



Published in final edited form as:

Neuron. 2016 May 4; 90(3): 622–634. doi:10.1016/j.neuron.2016.03.019.

Mnemonic functions for nonlinear dendritic integration in hippocampal pyramidal circuits

Patrick Kaifosh^{1,2} and Attila Losonczy^{1,3,4}

¹Department of Neuroscience, Columbia University, New York, NY 10032, USA

²Center for Theoretical Neuroscience, Columbia University, New York, NY 10032, USA

³Kavli Institute for Brain Science, Columbia University, New York, NY 10032, USA

⁴Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY 10032, USA

Summary

We present a model for neural circuit mechanisms underlying hippocampal memory. Central to this model are nonlinear interactions between anatomically and functionally segregated inputs onto dendrites of pyramidal cells in hippocampal areas CA3 and CA1. We study the consequences of such interactions using model neurons in which somatic burst-firing and synaptic plasticity are controlled by conjunctive processing of these separately integrated input pathways. We find that nonlinear dendritic input processing enhances the model's capacity to store and retrieve large numbers of similar memories. During memory encoding, CA3 stores heavily decorrelated engrams to prevent interference between similar memories, while CA1 pairs these engrams with information-rich memory representations that will later provide meaningful output signals during memory recall. While maintaining mathematical tractability, this model brings theoretical study of memory operations closer to the hippocampal circuit's anatomical and physiological properties, thus providing a framework for future experimental and theoretical study of hippocampal function.

Introduction

The mammalian hippocampus supports episodic memory formation and storage (Squire & Wixted, 2011) by passing information through its canonical trisynaptic circuit: from the dentate gyrus (DG) input node, to area CA3, and then to the CA1 output node. Theoretical and experimental studies predict that specialized computational operations are carried out by each of these subregions during memory processing (Marr, 1971; McClelland & Goddard, 1996; Nakazawa et al., 2002; Guzowski et al., 2004; Lee et al., 2004; Gold & Kesner, 2005;

Correspondence to: pwk2108@columbia.edu (P.K.), al2856@columbia.edu (A.L.).

Author Contributions

P.K. and A.L. conceived the study. P.K. developed the computational models and performed simulations and analyses. Both authors wrote the paper.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Kesner & Rolls, 2015). In particular, DG is implicated in input decorrelation (pattern separation), reducing interference between distinct memories of similar events (McHugh et al., 2007; Neunuebel & Knierim, 2014). The downstream CA3 area is thought to operate as a Hebbian autoassociative network, allowing memory storage and later recall of whole memories from partial cues (pattern completion). Various functions have been proposed for area CA1, including novelty detection (McClelland et al., 1995; Hasselmo et al., 2000; Lisman & Otmakhova, 2001; Vinogradova, 2001), and enrichment of the hippocampal output to the neocortex either by forming more information-rich re-encodings of CA3 ensembles during memory storage (McClelland & Goddard, 1996) or by redistributing information across a greater number of neurons during recall (Treves & Rolls, 1994).

A prominent feature of afferent connectivity to pyramidal cells (PCs) both in the CA3 and CA1 subregions is the anatomical segregation of functionally distinct input pathways. That is, synapses of the trisynaptic circuitry occupy the proximal dendrites of PCs (DG mossy fibers onto CA3 PCs, CA3 Schaffer collaterals onto CA1 PCs), while long range external inputs from the entorhinal cortex (EC, layer II to CA3, layer III to CA1) mainly innervate the distal dendrites (Andersen et al., 2006; Ahmed & Mehta, 2009). The electrical compartmentalization present within dendrites of hippocampal PCs (Spruston et al., 1994; Golding et al., 2005) indicates that these input pathways are initially processed independently. However, only a subset of theoretical models of hippocampal memory operations have considered this dual afferent connectivity (e.g. Treves & Rolls, 1992; McClelland & Goddard, 1996; Vinogradova, 2001), and none to our knowledge have accounted for the need for anatomical segregation of these input pathways. Even more strikingly, none of the hippocampal network models incorporates the assumption supported by extensive experimental evidence that dendritic input processing in hippocampal PCs is highly nonlinear, exhibiting several different types of dendritic spikes and plateau potentials (Golding & Spruston, 1998; Ariav et al., 2003; Gasparini et al., 2004; Jarsky et al., 2005; Losonczy & Magee, 2006; Spruston, 2008; Katz et al., 2009; Kim et al., 2012; Makara & Magee, 2013). Computational theories incorporating nonlinear dendritic processing into single-neuron or abstract network models suggest that nonlinear input processing within dendritic compartments can enhance the neurons' ability to process and store information (Koch et al., 1983; Archie & Mel, 2000; Poirazi & Mel, 2001; Poirazi et al., 2003a,b; Morita, 2008; Wu & Mel, 2009; Legenstein & Maass, 2011). Furthermore, nonlinear interactions between different input streams through dendritic spikes and backpropagating action potentials are known to induce long-lasting changes of synaptic strength and intrinsic excitability, and produce a distinct burst-firing output mode of PCs (Kamondi et al., 1998; Larkum et al., 1999; Golding et al., 2002; Jarsky et al., 2005; Sjöström & Häusser, 2006; Dudman et al., 2007; Tsay et al., 2007; Takahashi & Magee, 2009; Harvey et al., 2009; Epsztein et al., 2011; Xu et al., 2012; Larkum, 2013; Grienberger et al., 2014). Despite the overwhelming experimental evidence for the presence of dendritic nonlinearities and resulting burst-firing in PCs, the specific consequences of nonlinear dendritic processing for hippocampal memory operations at the circuit-level have, with few exceptions (Katz et al., 2007; Wu & Mel, 2009), remained unexplored.

Here, we present a model reflecting the separate nonlinear integration of and interaction between anatomically segregated excitatory inputs to hippocampal areas CA3 and CA1. We

show that nonlinear interaction between the intra-hippocampal and perforant path (EC) inputs provides a mechanism for the storage of partially decorrelated engrams. Nonlinear synaptic integration also provides a mechanism by which EC inputs to CA3 can trigger recall of the most related memory engrams, while noise in these inputs does not perturb the recalled activity pattern. The dependence of the degree of decorrelation on the nonlinear integration parameters allows for a scheme (Figure 1), analogous to that proposed by McClelland & Goddard (1996) in the context of a linear integration model, by which CA3 representations of similar memories can form non-interfering attractors, while CA1 representations can provide information-rich output to the neocortex. The model accounts for experimental findings independent of those which motivated its construction, including the different spatial remapping properties in CA3 and CA1 PCs (Leutgeb et al., 2007; Ziv et al., 2013), and the differential sensitivity of CA1 spatial representations to lesion of inputs from CA3 and EC (Brun et al., 2002, 2008).

Results

Formation of partially decorrelated engrams through nonlinear integration of segregated inputs to CA3 PCs

We first constructed an abstract model of CA3 with nonlinear interactions between the anatomically segregated distal (EC afferents) and proximal (DG afferents and CA3 recurrent collaterals) excitatory inputs to CA3 PCs. This model consists of a network of two-compartment neurons receiving separate external inputs to their proximal and distal compartments and connected recurrently through modifiable binary synapses onto their proximal compartments (Figure 2A). To reflect nonlinear dendritic processing, the inputs to each compartment are thresholded to produce compartmental outputs, with the values of 0 or 1 respectively corresponding to input levels below or above the local threshold for dendritic integration.

To store memories in this network, we developed a memory encoding scheme motivated by the capability for combined distal and proximal input to evoke burst-firing (Takahashi & Magee, 2009; Larkum et al., 1999, 2009) and plasticity at intra-hippocampal synapses (Dudman et al., 2007; Basu et al., 2013; Han & Heinemann, 2013). When encoding memories, we consider the CA3 recurrent synapses to be suppressed (Hasselmo et al., 1995), such that the proximal inputs are determined entirely from the DG afferents. Model neurons receiving suprathreshold distal and proximal inputs enter a burst-firing state (Figure 2B). Potentiation occurs at recurrent synapses between burst-firing neurons and is balanced by depression at synapses between burst-firing neurons and neurons not burst-firing (Figure 2C). Both synaptic potentiation and depression are applied probabilistically to reduce the rate at which stored memories are overwritten by the storage of more recent memories (Amit & Fusi, 1992, 1994). These plasticity processes create engrams consisting of the co-active burst-firing neurons.

We then considered the characteristics of engrams that would be formed for pairs encoding events in which the distal inputs – conveying the cortical representations – were similar, while the proximal inputs – conveying the output of DG pattern separation – were less similar. For this purpose, we report the correlation between burst-firing patterns evoked by

pairs of input patterns in which the proximal input patterns are less correlated than the distal input patterns (Figure 2D; Experimental Methods). The engram correlation decreases monotonically as the proximal correlation (c_p , between two thresholded proximal input patterns) decreases, and under the assumption that proximal correlation is less than the distal correlation (c_d , between two thresholded distal input patterns), the engram correlation is always less than distal correlation. Importantly, the engram correlation also depends on the fractions of neurons in which the distal and proximal compartments are activated, termed the distal (s_d) and proximal (s_p) sparsities, whose product determines the engram sparsity ($s = s_p s_d$). Burst-firing patterns are more constrained by, and thus have correlations generally more similar to those of the sparser of the two compartments. Overall, this analysis shows that nonlinear interactions between distal and proximal inputs provide a mechanism for creating attractor patterns that are decorrelated to an intermediate degree that depends on the distal and proximal sparsities.

For mathematical simplicity throughout the remainder of this study, we focus on the limit in which the proximal inputs are fully decorrelating, regardless of the degree of distal correlation. In this limit, the engram correlation, $c = c_d(1-s_d)/(1-s) = c_d(1-s/s_p)/(1-s)$, interpolates between the distal correlation and zero, as the distal sparsity varies between s and 1, or equivalently as the proximal sparsity varies between 1 and s .

Illustration of encoding and recall with pictorial network representations

We illustrate the storage and recall of correlated memories using pictorial network representations, in which each of the 128×128 pixels displays the state of a neuron in a fully connected recurrent network of 16384 neurons (Figure 3). We first show the contrasting case of a single network of one-compartment neurons with a single input pathway (Figure 3A). During memory encoding, external inputs activate ensembles of neurons, with Hebbian plasticity dictating the potentiation of synapses between co-active neurons. Following encoding of three correlated memory patterns, recall is tested: the external inputs initialize the network with a cue similar to a stored pattern, and then the network activity evolves according to the dynamics dictated by the recurrent connections (Experimental Procedures). As expected, the correlated memory engrams interfere: regardless of which cue initiates recall, the network converges to a mixed attractor that combines features of all three memories. We note that this single input pathway architecture is not compatible with decorrelation of patterns prior to storage, the reason being that any recall cue passing along this same pathway would then also become decorrelated and thus not resemble the corresponding engram sufficiently to initialize the network within the appropriate attractor basin.

We next show the case of our network of two-compartment neurons with separate distal and proximal input pathways (Figure 3B), with parameters set to match the engram sparsity of the single compartment case. During encoding, the distal (EC) inputs, encoding the content of the memory, combine with decorrelated proximal (DG) inputs to determine the pattern of burst-firing neurons that forms the engram. Following encoding of three engrams with correlated distal input patterns, the distal input pathway provides a recall cue similar to one of the input patterns that was provided during encoding. The proximal inputs play no role

during recall in this model because the same circuitry that decorrelates activity patterns for similar memories would be expected to produce a decorrelated, hence useless, recall cue. Under the influence of this distal recall cue, as explained in the following section, the network's recurrent activity then converges to the cued memory engram consisting of the set of neurons that were burst-firing together during memory encoding. The network of two-compartment neurons thus successfully stores and recalls engrams corresponding to similar memories.

Influences of nonlinear integration on recall dynamics

To model the role of nonlinear dendritic input integration during memory recall, we again focused on the ability of conjunctive distal and proximal input to evoke burst-firing. We modeled a recall state in which the CA3 recurrenents – no longer suppressed as during encoding (Hasselmo et al., 1995) – dominated over the DG inputs. such that the proximal input was determined entirely by recurrent connections. After initializing the network in a random state whose activity level matched the engram sparsity, we updated the binary state x^i of individual neurons asynchronously with the following update rule:

$$x^i = \begin{cases} 1 & \text{if } \sum_{j \neq i} (W_{ij} - W_I)(1 + bx_d^j)x^j > 0 \\ 0 & \text{if } \sum_{j \neq i} (W_{ij} - W_I)(1 + bx_d^j)x^j \leq 0 \end{cases}$$

where W_{ij} represents the binary recurrent synaptic weight from the j -th to the i -th neuron, W_I the effects of disinaptic inhibition, b the increase in output due to burst-firing, and x_d^j the thresholded distal input to the j -th neuron. According to this rule, proximal recurrent inputs determine which neurons become active, while distal (EC) inputs determine the level of output (non-bursting, 1; or bursting, $1 + b$) from each active neuron.

We applied this recall mechanism to the network represented in Figure 3B, with the same three stored engrams. When evolving according to these recall dynamics, the network activity would converge into one of the three attractors formed during encoding. We found that as the burst-firing ratio b increased, the probability of the network converging to the cued attractor increased (Figure 4A), and the number of asynchronous update cycles required for convergence decreased (Figure 4B). Even for relatively large burst-firing ratios, which gave high probabilities of recovering the cued attractor, the distal input cues did not perturb the attractor, which remained a fixed point of the dynamics whether the distal inputs were active or not.

For insight into this recall behavior, we related these modified recurrent dynamics to a pseudo-energy function (Experimental Procedures), analogous to the energy function whose local minima correspond to the dynamical attractors for networks with symmetric weights (Hopfield, 1982). We found that the distal inputs lower the energy of the stored engrams to a degree dependent on the burst-firing ratio and the overlap of the distal inputs with the engram (Figure 4C). Perfect recall cues – matching exactly the distal input patterns that were present during encoding – scale the recurrent input to each neuron by $1 + b$ when the network activity matches the cued engram, and because this scaling does not change the sign of the input to any neuron, the cued engram remains a fixed point of the dynamics. Noisy

recall cues, provided any correlations between the noise and the recurrent weights are negligible, leave the cued attractor unperturbed for similar reasons. Thus, the distal recall cues alter the pseudo-energy landscape to guide the dynamics toward cued attractors, but do so without changing the activity pattern at these fixed points to which the recall dynamics converge.

The results of this section show that, through nonlinear interactions with recurrent proximal inputs, the distal (EC) inputs can bias the recall dynamics to favor recovery of the memory engrams with which they overlap most strongly.

The interference-information trade-off

The production of partially decorrelated memory engrams through the combination of correlated distal (EC) inputs with decorrelated proximal (DG) inputs has been previously explored (McClelland & Goddard, 1996), though in the context of single-compartment neurons with linear input integration (Figure 5A). In that context, the degree of decorrelation is determined by the relative variances of the distal and proximal input (McClelland & Goddard, 1996), rather than the proximal and distal sparsities as in our model with nonlinear input integration. We sought to investigate the implications of these differing input integration schemes for storage and recall of similar memories.

Networks with either integration scheme are constrained by a trade-off between the amount of information that engrams contain about the cortical state (EC inputs) during encoding, and the degree to which engrams for similar memories are decorrelated to prevent interference (McClelland & Goddard, 1996). For example, perfectly decorrelated engrams would not interfere with each other, but once recalled would provide no meaningful information to downstream areas because of their complete lack of correlation with engram representing similar memories. To quantitatively compare the two integration schemes in terms of this trade-off, we computed the relationship between the pairwise engram correlation – in the limit of highly similar distal inputs – and the information that each engram contains about the distal input pattern that produced it (Figure 5B; Experimental Procedures). We calculated this relationship for different fixed levels of engram sparsity, while varying the distal and proximal sparsities or variances in the nonlinear or linear model, respectively. This relationship was qualitatively similar across the two models, although for low correlations, activity patterns in the two-compartment model contain more information about the distal input pattern.

Storage and recall of multiple similar memories

While the above analysis may indicate comparable performance of the linear and nonlinear integration schemes when pairs of similar memories are stored, it does not guarantee comparable performance when larger sets of similar memories are stored. To examine such scenarios, we studied networks in which the proximal inputs provided the dominant contribution to the engram so as to provide a high capacity for the storage of similar memories. In the model with nonlinear integration, we set the proximal sparsity to $s_p = 0.05$ and distal sparsity to $s_d = 0.5$. We set the parameters in the linear integration model to match

the nonlinear integration model with regard to engram sparsity and pairwise correlation between engrams for similar memories.

We first examined, for each integration scheme, the probability with which each active neuron in an engram would also be active in the engram for a highly similar memory (Figure 5C), modeled by taking the limit of perfect distal input correlation. In the case of nonlinear integration, all neurons within an engram have the same probability, equal to the proximal sparsity, of recurring in an engram for a highly similar memory. In the case of linear integration, however, the recurrence frequency is distributed over a range of values; most neurons have a low probability, less than the proximal sparsity, of recurring in a similar engram, while a smaller fraction of neurons recur with much higher probability. Since this recurrence probability is also the fraction of similar memory engrams in which the neuron is expected to be active, we hypothesized that these different distributions would imply different network behavior when large numbers of similar memories were stored.

To test this hypothesis, we created a network of 10000 neurons for each integration scheme. In each network we stored 200 engrams with independent proximal inputs, to reflect DG pattern separation, but identical distal (EC) inputs, to model the limiting case of high memory similarity. After storing these patterns, we assessed the integrity of the attractor for each of these engrams. To do this, we initialized the network with the activity pattern of each of the stored engrams, allowed the network activity to evolve according to the recurrent dynamics determined by the learned synaptic weights, and then measured the correlation between the resulting activity pattern and each of the encoded engrams (Figure 5D). With the linear integration scheme, the network activity evolved away from most recall cues, as evidenced by the paucity of strong correlations along the diagonal, and into a single common attractor, as evidenced by the horizontal banding pattern. In contrast, the network with the nonlinear integration typically remained near the location of the cue, indicating the presence of an intact memory-specific attractor. Network dynamics following initialization from perturbed version of encoded patterns and with asymmetric plasticity rules are explored in Supplementary Figures 1 and 2.

Examining the common attractor to which the network with linear input integration converged, we found that the active neurons were indeed those which had occurred in a high proportion of the stored engrams (Figure 5E), and we therefore refer to this attractor as a “composite attractor” henceforth. To track the formation of this attractor as progressively many patterns were stored, we determined a composite pattern after each new storage event by selecting the n neurons that had been active in the highest fraction of stored patterns, choosing n in each case to minimize the pseudo-energy of this composite pattern. The pseudo-energy of the composite pattern progressively decreases and drops well below that of the most recently stored engram (Figure 5F). The same analysis applied to the network with linear input integration shows that pseudo-energy of the composite state remains well above that of the most recently stored pattern. Calculations regarding the parameter conditions under which composite attractors are formed are presented in the Experimental Procedures. Asymmetric plasticity rules do not prevent the formation of the composite attractor in the linear integration model but can support the coexistence of memory-specific attractors despite the presence of this composite attractor (Supplementary Figures 1 and 2).

Overall, our comparison of the linear and nonlinear integration schemes indicates that while both the nonlinear and linear networks effectively decorrelate pairs of engrams, only the nonlinear network effectively decorrelates larger numbers of similar patterns. While all neurons recur in similar memories with equal probabilities in the nonlinear model, the uneven distribution of recurrence probabilities in the linear model promotes the formation of a spurious attractor that interferes with the recall of individual engrams.

Distinct roles for CA3 and CA1

We have observed that the parameters that maximize the capacity to store non-interfering memory engrams are those that also minimize the information that these engrams encode (Figure 5B). This antagonism, along with the lack of recurrent activity in hippocampal area CA1, suggests distinct roles for areas CA3 and CA1, as also noted by those studying the dual-input model with linear integration (McClelland & Goddard, 1996). In the recurrently connected area CA3, engrams should be strongly decorrelated to prevent similar memories from forming interfering attractors, but these engrams need not have high information content, provided they are paired with more information-rich representations in downstream CA1. Lacking strong excitatory recurrent connectivity, area CA1 can store correlated engrams without risk of interference during recall; the high information content of these engrams would suit area CA1's role in providing the major excitatory output from the hippocampus. These differing demands of the CA3 and CA1 networks can be satisfied if proximal activation patterns are sparser than distal activation patterns (i.e. $s_p < s_d$) in CA3, and distal activation patterns are sparser than proximal activation patterns (i.e. $s_d < s_p$) in CA1.

The combined CA3-CA1 hippocampal memory system can then be modeled as follows. During encoding (Figure 6A), sparse proximal activation patterns determined by DG inputs combine with less sparse distal activation patterns determined by EC inputs to produce the pattern of burst-firing neurons in area CA3. The resulting heavily decorrelated CA3 activity pattern provides feed-forward input to the proximal compartment of CA1 PCs, which is thresholded to produce a non-sparse proximal pattern. The CA1 PC burst-firing pattern results from the combination of this proximal pattern with a sparse distal activation pattern determined by EC inputs to CA1. Synapses from CA3 PCs onto CA3 and CA1 PCs are then modified according to the aforementioned Hebbian plasticity rule (Figure 2C). During recall (Figure 6B), distal inputs from EC provide a recall cue to area CA3, which then evolves according to the network dynamics dictated by its recurrent activity. From a modeling perspective, we can view CA3 and CA1 during recall as one combined recurrent network with some neurons, those of CA1, having no outgoing recurrent synapses. The CA1 activity thus tracks the CA3 activity such that, when the CA3 network converges to an activity pattern resembling a stored engram, the CA1 network likewise converges to the activity pattern that was co-active with that CA3 pattern during encoding. We note that, in this model, the distal inputs to area CA1 do not influence which CA1 activity pattern is recovered during recall.

Distal and proximal contributions to spatial tuning in CA3 and CA1

Lastly, we sought to relate this model of hippocampal memory function to experimental results regarding place cell properties of PCs in hippocampal areas CA3 and CA1. In particular, we wished to examine the model's predictions regarding differences in CA3 and CA1 spatial coding resulting from the distinct properties of distal and proximal input processing by CA3 and CA1 networks.

For this purpose, we constructed random input patterns that varied continuously across position in a periodic one-dimensional model environment (Experimental Procedures). We then thresholded these inputs to produce spatially dependent distal and proximal activation patterns for each neuron. For the model CA3 network, we set the distal and proximal thresholds such that at each location the fractions of neurons receiving suprathreshold distal and proximal input were $s_d = 0.5$ and $s_p = 0.05$, respectively, resulting in strongly decorrelated representations ($c = \frac{1}{39}$ for highly similar distal inputs) encoding a low level of information about the EC inputs ($I = 0.025$ bits/neuron). These thresholds and corresponding compartmental sparsities were reversed for the model CA1 network, resulting in weakly decorrelated representations ($c = \frac{19}{39}$ for highly similar distal inputs) encoding greater information about the distal input pattern ($I = 0.12$ bits/neuron).

The resulting CA3 and CA1 spatial tuning patterns, alongside the distal and proximal activation patterns giving rise to these tunings, are displayed in Figure 6C. The similarity of the CA1 activity patterns to their distal activations is notable in light of experimental reports that place field properties are maintained when proximal (Brun et al., 2002; but see Nakashiba et al., 2008) but not distal (Brun et al., 2008) excitatory inputs to this network are lesioned, the latter lesion resulting in larger and more dispersed CA1 place fields. Together with a compensatory mechanism that eliminated the requirement for suprathreshold input to the compartment receiving lesioned inputs while preserving nonlinear processing in the compartment receiving intact inputs, this model would account for these experimental results; the CA1 output would come to resemble the thresholded input coming from the non-lesioned pathway.

We next assessed the model's predictions for remapping of spatial representations in similar environments (Figure 6D,E). We simulated similar environments by producing dissimilar proximal input patterns, and distal patterns with varying degrees of similarity (Experimental Procedures). We then evaluated the level of correlation (Figure 6E) between activity patterns in the original environment (Figure 6C) and the similar environments (Figure 6D). In striking agreement with experimental studies of different remapping in CA3 and CA1 (Leutgeb et al., 2004), the model CA3 network remapped substantially in all three similar environments, while the model CA1 network remapped progressively with decreasing environmental similarity.

We thus see that this model – motivated entirely by considerations of circuit anatomy, nonlinear dendritic integration, and hippocampal memory function – accurately accounts for experimentally observed differences in the remapping properties of CA3 and CA1 place cells, as well as the differential importance of proximal and distal inputs to CA1 place field properties.

Discussion

Comparison with other models of CA3 and CA1

Our consideration of the nonlinear synaptic integration of segregated inputs to hippocampal PCs led us to a theory in which the roles of the various hippocampal subfield elements are similar to those proposed by McClelland & Goddard (1996). Area CA3 stores attractors that are decor-related by the orthogonalizing DG inputs, but can be recalled based on cues from EC inputs, while area CA1 pairs CA3 representations with recoded representations containing greater information about the cortical state at the time of encoding. Our work builds on this understanding by demonstrating that nonlinear dendritic integration enhances the ability of the network to store greater numbers of similar memories, and can serve as a recall cue not only by initializing the state of the recurrent CA3 network, but also by lowering the energy state of the cued engram without introducing artifacts due to the noisiness of the cue.

This model differs from other attractor based hippocampal theories (Treves & Rolls, 1992, 1994; Rolls & Treves, 1994). Regarding the roles of the separate input pathways to CA3, we propose that DG serves not to increase the information content of CA3 engrams (Treves & Rolls, 1992), but instead to orthogonalize and as a consequence decrease the information content of CA3 engrams. We consider the mutual information with the EC input pattern, not the DG input pattern (Treves & Rolls, 1992), to be the proper measure of the attractor's information content and interpretability. Instead of requiring associative plasticity of the EC inputs to CA3 (Treves & Rolls, 1992), we instead require these inputs to influence plasticity at recurrent synapses within CA3. In contrast to previous arguments for the needs for weak plastic EC inputs to cue recall and for strong DG inputs to overcome CA3 recurrence during encoding (Treves & Rolls, 1992), our analysis of the trade-off between attractor interference and information (McClelland & Goddard, 1996) explains why no single input pathway can suffice, whether it be plastic, nonplastic, strong, weak, or even modulated (Hasselmo et al., 1995). In our model, dendritic nonlinearities do not simply compensate for passive attenuation (Treves & Rolls, 1992), but are central to interactions between the dual input pathways to CA3 PCs.

We propose that area CA1 functions not to redistribute the information content of CA3 attractors across a larger number of cells (Treves & Rolls, 1994), but instead to attach mutual information to an otherwise uninformative CA3 attractor pattern. Rather than providing richer information about the cued aspects of the event during recall (Treves & Rolls, 1994), EC inputs to CA1 are in our model essential only during encoding, when they control plasticity at CA3-CA1 synapses to pair the information-poor attractor pattern in CA3 with an information-rich representation in CA1. This proposed function is arguably more useful, in that it ultimately enhances recalled information that is not already represented in cortex.

The proposed recall mechanism, by which orthogonalized CA3 activity patterns reactivate information-rich CA1 patterns, resembles Marr's proposal that hippocampal "codon" representations reactivating cortical representations through hippocampal-cortical synapses (Marr, 1971). The key difference here, of an intermediate CA1 layer between the codon representation and the cortical representations, has important implications for memory

capacity. The orthogonalization of representations in CA3 “codons” necessitates a substantial plasticity at synapses emanating from CA3 during memory formation. Given the stringent constraints on and memory capacity during online learning (Amit & Fusi, 1992, 1994), the extra synaptic resources provided by a CA1 region, whose entire set of Schaffer collateral synapses can be dedicated to learning such associations, would allow for a memory capacity not possible if such associations had to be learned through the small fraction of synapses that cortical neurons receive from the hippocampus.

Here, we have focused on a role for CA1 distinct from novelty detection (Hasselmo et al., 2000; Lisman & Otmakhova, 2001; McClelland et al., 1995). However, the model we present is not incompatible with such a role for CA1. The proposed associative potentiation of CA3-CA1 synapses during encoding would cause later repetition of the same CA3 activity pattern to provide much stronger feed-forward input to CA1. Nonlinear interaction between the CA3 and EC inputs to CA1 would also provide a mechanism for assessing the similarity of the recovered memory (CA3 inputs) to the recall cue (EC inputs).

Model predictions and relations to experimental data

Nonlinear interactions between spatially segregated and functionally distinct inputs, and the instructive role of the resulting dendritic spiking and somatic burst firing for synaptic and intrinsic plasticity, have been extensively demonstrated across various neocortical and hippocampal pyramidal circuits (Golding & Spruston, 1998; Larkum et al., 1999; Golding et al., 2002; Jarsky et al., 2005; Sjöström & Häusser, 2006; Dudman et al., 2007; Tsay et al., 2007; Takahashi & Magee, 2009; Xu et al., 2012; Basu et al., 2013; Larkum, 2013; Gambino et al., 2014). Recent studies have also provided compelling evidence for the presence of dendritic plateau potentials and somatic burst spiking in hippocampal pyramidal circuits in vivo (Kamondi et al., 1998; Harvey et al., 2009; Epsztein et al., 2011; Grienberger et al., 2014). Our model gives insight into how these cellular-level phenomena can allow dynamic coupling of two functionally and spatially distinct pathways, and provides experimental predictions about the resulting cellular and network dynamics during hippocampal spatial and memory operations.

Inherent in this proposed model for hippocampal memory are specific predictions for epoch-dependent requirements of specific input pathways and plasticity rules. DG inputs are predicted to be important specifically during encoding, with their disruption – provided sufficient compensatory mechanisms to ensure EC-driven CA3 attractor formation – leading to the interference of similar memories. EC inputs to CA3 are essential for both encoding and recall, whereas EC inputs to CA1 are required only for encoding. The latter prediction has been verified, though so far only in the case of hippocampal-dependent temporal associative learning (trace fear conditioning), for which EC layer 3 input to CA1 was found to be necessary during encoding but not recall (Suh et al., 2011).

We also predict that memory formation depends critically on plasticity at synapses from CA3 PCs to both CA3 and CA1 PCs, and that this plasticity is regulated by nonlinear interactions between distal and proximal inputs. Heterosynaptic control exerted by distal (EC) inputs over proximal (CA3) inputs has been observed in CA1 PCs (Takahashi & Magee, 2009; Dudman et al., 2007; Basu et al., 2013; Han & Heinemann, 2013). Moreover,

dendritic plateau potentials require inputs from EC layer III and are sufficient to induce rapid place field formation in CA1 PCs in vivo (Bittner et al., 2015). Similar behavior of CA3 PCs remains to be confirmed, although active dendritic input integration in CA3 PCs has been recently demonstrated (Kim et al., 2012; Makara & Magee, 2013), and quantitative arguments have been made for the interaction of EC and DG inputs driving CA3 PC activity (Lisman, 1999). Plasticity of neither DG nor EC inputs is predicted to be essential during the formation of memories, though such plasticity rules could serve ancillary network maintenance functions.

The model also makes predictions regarding activity of hippocampal circuit elements following plasticity and input manipulation. For example, under in vivo conditions, a higher fraction of CA3 PCs should respond to proximal stimulation with burst firing due to the higher fraction receiving strong distal input. Consistent with the reported sensitivity of CA1 PC spatial tuning properties to lesion of inputs from EC but not from CA3 (Brun et al., 2002, 2008), the predicted differences in distal and proximal sparsity also imply a greater dependence of CA1 tuning on distal inputs, with CA3 tuning being instead more dependent on proximal inputs. Accordingly, environmental perturbations, evoking stronger changes in proximal than distal inputs (Neunuebel & Knierim, 2014), should cause greater changes in CA3 than in CA1 (Leutgeb et al., 2007). Viewing elapsed time as a form of environmental perturbation, we predict that long-term spatial remapping in CA3 should be much stronger than in CA1. While experimentally characterized long-term place field dynamics in CA1 (Ziv et al., 2013) match our predictions for remapping between similar environments (Figure 6B), CA3 place fields are reported to be more stable than CA1 place fields, though on much shorter time-scales and in familiar environments (Mankin et al., 2012), where recurrent dynamics within CA3 might compensate for temporal-coding-related changes in DG and EC inputs.

Outlook

We have presented a hippocampal memory model motivated by the anatomical segregation of excitatory inputs to hippocampal areas CA3 and CA1. This model builds on previous attractor-based models of memory, but extends them through consideration of nonlinear interactions between separate input pathways during PC input integration. Our model allows for the storage and recall of information rich memory engrams, which do not interfere even when large numbers of highly similar memories are stored. By reducing the gap between the anatomical details of the hippocampus and abstract models of hippocampal memory, this work provides a framework for future experimental and modeling investigations into hippocampal functions.

Experimental Procedures

All simulations, numerical integrations, and visualizations were performed with Python, Scipy, and matplotlib.

Memory encoding

We initialize the recurrent synapses with random weights sampled from a Bernoulli (i.e. 0 or 1) distribution with mean equal to the equilibrium fraction of potentiated synapses. Following the presentation of each input pattern, weights are updated according to a stochastic Hebbian rule: the transition probabilities for the weight W_{ij} of the synapse from the j -th neuron to the i -th neuron are

$$p(W_{ij}^{\text{new}}=1|W_{ij}^{\text{old}}=0)=p_+x_d^ix_p^ix_d^jx_p^j$$

$$p(W_{ij}^{\text{new}}=0|W_{ij}^{\text{old}}=1)=p_-[(1-x_d^ix_p^i)x_d^jx_p^j+x_d^ix_p^i(1-x_d^jx_p^j)],$$

where the binary variables X^i and X^j represent the burst-firing state of the neurons. In our simulations, we set $p_+ = 0.25$ and $p_- = 2s_p p_+$. For the asymmetries plasticity rules in Supplementary Figures 1 and 2, the value of $p(W_{ij}^{\text{new}}=0|W_{ij}^{\text{old}}=1)$ was changed to either $2p_-(1-x_d^ix_p^i)x_d^jx_p^j$ or $2p_-x_d^ix_p^i(1-x_d^jx_p^j)$.

In the case of the two-compartment model, the burst-firing state X^i is equal to the product $x_d^ix_p^i$, where x_d^i and x_p^i , which take values of 0 or 1, represent respectively the distal and proximal activations (i.e. thresholded inputs) of the i -th neuron.

In the one-compartment model, the distal and proximal inputs are each Gaussian distributed with mean zero and variance σ_d^2 or $\sigma_p^2=1-\sigma_d^2$, respectively. A neuron is activated ($X^i = 1$) if the summed proximal and distal inputs exceeded a threshold, which we set to $\theta = \sqrt{2} \operatorname{erf}^{-1}(1-2s)$ to impose a coding level s ; otherwise, the neuron is inactive ($X^i = 0$).

Here, erf denotes the error function defined by $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$.

Recall Dynamics

For recall dynamics without distal inputs, we followed standard approaches for attractor networks with non-negative weights and activity states (Denker, 1986). The recurrent input to the i -th neuron was calculated as $v_i = \sum_j (W_{ij} - W_j) x^j$, where the binary variable $x^j \in \{0, 1\}$ represents the activity of the j -th neuron. The variable W_j , representing disinhibitory inhibition, was set to $(1.2)s_p p_+ / [s_p p_+ + 2(1 - s_p)p_-]$ based on considerations outlined in the subsection below on synapse equilibria. We repeatedly cycled through the neurons, setting x^i to 1 whenever $v_i > 0$, and setting x^i to 0 otherwise. We continued this process until convergence to a fixed point or limit cycle. When considering the contribution of distal inputs to recall dynamics in the two-compartment model, we modified the equation input received by the i -th neuron as follows: $v_i = \sum_{j \neq i} (W_{ij} - W_i)(1 + bx_d^j)x^j$, where x_d^j is the binary thresholded distal input to the j -th neuron, and b is a constant representing the amount of extra output due to burst-firing in the presence of suprathreshold distal input.

Pseudo-energy calculations

In the case of symmetric synaptic weights and no distal input, the above recall dynamics correspond to the following energy function:

$$E = - \sum_{\substack{i, j=1 \\ i \neq j}} (W_{ij} - W_I) x^i x^j.$$

To reflect the modified dynamics of the two-compartment network in the presence of distal inputs, this energy function can be altered as follows:

$$E' = - \sum_{\substack{i, j=1 \\ i \neq j}} (W_{ij} - W_I) (1 + bx_d^i) x^i (1 + bx_d^j) x^j.$$

Here, the modification $(1 + bx_d^i)$ is interpreted – in the sense that this energy model is used with stochastic dynamics at finite temperature – as reflecting the increased likelihood that a neuron will respond to proximal input if it also receives distal input, and the modification $(1 + bx_d^j)$ can represent the increased output magnitude (burst-firing) of a neuron receiving distal input. When plotting the pseudo-energies (Figures 4B, 5F), we normalized both of these quantities by $N^2 s^2 (1 - W_I)$, with N representing the number of neurons.

In the interest of biological plausibility, we did not enforce symmetric weight changes resulting from the probabilistic weight updates. As a consequence of the resulting random asymmetries in synaptic weights, these energy functions only approximately govern the recall dynamics (with the approximation quality increasing with network size) and are therefore referred to as the pseudo-energies throughout.

Synapse equilibria and conditions for engram recall

Under this model, the equilibrium mean synapse weight is $\bar{W}_{ij} = \frac{p_+ s}{2p_- (1-s) + p_+ s} \approx \frac{p_+ s}{2p_-}$. Immediately following storage of a new engram, the fraction of potentiated synapses between neurons active within the engram becomes $\tilde{W}_{ij} = \bar{W}_{ij} + (1 - \bar{W}_{ij}) p_+$. Also, when a large number of engrams with highly similar distal inputs are stored, the mean weight of synapses between neurons receiving suprathreshold distal input converges to

$$\tilde{W}_{ij} = \frac{p_+ s p}{2p_- (1-s p) + p_+ s p} \approx \frac{p_+ s p}{2p_-}.$$

Conditions for recall can be derived in the limit of large numbers of neurons by considering two requirements. First, neurons within an engram must provide net excitation to each other, i.e. $\tilde{W}_{ij} > W_I$. Second, neurons within an engram must provide net inhibition to neurons that are not active as part of the engram, but which may be active in engrams for highly similar memories, resulting in the condition, resulting in the inequality $\tilde{W}_{ij} < W_I$. The simultaneous satisfiability of these two conditions requires that $\tilde{W}_{ij} < \hat{W}_{ij}$, which is equivalent to the

$$\text{condition } p_+ > \frac{\tilde{W}_{ij} - \bar{W}_{ij}}{1 - \bar{W}_{ij}}.$$

The second condition can be adapted to the case of the one-compartment model by replacing s_p with the recurrence frequency (Figure 5C) in the calculation of \tilde{W}_{ij} . The high values for \tilde{W}_{ij} for the neurons with the highest recurrence probabilities (i.e. the upper tail of the distribution in Figure 5C) results in a tendency for these neurons to become activated during the recall of engrams in which they were not activated, ultimately leading to the formation of composite attractors. The conditions above indicate that the avoidance of such composite attractors would require higher values of W_I and p_+ ; however, higher rates of synaptic turnover would increase the rate at which memories were over-written and (Amit & Fusi, 1992, 1994) and thus limit the memory capacity of the network.

Recurrence frequencies

For the two-compartment model, the expected frequency with which a neuron recurs in engrams for similar memories (the recurrence frequency, Figure 5C) is equal to the proximal sparsity s_p . For the one-compartment model, the probability density of the expected recurrence frequency is given by

$$p(x) = \frac{\sigma_p}{\sigma_d} \exp \left\{ (\text{erf}^{-1}(2x-1))^2 - \frac{(\theta + \sqrt{2\sigma_p^2} \text{erf}^{-1}(2x-1))^2}{2\sigma_d^2} \right\}.$$

This expression is obtained by considering the distribution of distal input levels conditioned on a neuron being active, and the probability of activation given each level of distal input.

Engram correlations

Since each neuron is independent during encoding, correlation between two engrams is equivalent to the correlation between the activity of a single neuron in each of the engrams. In general, the correlation c between two Bernoulli random variables X^a , X^b each with mean m can be calculated in terms of the expectation of their product:

$$c = \frac{E[X^a X^b] - m^2}{m(1-m)}.$$

Applying this formula to the two-compartment model, we set $X^a = x_d^a x_p^a$ and $X^b = x_d^b x_p^b$, and noting the independence of the proximal and distal activations, calculate

$E[X^a X^b] = E[x_d^a x_d^b] E[x_p^a x_p^b]$. When we can then reapply the same formula to calculate the product expectations in terms of the distal and proximal correlations – c_d and c_p , respectively – to obtain an expression for the engram correlation in terms of the correlations of the distal and proximal activations:

$$c = \frac{c_p c_d (1-s_p)(1-s_d) + c_p s_d (1-s_p) + c_d s_p (1-s_d)}{1-s}.$$

In Figure 5B, we consider the limit of perfect distal correlation ($c_d = 1$) and perfect proximal decorrelation ($c_p = 0$), in which case the above formula simplifies to $c = (s_p - s)(1 - s)$.

For the one-compartment network, in the limit of perfectly correlated distal inputs and perfectly decorrelated proximal inputs, the product expectation can be calculated by integrating over the distribution of the distal input values:

$$E[X^a X^b] = \frac{1}{\sqrt{2\pi\sigma_d^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x_d^2}{2\sigma_d^2}\right) \left(\frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{\theta - x_d}{\sqrt{2\sigma_p^2}}\right)\right]\right)^2 dx_d.$$

We obtained the correlation values in Figure 5B by calculating this integral numerically and substituting the result into the general formula above.

Mutual information

Since each neuron is independent during encoding, the mutual information between the engram and the distal input pattern is linear in the number of neurons. We therefore focus on calculating the mutual information between the activity of a single neuron and its distal input. For the two-compartment network, this quantity is calculated as

$$-I(x_d x_p, x_d) = H(x_d x_p) - H(x_d x_p | x_d) = h(s) - s_d h(s_p),$$

where $h(x) = -x \log_2(x) - (1 - x) \log_2(1 - x)$ is the entropy of a Bernoulli random variable with mean x . For the one-compartment network, this quantity is calculated as

$$I(x_d + x_p > \theta, x_d) = H(x_d + x_p > \theta) - H(x_d + x_p > \theta, x_d) \\ = h(s) - \frac{1}{\sqrt{2\pi\sigma_d^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x_d^2}{2\sigma_d^2}\right) h\left(\frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{\theta - x_d}{\sqrt{2\sigma_p^2}}\right)\right]\right) dx_d,$$

which we integrated numerically when creating Figure 5B. To obtain the mutual information per active neuron, we divide these quantities by the engram sparsity s .

The parametric curves in Figure 5B were obtained by varying s_p and s_d under the restriction that $s_p s_d = s$ in the case of the two compartment model, or by varying s_p^2 and s_d^2 under the restriction that $s_p^2 + s_d^2 = 1$ and with a fixed threshold θ in the case of the one-compartment model.

Spatial Tunings

We generated random spatial input patterns that varied continuously across the unit circle $[0, 2\pi)$. The continuous valued distal and proximal inputs to the k -th neuron are determined as

$$I_\alpha^k(\theta) = \sum_{i=1}^{100} f(\theta | \theta_\alpha^{i,k}, \frac{100}{\pi^2}), \text{ where } f(\theta | \mu, \kappa) \text{ is the probability density function of a von Mises distribution centered at } \mu \text{ and with concentration } \kappa, \text{ and where each } \theta_i \text{ is independently}$$

sampled from a uniform distribution over $[0, 2\pi)$. Then index α takes values d and p to denote the distal and proximal inputs, respectively. We determined distal and proximal threshold such that, at a given position along the unit circle, the probability of the input being above threshold was equal to s_d or s_p , respectively. We then applied these thresholds to the randomly generated continuous functions to obtain binary distal and proximal activation patterns for each model neuron.

Tunings for similar environments were created by regenerating entirely new proximal input patterns and perturbing the distal input patterns. The perturbed distal activation patterns were constructed by thresholding continuous input patterns of the form

$\tilde{I}_p^k(\theta) = \sum_{i=1}^{100} f(\theta | \theta_\alpha^{i,k} + \delta\theta_\alpha^{i,k}, \frac{100}{\pi^2})$, where the values of $\theta_\alpha^{i,k}$ were the same as used to generate the inputs in the original environment, and where the perturbations $\delta\theta_\alpha^{i,k}$ were sampled from a von Mises distribution with a mean of zero and concentration (κ) of 1000, 10, and 1 for the environments with high, medium, and low similarity, respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank L.F. Abbott, N.B. Danielson, K.D. Miller, and J.D. Zaremba for helpful discussions and comments on the manuscript. P.K. was supported by a Howard Hughes Medical Institute International Student Research Fellowship and NSERC PGS M and PGS D awards during work on this project. A.L. is supported by the McKnight Memory and Cognitive Disorders Award, the Human Frontiers Science Program, the Brain and Behavior Research Foundation, NIMH1R01MH100631, and NINDS1U01NS090583.

References

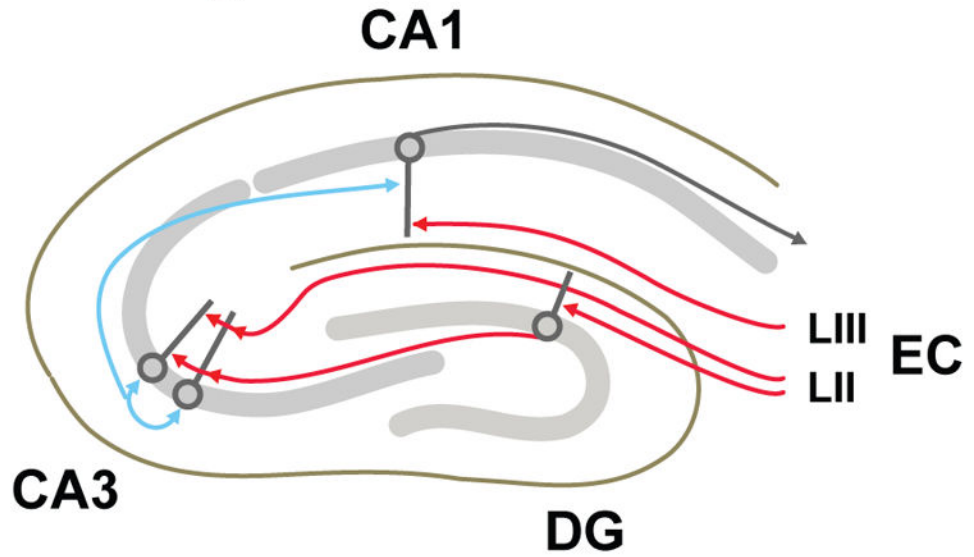
- Ahmed OJ, Mehta MR. The hippocampal rate code: anatomy, physiology and theory. *Trends Neurosci.* 2009; 32:329–338. [PubMed: 19406485]
- Amit DJ, Fusi S. Constraints on learning in dynamic synapses. *Network: Computation in Neural Systems.* 1992; 3:443–464.
- Amit DJ, Fusi S. Learning in neural networks with material synapses. *Neural Computation.* 1994; 6:957–982.
- Andersen, P.; Morris, R.; Amaral, D.; Bliss, T.; O'Keefe, J. *The hippocampus book.* Oxford University Press; USA: 2006.
- Archie KA, Mel BW. A model for intradendritic computation of binocular disparity. *Nat Neurosci.* 2000; 3:54–63. [PubMed: 10607395]
- Ariav G, Polsky A, Schiller J. Submillisecond precision of the input-output transformation function mediated by fast sodium dendritic spikes in basal dendrites of CA1 pyramidal neurons. *J Neurosci.* 2003; 23:7750–7758. [PubMed: 12944503]
- Basu J, Srinivas KV, Cheung SK, Taniguchi H, Huang ZJ, Siegelbaum SA. A cortico-hippocampal learning rule shapes inhibitory microcircuit activity to enhance hippocampal information flow. *Neuron.* 2013; 79:1208–1221. [PubMed: 24050406]
- Bittner KC, Grienberger C, Vaidya SP, Milstein AD, Macklin JJ, Suh J, Tonegawa S, Magee JC. Conjunctive input processing drives feature selectivity in hippocampal ca1 neurons. *Nature neuroscience.* 2015; 18:1133–1142. [PubMed: 26167906]
- Brun VH, Leutgeb S, Wu HQ, Schwarcz R, Witter MP, Moser EI, Moser MB. Impaired spatial representation in CA1 after lesion of direct input from entorhinal cortex. *Neuron.* 2008; 57:290–302. [PubMed: 18215625]

- Brun VH, Otnass MK, Molden S, Steffenach HA, Witter MP, Moser MB, Moser EI. Place cells and place recognition maintained by direct entorhinal-hippocampal circuitry. *Science*. 2002; 296:2243–2246. [PubMed: 12077421]
- Denker JS. Neural network refinements and extensions. *AIP Conf Proc*. 1986; 151:121–128.
- Dudman JT, Tsay D, Siegelbaum SA. A role for synaptic inputs at distal dendrites: instructive signals for hippocampal long-term plasticity. *Neuron*. 2007; 56:866–879. [PubMed: 18054862]
- Epsztein J, Brecht M, Lee AK. Intracellular determinants of hippocampal CA1 place and silent cell activity in a novel environment. *Neuron*. 2011; 70:109–120. [PubMed: 21482360]
- Gambino F, Pagès S, Kehayas V, Baptista D, Tatti R, Carleton A, Holtmaat A. Sensory-evoked Itp driven by dendritic plateau potentials in vivo. *Nature*. 2014; 515:116–119. [PubMed: 25174710]
- Gasparini S, Migliore M, Magee JC. On the initiation and propagation of dendritic spikes in CA1 pyramidal neurons. *J Neurosci*. 2004; 24:11046–11056. [PubMed: 15590921]
- Gold AE, Kesner RP. The role of the CA3 subregion of the dorsal hippocampus in spatial pattern completion in the rat. *Hippocampus*. 2005; 15:808–814. [PubMed: 16010664]
- Golding NL, Mickus TJ, Katz Y, Kath WL, Spruston N. Factors mediating powerful voltage attenuation along ca1 pyramidal neuron dendrites. *J Physiol*. 2005; 568:69–82. [PubMed: 16002454]
- Golding NL, Spruston N. Dendritic sodium spikes are variable triggers of axonal action potentials in hippocampal ca1 pyramidal neurons. *Neuron*. 1998; 21:1189–1200. [PubMed: 9856473]
- Golding NL, Staff NP, Spruston N. Dendritic spikes as a mechanism for cooperative long-term potentiation. *Nature*. 2002; 418:326–331. [PubMed: 12124625]
- Grienberger C, Chen X, Konnerth A. NMDA receptor-dependent multidendrite Ca²⁺ spikes required for hippocampal burst firing in vivo. *Neuron*. 2014; 81:1274–1281. [PubMed: 24560703]
- Guzowski JF, Knierim JJ, Moser EI. Ensemble dynamics of hippocampal regions CA3 and CA1. *Neuron*. 2004; 44:581–584. [PubMed: 15541306]
- Han EB, Heinemann SF. Distal dendritic inputs control neuronal activity by heterosynaptic potentiation of proximal inputs. *J Neurosci*. 2013; 33:1314–1325. [PubMed: 23345207]
- Harvey CD, Collman F, Dombeck DA, Tank DW. Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature*. 2009; 461:941–946. [PubMed: 19829374]
- Hasselmo ME, Fransen E, Dickson C, Alonso AA. Computational modeling of entorhinal cortex. *Ann N Y Acad Sci*. 2000; 911:418–446. [PubMed: 10911889]
- Hasselmo ME, Schnell E, Barkai E. Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *J Neurosci*. 1995; 15:5249–5262. [PubMed: 7623149]
- Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA*. 1982; 79:2554–2558. [PubMed: 6953413]
- Jarsky T, Roxin A, Kath WL, Spruston N. Conditional dendritic spike propagation following distal synaptic activation of hippocampal CA1 pyramidal neurons. *Nat Neurosci*. 2005; 8:1667–1676. [PubMed: 16299501]
- Kamondi A, Acsády L, Buzsáki G. Dendritic spikes are enhanced by cooperative network activity in the intact hippocampus. *J Neurosci*. 1998; 18:3919–3928. [PubMed: 9570819]
- Katz Y, Kath WL, Spruston N, Hasselmo ME. Coincidence detection of place and temporal context in a network model of spiking hippocampal neurons. *PLoS Comput Biol*. 2007; 3:e234. [PubMed: 18085816]
- Katz Y, Menon V, Nicholson DA, Geinisman Y, Kath WL, Spruston N. Synapse distribution suggests a two-stage model of dendritic integration in ca1 pyramidal neurons. *Neuron*. 2009; 63:171–177. [PubMed: 19640476]
- Kesner RP, Rolls ET. A computational theory of hippocampal function, and tests of the theory: new developments. *Neurosci Biobehav Rev*. 2015; 48:92–147. [PubMed: 25446947]
- Kim S, Guzman SJ, Hu H, Jonas P. Active dendrites support efficient initiation of dendritic spikes in hippocampal CA3 pyramidal neurons. *Nat Neurosci*. 2012; 15:600–606. [PubMed: 22388958]
- Koch C, Poggio T, Torre V. Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proc Natl Acad Sci USA*. 1983; 80:2799–2802. [PubMed: 6573680]

- Larkum M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* 2013; 36:141–151. [PubMed: 23273272]
- Larkum ME, Nevian T, Sandler M, Polsky A, Schiller J. Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. *Science.* 2009; 325:756–760. [PubMed: 19661433]
- Larkum ME, Zhu JJ, Sakmann B. A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature.* 1999; 398:338–341. [PubMed: 10192334]
- Lee I, Yoganarasimha D, Rao G, Knierim JJ. Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature.* 2004; 430:456–459. [PubMed: 15229614]
- Legenstein R, Maass W. Branch-specific plasticity enables self-organization of non-linear computation in single neurons. *J Neurosci.* 2011; 31:10787–10802. [PubMed: 21795531]
- Leutgeb JK, Leutgeb S, Moser MB, Moser EI. Pattern separation in the dentate gyrus and ca3 of the hippocampus. *Science.* 2007; 315:961–966. [PubMed: 17303747]
- Leutgeb S, Leutgeb JK, Treves A, Moser MB, Moser EI. Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science.* 2004; 305:1295–1298. [PubMed: 15272123]
- Lisman JE. Relating hippocampal circuitry to function: recall of memory sequences by reciprocal dentate-CA3 interactions. *Neuron.* 1999; 22:233–242. [PubMed: 10069330]
- Lisman JE, Otmakhova NA. Storage, recall, and novelty detection of sequences by the hippocampus: elaborating on the socratic model to account for normal and aberrant effects of dopamine. *Hippocampus.* 2001; 11:551–568. [PubMed: 11732708]
- Losonczy A, Magee JC. Integrative properties of radial oblique dendrites in hippocampal CA1 pyramidal neurons. *Neuron.* 2006; 50:291–307. [PubMed: 16630839]
- Makara JK, Magee JC. Variable dendritic integration in hippocampal CA3 pyramidal neurons. *Neuron.* 2013; 80:1438–1450. [PubMed: 24360546]
- Mankin EA, Sparks FT, Slayyeh B, Sutherland RJ, Leutgeb S, Leutgeb JK. Neuronal code for extended time in the hippocampus. *Proc Natl Acad Sci USA.* 2012; 109:19462–19467. [PubMed: 23132944]
- Marr D. Simple memory: a theory for archicortex. *Philos Trans R Soc Lond B Biol Sci.* 1971; 262:23–81. [PubMed: 4399412]
- McClelland JL, Goddard NH. Considerations arising from a complementary learning systems perspective on hippocampus and neocortex. *Hippocampus.* 1996; 6:654–665. [PubMed: 9034852]
- McClelland JL, McNaughton BL, O'Reilly RC. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev.* 1995; 102:419–457. [PubMed: 7624455]
- McHugh TJ, Jones MW, Quinn JJ, Balthasar N, Coppari R, Elmquist JK, Lowell BB, Fanselow MS, Wilson MA, Tonegawa S. Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. *Science.* 2007; 317:94–99. [PubMed: 17556551]
- Morita K. Possible role of dendritic compartmentalization in the spatial working memory circuit. *J Neurosci.* 2008; 28:7699–7724. [PubMed: 18650346]
- Nakashiba T, Young JZ, McHugh TJ, Buhl DL, Tonegawa S. Transgenic inhibition of synaptic transmission reveals role of CA3 output in hippocampal learning. *Science.* 2008; 319:1260–1264. [PubMed: 18218862]
- Nakazawa K, Quirk MC, Chitwood RA, Watanabe M, Yeckel MF, Sun LD, Kato A, Carr CA, Johnston D, Wilson MA, Tonegawa S. Requirement for hippocampal CA3 NMDA receptors in associative memory recall. *Science.* 2002; 297:211–218. [PubMed: 12040087]
- Neunuebel JP, Knierim JJ. CA3 retrieves coherent representations from degraded input: Direct evidence for CA3 pattern completion and dentate gyrus pattern separation. *Neuron.* 2014; 81:416–427. [PubMed: 24462102]
- Poirazi P, Brannon T, Mel BW. Arithmetic of subthreshold synaptic summation in a model CA1 pyramidal cell. *Neuron.* 2003a; 37:977–987. [PubMed: 12670426]
- Poirazi P, Brannon T, Mel BW. Pyramidal neuron as two-layer neural network. *Neuron.* 2003b; 37:989–999. [PubMed: 12670427]
- Poirazi P, Mel BW. Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron.* 2001; 29:779–796. [PubMed: 11301036]

- Rolls ET, Treves A. Neural networks in the brain involved in memory and recall. *Prog Brain Res.* 1994; 102:335–341. [PubMed: 7800823]
- Sjöström PJ, Häusser M. A cooperative switch determines the sign of synaptic plasticity in distal dendrites of neocortical pyramidal neurons. *Neuron.* 2006; 51:227–238. [PubMed: 16846857]
- Spruston N. Pyramidal neurons: dendritic structure and synaptic integration. *Nat Rev Neurosci.* 2008; 9:206–221. [PubMed: 18270515]
- Spruston N, Jaffe DB, Johnston D. Dendritic attenuation of synaptic potentials and currents: the role of passive membrane properties. *Trends Neurosci.* 1994; 17:161–166. [PubMed: 7517596]
- Squire LR, Zola-Morgan JT. The cognitive neuroscience of human memory since H.M. *Annu Rev Neurosci.* 2011; 34:259–288. [PubMed: 21456960]
- Suh J, Rivest AJ, Nakashiba T, Tominaga T, Tonegawa S. Entorhinal cortex layer III input to the hippocampus is crucial for temporal association memory. *Science.* 2011; 334:1415–1420. [PubMed: 22052975]
- Takahashi H, Magee JC. Pathway interactions and synaptic plasticity in the dendritic tuft regions of CA1 pyramidal neurons. *Neuron.* 2009; 62:102–111. [PubMed: 19376070]
- Treves A, Rolls ET. Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus.* 1992; 2:189–199. [PubMed: 1308182]
- Treves A, Rolls ET. Computational analysis of the role of the hippocampus in memory. *Hippocampus.* 1994; 4:374–391. [PubMed: 7842058]
- Tsay D, Dudman JT, Siegelbaum SA. HCN1 channels constrain synaptically evoked Ca^{2+} spikes in distal dendrites of CA1 pyramidal neurons. *Neuron.* 2007; 56:1076–1089. [PubMed: 18093528]
- Vinogradova OS. Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus.* 2001; 11:578–598. [PubMed: 11732710]
- Wu XE, Mel BW. Capacity-enhancing synaptic learning rules in a medial temporal lobe online learning model. *Neuron.* 2009; 62:31–41. [PubMed: 19376065]
- Xu, N-I; Harnett, MT.; Williams, SR.; Huber, D.; O'Connor, DH.; Svoboda, K.; Magee, JC. Nonlinear dendritic integration of sensory and motor input during an active sensing task. *Nature.* 2012; 492:247–251. [PubMed: 23143335]
- Ziv Y, Burns LD, Cocker ED, Hamel EO, Ghosh KK, Kitch LJ, Gamal AE, Schnitzer MJ. Long-term dynamics of CA1 hippocampal place codes. *Nat Neurosci.* 2013; 16:264–266. [PubMed: 23396101]

Encoding



Recall

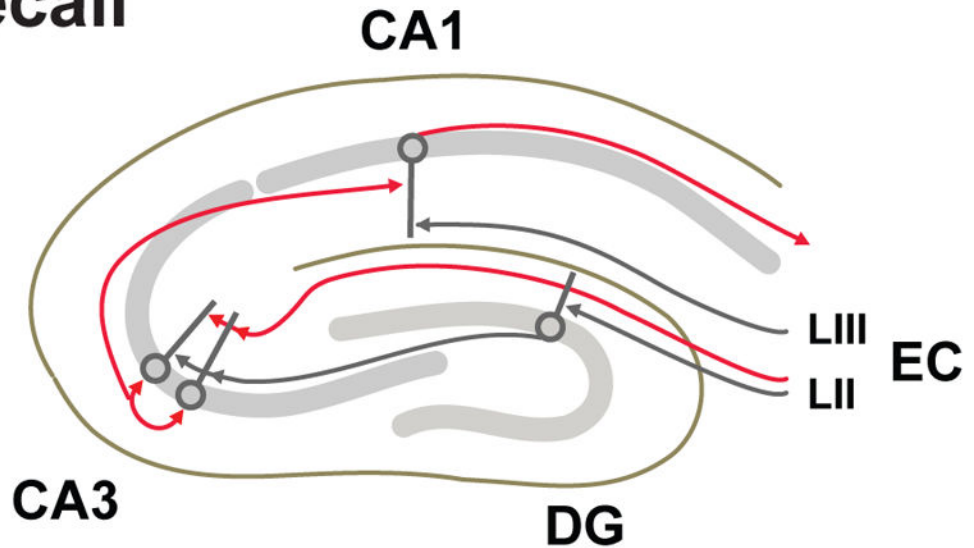


Figure 1. Roles of hippocampal subfields, inputs, and connections

Schematic diagram outlining the roles for specific input pathways and hippocampal subfields during memory encoding and recall phases considered here and by McClelland & Goddard (1996). Pathways with essential roles at each stage are indicated in red or blue, with blue indicating pathways in which synaptic plasticity occurs. *Encoding*: EC and DG inputs to produce a heavily decorrelated CA3 engram, which is stored through plasticity at recurrent connections. Plasticity at synapses between CA3 and CA1, modulated by direct EC inputs, associates the decorrelated CA3 engram with a more information-rich CA1 engram. *Recall*: EC layer II (LII) inputs to CA3 determine which engram will be reactivated

by CA3 recurrent dynamics. The reactivated CA3 engram in turn activates the associated CA1 engram, which provides an informatin-rich hippocampal output representing the stored memory.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

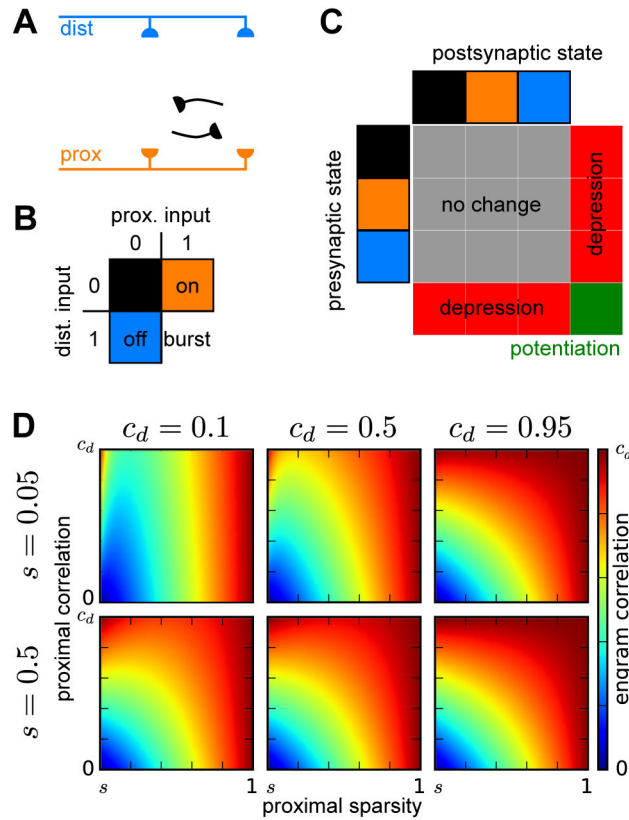


Figure 2. Memory encoding in an attractor network of two-compartment neurons

(A) Model architecture. Two-compartment model neurons receive external inputs to their distal and proximal compartments and recurrent connections to their proximal compartments.

(B) Nonlinear integration rule. Neurons receiving receiving suprathreshold proximal input are active, and those that additionally receive suprathreshold distal input enter a burst-firing state that engages plasticity mechanisms. (C) Dependence of synaptic plasticity on the combined inputs to presynaptic and postsynaptic neurons. Synapses between burst-firing neurons undergo potentiation, while synapses between burst-firing neurons and non-burst-firing neurons undergo depression.

(D) Correlation between engrams created through integration of correlated distal input patterns and less correlated proximal input patterns. Each plot, corresponding to a different pair of values for the total sparsity (s) and correlation between distal activations (c_d), shows how the correlation between engrams, ranging from 0 to c_d , depends on the proximal sparsity, ranging from s to 1, and the correlation between proximal activations, ranging from 0 to c_d .

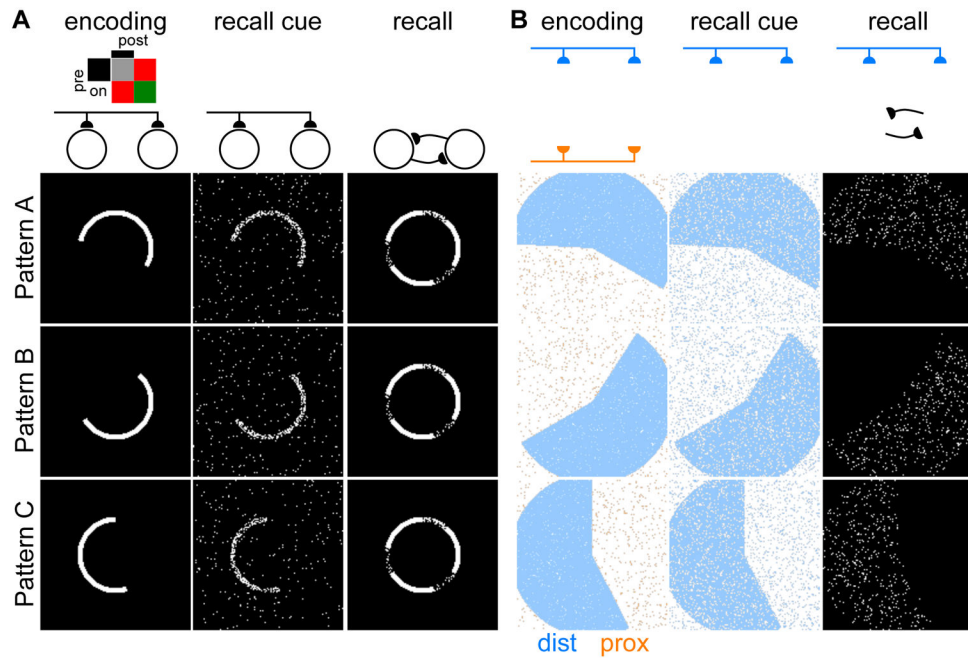


Figure 3. Storage and recall of correlated memories illustrated with pictorial network representations

(A–B) Example storage of correlated patterns (left), followed by presentation of recall cues (center), each resulting in the network converging upon a recall activity pattern (right). Schematic diagrams (above) indicate which inputs are active (dark) or suppressed (faded) at each stage.

(A) The attractor network with a single input pathway consistently converges to a common mixed attractor combining features of all three stored patterns. Inset: Plasticity rule, analogous to Figure 2C, for the case of a single input pathway.

(B) The attractor network of two-compartment neurons maintains separate engrams for each stored pattern. Encoding (left) occurs through combination of distal and proximal inputs, with pixels colored according to input combinations (Figure 2B). Recall cues (center) are provided through the distal inputs. The recalled activity patterns (right) match the patterns of burst-firing neurons during encoding.

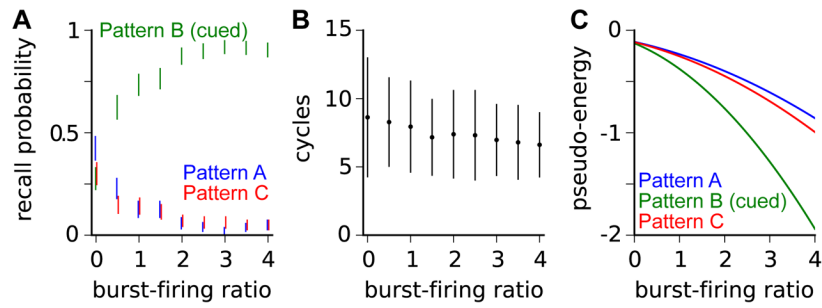


Figure 4. Recall with nonlinear dendritic integration

(A) Probabilities of recall dynamics converging to each engram (Pattern A, blue; Pattern B, green; Pattern C, red) as a function of the burst-firing ratio b . Throughout, the recall cue is a noisy version of the distal input pattern for Pattern B (Figure 3B). Bars indicate 95% confidence intervals estimated with bootstrapping after 250 recall simulations.

(B) The number of asynchronous update cycles required for convergence. Bars indicate mean \pm standard deviation. (C) Pseudo-energies of the three stored engrams as a function of the burst-firing ratio b , with a noisy version of the distal inputs for Pattern B as the recall cue. Since this cue has overlap with all three engrams, the pseudo energy of each engram decreases as the burst-firing ratio increases, and that of the most overlapping pattern decreases most.

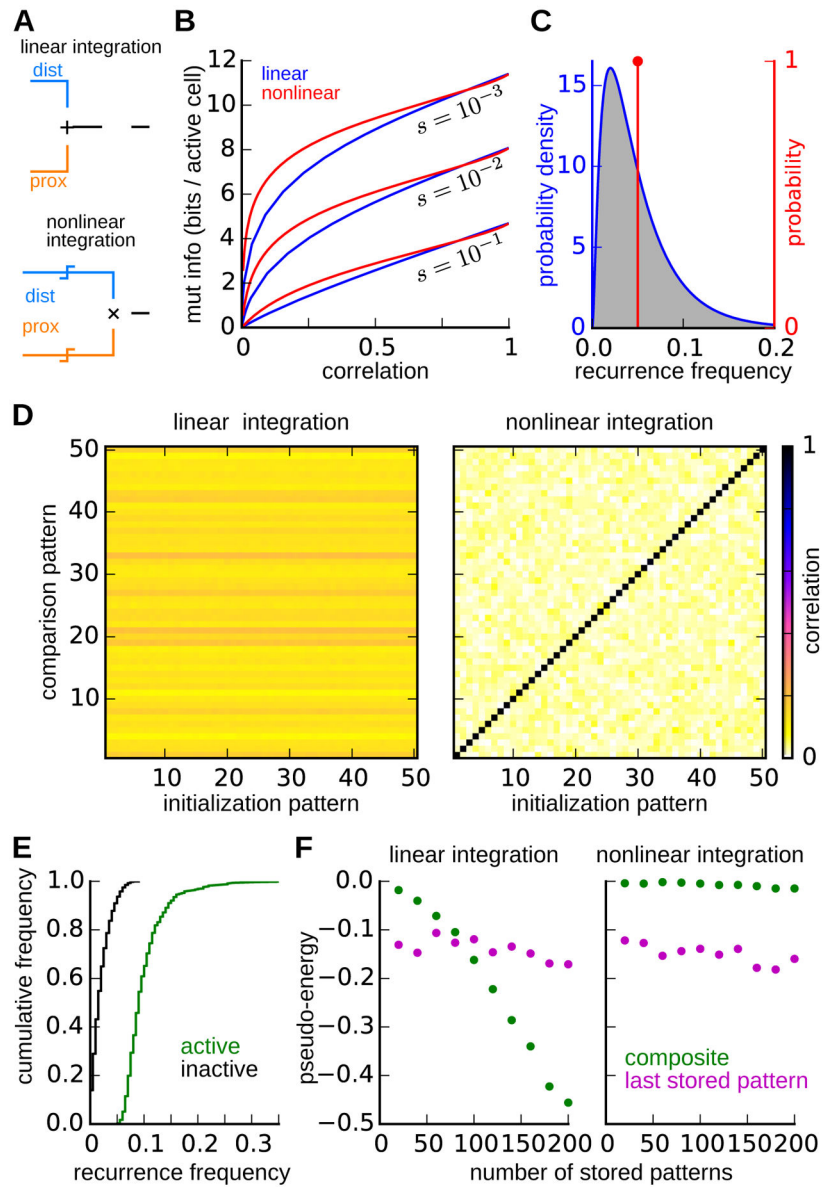


Figure 5. Comparison of linear and nonlinear integration of distal and proximal inputs
 (A) Schematic diagrams of contrasting input integration schemes. For linear integration (top), distal and proximal inputs are summed before thresholding is applied. For nonlinear integration (bottom), the proximal and distal inputs are each thresholded separately, with an AND operation applied to the output of these threshold operations.
 (B) Trade-off between engram decorrelation and information. Parametric curves show the relationship between the correlation between engrams for pairs of highly similar memories, and the mutual information that an engram contains about the distal input pattern at the time of encoding. Separate curves correspond to networks with linear (blue) and nonlinear (red) integration for three different levels of engram sparsity.

(C) Distribution of recurrence frequency – i.e. the fraction of highly similar engrams in which a neuron is active – for neurons in engrams of networks with linear (blue) and nonlinear (red) integration.

(D) Evaluation of engram attractors following the storage of 200 similar (i.e. same distal inputs) memories. The plots show the correlation between the 50 most recently stored engrams (comparison patterns, ordered with most recent first) and the activity state to which the recall dynamics converge when the network is initialized with each of these same engrams (initialization patterns). The network with linear input integration (left) converges into a common attractor independent of the initialization. The correlation of this common attractor with each of the engrams determines the horizontal banding pattern. The network with nonlinear integration (right) maintains distinct attractors for each pattern, as indicated by the high correlations along the diagonal. The two networks were matched with regard to both sparsity and the pairwise correlation between similar engrams.

(E) The recurrence frequency distributions for neurons that are active (red) or inactive (black) in the composite attractor of the linear integration network in (D). The composite attractor consists of the neurons that have been active in the highest fraction of stored patterns.

(F) Pseudo-energies of the most recently stored engram and the composite activity pattern following storage of progressively many similar patterns. In the network with linear integration (left), but not in the network with nonlinear integration (right), the pseudo-energy of the composite pattern falls below that of the most similar stored pattern.

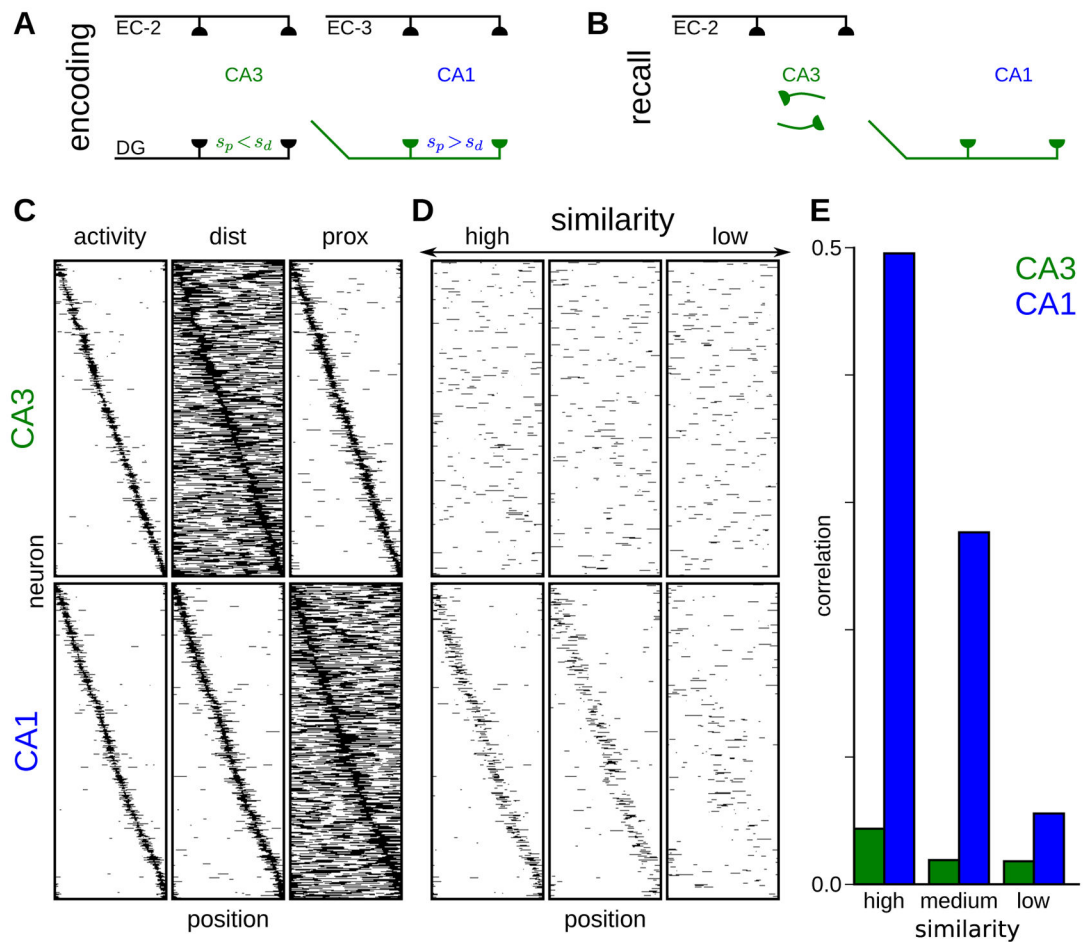


Figure 6. Distinct memory roles correspond to distinct place cell properties in CA3 and CA1 (A–B) Schematic of the combined CA3-CA1 network during encoding (A) and recall (B) phases. The synaptic connections determining activity in each phase are darkly shaded, with other pathways faded.

(A) Encoding phase: Distal and proximal inputs combine to determine burst-firing neurons in both CA3 and CA1, with synapses from burst-firing CA3 neurons to burst-firing or non-burst-firing CA3 and CA1 neurons potentiated or depressed, respectively. For the CA3 network, the proximal sparsity is lower than the distal sparsity, whereas the opposite is true for the CA1 network.

(B) Recall phase. EC layer II (EC-2) inputs promote recall of a similar engram through recurrent activity within CA3. The CA3 engram provides feed-forward input to CA1 and thus reactivates the CA1 ensemble with which it was co-active during encoding.

(C) Distal and proximal contributions to spatial tuning. Simulated CA3 (top) and CA1 (bottom) place cells sorted according to their spatial tuning (left). Spatial tunings of the distal (center) and proximal (right) activations are shown for each cell.

(D) Remapping of spatial tunings across similar environments. Spatial tunings are shown for environments with high (left), medium (center), and low (right) similarity to the original environment (C), with neurons sorted in the same order. CA3 spatial representations remap

almost entirely across all environments, whereas CA1 representations remap progressively with changes to the environment.

(E) Correlations between the spatial tunings in the original environment (C) with those in environments of progressively decreasing similarity (D).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript