300.

Guzman, S. J., Schlögl, A., Frotscher, M., and Jonas, P. (2016). Synaptic mechanisms of pattern completion in the hippocampal CA3 network. *Science*, 353(6304):109–120.

Hájos, N., Pálhalmi, J., Mann, E. O., Németh, B., Paulsen, O., and Freund, T. F. (2004). Spike timing of distinct types of GABAergic interneuron during hippocampal gamma oscillations in vitro. *Journal of Neuroscience*, 24(41):9127–37.

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience*, 11(7):523–32.

Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(April):2554–2558.

Jensen, O. and Lisman, J. E. (1996). Hippocampal CA3 region predicts memory sequences: accounting for the phase precession of place cells. *Learning & Memory*, 3(2-3):279–87.

Lisman, J. (2005). The theta/gamma discrete phase code occuring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus*, 15(7):913–22.

Lisman, J. and Idiart, M. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science*, 267(5203):1512–5.

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

Lisman, J. E., Talamini, L. M., and Raffone, A. (2005). Recall of memory sequences by interaction of the dentate and CA3: a revised model of the phase precession. *Neural Networks*, 18(9):1191–201.

Lytton, W. W. (1998). Adapting a feedforward heteroassociative network to Hodgkin-Huxley dynamics. *Journal of Computational Neuroscience*, 5(4):353–64.

Meeks, J. P. and Mennerick, S. (2007). Action potential initiation and propagation in CA3 pyramidal axons. *Journal of Neurophysiology*, 97(5):3460–72.

Miles, R. and Wong, R. (1986). Excitatory synaptic interactions between CA3 neurones in the guinea-pig hippocampus. *Journal of Physiology*, 373:397–418.

Miyata, R., Ota, K., and Aonishi, T. (2013). Optimal design for hetero-associative memory: hippocampal CA1 phase response curve and spike-timing-dependent plasticity. *PLoS One*, 8(10):e77395.

Nadal, J. P., Toulouse, G., Changeux, J. P., and Dehaene, S. (1986). Networks of Formal Neurons and Memory Palimpsests. *Europhys. Lett*, 1(10):535–542.

Nakazawa, K., Quirk, M. C., Chitwood, R., atanabe, M., Yeckel, M. F., Sun, L., Kato, A., Carr, C., Johnston, D., Wilson, M., and Tonegawa, S. (2002). Requirement for hippocampal CA3 NMDA receptors in associative memory recall. *Science*, 297(5579):211–8.

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

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.

Winocura, G. and Moscovitch, M. (2011). Memory Transformation and Systems Consolidation. *Journal of the International Neuropsychological Society*, 17(05):766–780.

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