Memorization of novel patterns in working memory in a model based on dendritic bistability

Jiacheng Xu,[1*] Daniel L Cox,[1#] Steven J Luck,[2#]

[1]Department of Physics and Astronomy,
University of California Davis, Davis CA 95616 USA
[2]Center for Mind & Brain and Department of Psychology,
University of California Davis, Davis CA 95616 USA

[*]To whom correspondence should be addressed; E-mail: jiacheng.xu@dartmouth.edu
[#]Senior authors, listed alphabetically

Working memory can hold many types of information and is crucial for cognition. Commonly, models of working memory maintain information such as hues or words by forming memory attractors through structured connectivities. However, real-world information can be novel, making it infeasible to use pre-trained attractors. In addition, most models—with or without attractors—have focused on maintaining binary categories instead of continuous information in each neuron. In the present study, we investigate how the brain might maintain working memory representations of arbitrary novel patterns with graded values. We propose an unstructured network model in which each neuron has multiple bistable dendrites. Each dendrite effectively implements fast Hebbian plasticity due to dendritic dynamics and dendrite-soma interactions. This network can maintain novel graded patterns under various perturbations without fine tuning of parameters. Through analytical char-

acterization of network dynamics during the encoding and memory periods, we identify different conditions that yield either perfect memories or several types of memory errors. We also demonstrate memory robustness under various conditions and resilience to temporal inhibitory perturbations. Thus, this architecture provides robust and analytically tractable storage of novel graded patterns in working memory.

Introduction

Working memory can actively maintain a variety of content types. Most previous studies have focused on simple content, such as location, color, shape, or letter identity. But working memory can also hold complex content, such as a newly-defined word whose meaning is based on other learned words, or a beach scene containing both basic elements (sand, cliff, sea or bird) and additional information (the inclined angle of the cliff or the speed of flying birds). The content can also be more abstract, such as a new odor that elicits a distributed firing pattern in the cortex, or the rules of a new task that may be represented in some encrypted pattern of neural firing.

The underlying mechanisms of working memory with complex content are particularly poorly understood. To approach this question, we assume that complex information could be represented by a graded firing pattern, with each neuron maintaining a specific firing rate across a range of possible values. Given the flexibility of working memory content, we explore the mechanism that allows arbitrary graded firing patterns to be maintained. Because there are exponentially many such patterns, it is important that the memory mechanisms can store novel patterns without pre-trained attractors.

However, the majority of working memory models use attractor networks to store information about a predefined and limited space of possibilities. For example, an eye position can be

maintained by the intensity of memory activity [33, 16, 24, 7], and a hue on a color wheel can be maintained by local activity among neurons arranged as a ring [1, 8, 9, 5]. In these models, network connectivities are explicitly structured or partially structured [5] as attractors, allowing stable representations. These models assume that the network connectivity is predefined, although in theory they could be learned via extensive training.

Holding novel information in working memory is important for human cognition. In everyday life, we often encounter novel information and would need such an ability to make flexible decisions. In addition, in cases where attractors are valuable, maintaining novel information may be important for attractor formation. That is, holding the novel pattern in working memory for a period of time may provide the extended time needed for slow attractor formation and long-term storage [20].

There have been limited efforts to model the maintenance of novel information. To perform this task, the system needs to form a stable memory representation in one shot based on a brief signal presentation. However, canonical Hebbian plasticity [20] cannot readily achieve this because the synaptic weight typically changes slowly, requiring a long training period to form attractors. People have therefore proposed fast Hebbian plasticity [23], which is Hebbian plasticity but with faster changes in synaptic weight (within 100 ms). However, two problems can happen with existing fast Hebbian plasticity models [31, 12]. First, the memory only maintains binary activity, discarding any continuous information (e.g., input intensity). Second, experimental evidence for fast Hebbian plasticity is not strong, despite some support [30, 11, 29]. In addition, people have also suggested that NMDA receptors, due to the pre- and postsynaptic voltage dependence of their conductance, can effectively achieve fast Hebbian plasticity through voltage change. However, models based on NMDA receptors still maintain only binary activity [25, 36, 32]. See Discussion for further comparison with other models.

In this paper, we propose a network model of working memory that can encode and maintain

novel patterns with graded firing rates across neurons. The model consists of a homogeneous network where each neuron possesses multiple bistable dendrites, each of which effectively exhibits a form of fast Hebbian plasticity through dendritic activity. These dendrites give each neuron multistability in their firing rates, making it possible to maintain graded values. The same modeling idea can also be implemented for bistable synapses or bistable neurons as discussed below.

The rest of the structure of the paper is as follows: First, we quantify the basic performance of the network and show its ability to robustly maintain information under various perturbations. Next, we explain the detailed network dynamics during the encoding and memory periods. Then we analytically derive the input function for a perfect memory and identify various mechanisms for error. Additionally, we show memory robustness and resilience to a strong, localized external inhibitory input. Finally, we discuss biological plausibility, different model interpretations, and comparisons with other models.

Results

Network structure and basic performance

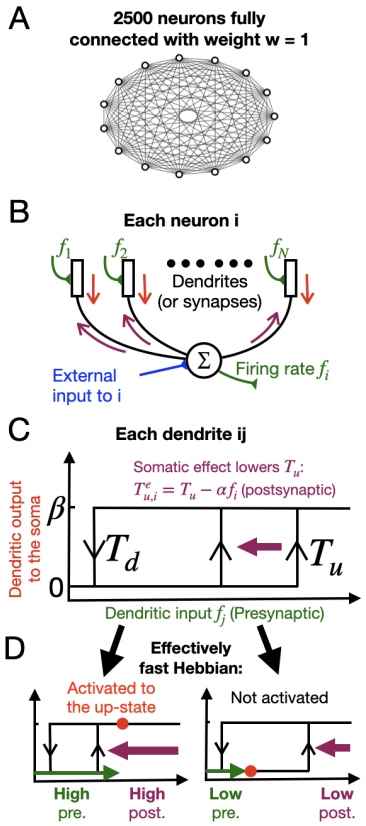

Figure 1: **Network structure.** **A**, Unstructured network. N = 2500 neurons are fully connected with uniform weight w = 1. **B**,Multi-Compartment neuron. Each neuron has N separate dendrites. Each dendrite receives dendritic input from other neurons (with self-connection) and outputs to the soma. The firing rate of the neuron equals the summation of all dendritic outputs and the external input in the soma. It is also possible to implement the model with separate synapses instead of separate dendrites. See Discussion for details about biological plausibility. **C**, Bistable dendrite input-output relationship. The dendritic input and output follow a bistable relation with up- and down-thresholds T_u and T_d. The up-state contributes a factor of β to the somatic firing rate. In addition, the somatic firing f_i can effectively lower T_u. The effective up-threshold is modeled by $T^e_{u,i} = T_u - \alpha f_i$, which qualitatively captures the effect from the voltage backpropagation from the soma to dendrites. A spiking network implementation is provided later. **D**, Dendritic activation is similar to fast Hebbian plasticity. Left, when both pre- and postsynaptic firing rates are high, the dendritic input exceeds the low up-threshold, producing dendritic activation with the stabilization point indicated by the red dot. Right, low pre- and postsynaptic firing rates can not make the dendritic input exceed the up-threshold.

We first introduce the rate-based network model. The network has N neurons, all-to-all connected with weights equal to 1 (Fig. 1A). Each neuron is identical and has N identical dendrites (Fig. 1B). Each dendrite receives a different dendritic input from all other neurons. All dendritic outputs are summed in the soma along with the external input to determine the

firing rate. The key elements of the model are the dendrites, which can be either in an up-state or a down-state. Specifically, they have a bistable input-output relation with up- and down-thresholds T_u, T_d (Fig. 1C). Each dendritic up-state contributes β to the somatic firing, which qualitatively captures the impact of dendritic voltage to the soma. Importantly, the somatic firing f_i of a neuron can lower the effective up-threshold $T^e_{u,i}$ of all its dendrites, which we refer to as the *somatic effect*. It is modeled as a linear relation:

$$T^e_{u,i} = T_u - \alpha f_i \tag{1}$$

where α is the strength factor for the somatic effect. We also require $T^e_{u,i} > T_d$. This bistable dendrite, whose dynamics are determined by two step functions and the somatic effect, qualitatively captures the performance of previous conductance-based models [25, 32] based on NMDA receptors and back-propagated action potentials (bAPs), which is also reproduced in the spiking model below. Note that the spiking model suggests that T_d is much less affected by bAPs, such that it is treated as a constant in the rate model.

These bistable dendrites function as basic associative working memory units. During the encoding period, a dendritic up-state can be activated by the combination of a high postsynaptic firing rate in the soma, which lowers T^e_u, and a high presynaptic input to the dendrite (Fig. 1D). However, low pre- and postsynaptic firing rates cannot activate the dendrite. This effectively acts as fast Hebbian plasticity with binary values, but is based on dendritic activation change instead of actual weight change. During the memory period, as long as the dendritic input remains higher than T_d, a dendritic up-state is maintained.

The firing rate of a neuron is proportional to the number of its up-state dendrites. Thus, although each dendrite stores a binary value, the number of dendrites in the up- or down-state for a given neuron can vary depending on the stimulus presented to the network, allowing an approximately continuous set of firing rates in the soma.

Given the network structure, the dynamics for each neuron is governed by

$$\tau \frac{df_i}{dt} = -f_i + \sum_{j=1}^{N} D(f_i, f_j) + I_i, \tag{2}$$

where I_i is the external input pattern and $D(f_i, f_j)$ is the dendritic bistable relation in Fig. 1C, which is mathematically expressed:

$$B(f_i, f_j) = \begin{cases} \beta\Theta(f_j - T^e_{u,i}) & \text{Activation from down- to up-state} \\ \beta\Theta(f_j - T_d) & \text{Deactivation from up- to down-state} \end{cases} \tag{3}$$

where $\Theta(x)$ is a Heaviside step function.

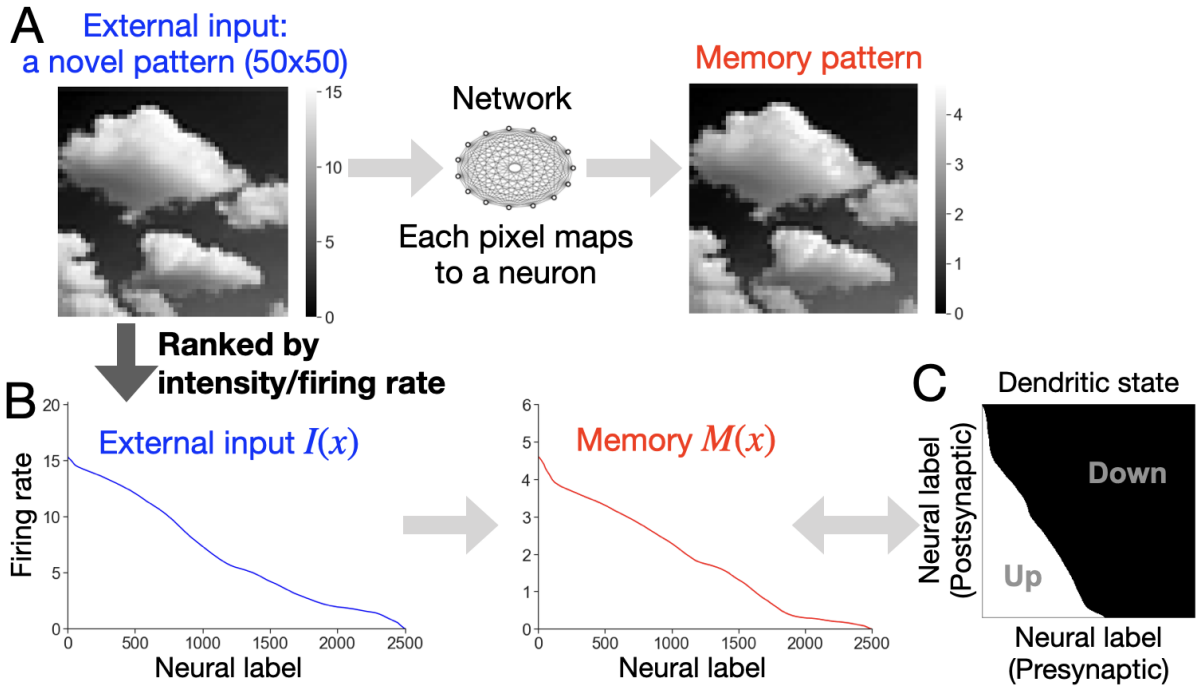

Figure 2: **Memorization of an example novel pattern.** **A**, Network performance for an input with a two-dimensional representation. Left, a graded cloud-like image (50x50) serves as the input. Each pixel maps to a neuron in the network, with its intensity equal to the neural firing rate. Right, when the input is removed, the resulting memory maintains a similar graded pattern despite some small distortion. **B**, Network performance in a one-dimensional representation. The neurons are anatomically identical, so they were re-ordered according to the input firing rate/intensity. In this way, the external input and the memory can be described by two functions, I(x) and M(x), respectively. **C**, Dendritic states when the image is represented in memory. All dendritic states, using white for up and black for down, are plotted for the memory in B. Each individual dendrite in this image is indicated by its pre- and postsynaptic neurons, which are ordered in the same way as in B. Note that the memory activity is proportional to the summed up-states along the vertical axis.

Fig. 2A shows the memory performance of this network for an example novel input pattern. This input pattern is a graded set of varying values in each of 50×50 pixels, which resembles a

cloud. Each pixel maps to a separate neuron within the network, with in total $N = 2500$ neurons. The network starts with all dendrites in the down-state. During the encoding period, the external input strength received by each neuron equals the corresponding pixel intensity. During the memory period with the external input turned off, a stable activity pattern is maintained, with graded firing rates for the neurons that form a pattern similar to the original input with small distortion. See Appendix A for more simulation details.

For ease of analysis, we convert both the external input and the resulting memory into a one-dimensional vector. Because the present network is homogeneous, each neuron is anatomically equivalent, we can relabel the neurons in any arbitrary manner without loss of generality. It turns out that it is convenient to relabel them in descending order according to the external input intensity, resulting in a vector of non-increasing intensity values I_i, which is approximated to be a continuous function $I(x)$ (Fig. 2B). In addition, the network activity during the memory period is also a non-increasing vector $M(x)$ under the same set of reordered neural labels. This is because a larger input leads to a lower T_u^e, which activates more dendrites and induces greater activity during the maintenance period. The detailed dynamics is explained below.

Memory activity is not just represented in firing rates, but equivalently represented in a functional network based on dendritic states. As shown in Fig. 2C, during the memory period, each dendrite (pixel), specified by the label of the neuron it belongs to (postsynaptic) and the neuron it receives a projection from (presynaptic), can be in the up-state (white) with recurrent interaction or in the down-state, effectively disconnected (black). With both the pre- and postsynaptic labels ordered in the same way as in Fig. 2B, Fig. 2C shows that neurons with greater pre- and postsynaptic firing rates (i.e., those toward the left lower side of the plot) have more up-state dendrites between them. For future use, we define in-degree as the number of up-state dendrites a neuron has, and out-degree as the number of up-state dendrites a neuron projected to other neurons. Note that the summation of in-degree for each neuron is proportional to memory firing

rate.

Some clarifications on these results:

1. We chose a large value of N for ease of analysis. While $N = 2500$ separate dendrites may be too large for a real neuron, and a novel pattern with $N = 2500$ entries may also be too hard for human memory, a large N value allows the performance of this discrete network simulation to approach its continuum limit, simplifying analysis. See Discussion for equivalent model implementations at the synaptic or cellular level, which allow for more computational units. We also include a spiking model below with fewer neurons and dendrites.

2. The cloud image used as the input here is just an example of a novel graded pattern, and it was chosen simply because it makes the input and output easy to visualize. We are not implying that this is how a cloud image is actually encoded in the brain, which involves multilayer processing from the photoreceptors of the eye to visual cortex.

3. The minimum external input is set to zero for simplicity. The following results still hold if a nonzero constant I_c adds to the external input, in which case it is equivalent to reset the constant T_u to a lower value $T_u - \alpha I_c$.

4. Note that firing rates are on average higher during the input period than the memory period. In other words, the memory activity loses information about the overall input strength, despite memorizing graded values. This overall firing difference can be captured by a multiplicative scaling factor, and can be memorized by a separate parametric working memory system [21, 6, 16], which is not the focus in this work.

Memory robustness under perturbations

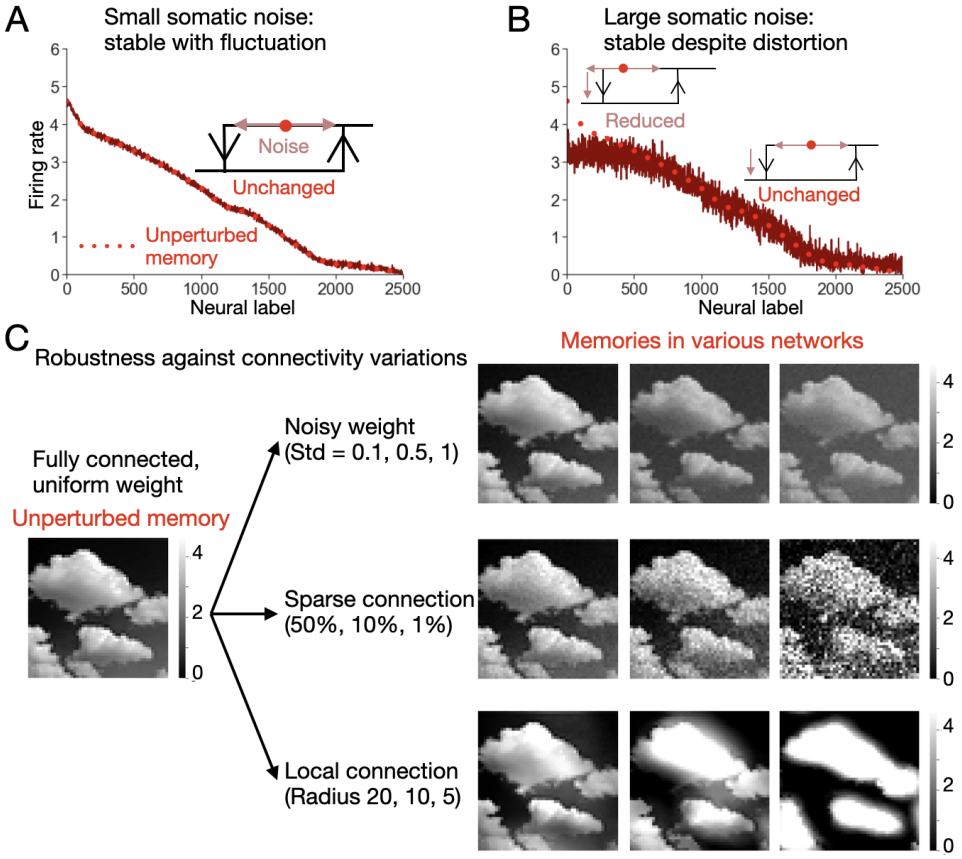

Figure 3: **Memory performance with somatic noise or connectivity variations.** **A**, Network performance in the presence of somatic noise. Gaussian random noise is applied to each soma independently. The dark red line shows the averaged memory over 500 ms. The dotted red line is the unperturbed memory replotted from Fig. 2 for comparison. Inset: A schematic plot illustrating how noise can drive a stable point in a bistable dendrite back and forth along its original stable point. **B**, Same as A but with a larger noise. Noise causes the high firing rate part of the memory to drop, in addition to noisy fluctuation. Inset: Similar to A, but now some dendrites flip to the down-state due to noise. **C**, Network performance under connectivity variations. Three conditions are implemented. Left: The unperturbed memory replotted from Fig. 2. Top: Memories resulting from noisy weight values from a Gaussian distribution of connection strengths with mean 1 and standard deviation 0.1, 0.5, or 1, respectively. Middle: Memories resulting from uniform weight sparse connections with a connection probability of 50%, 10%, or 1%. Bottom: Memories resulting from a connectivity pattern in which connections only exist between local neurons, within a radius of 20, 10, or 5, in a space defined by two-dimensional pixels. The memory is robust for all but the most extreme connectivity modifications. The same graded scale bar is used to help direct comparison. See simulation details in Appendix A.

We next tested whether our model is robust against various perturbations without finely tuned model parameters. In the following, we consider two magnitudes of somatic Gaussian noise during the memory period and alternative connectivity patterns to the network.

We first consider robustness against relatively small levels of somatic Gaussian noise during

the memory period. We found that the pattern of activity across units in the presence of noise (Fig. 3A, dark red line) was nearly identical to the pattern observed without noise (Fig. 3A, dotted red line), despite some random fluctuation. The reason for this robustness against noise lies in dendritic bistability: the noise changes the dendritic input, shifting the stable point (Fig. 3A, inset, red) back and forth, but this small noise is insufficient to drive the stable point across T_u or T_d. Consequently, small noise does not cause dendrites to shift between the up- and down-state, so the input from the dendrites to the soma remains unchanged by small noise.

We next examined the effects of larger somatic noise, which we found produces a more significant distortion of the memory. An example is shown by the dark red line in Fig. 3B. Negative noise fluctuation inactives some dendrites in high-firing neurons from the up-state to the down-state, this causes a reduction in the somatic firing rate, which may further cause dendritic deactivation. But this positive feedback loop does not completely destabilize the memory as the example shown. Meanwhile, the same noise level does not flip the dendritic states of low-firing neurons. Altogether, although large noise did cause some distortion of the memory, the distortion was rather modest despite the fact that the network was not fine tuned.

Why does the noise induced activity reduction not lead to further recurrent reduction and eventually silencing the entire memory? Mechanistic details are explained in Fig. 4 and Appendix B, but the intuition is simple. Upon the receiving of a recurrent activity reduction, the output of a dendrite may be unaffected, as long as the reduction is not strong enough to drive this dendrite to deactivation by crossing T_d. Therefore, the threat of a positive feedback loop is mitigated by zero feedback. The same idea works for positive noise fluctuation.

Next we examined memory robustness against variations in connectivity. Three common variations were considered (Fig. 3C): random weight values in a uniformly connected network (top), sparse connectivity (middle), and local connectivity, where connections only exist between neurons that are within a specific distance in two-dimensional space (bottom). For visual

clarity, memories are presented as two-dimensional pixel arrays. In each case, the network robustly maintained a novel pattern similar to the unperturbed case (left), unless the variations are extreme. For example, the memory was clearly degraded when the standard deviation of weights approached its mean value 1 (top, rightmost); when the connection probability dropped to 1

Detailed network dynamics during the encoding and the memory periods

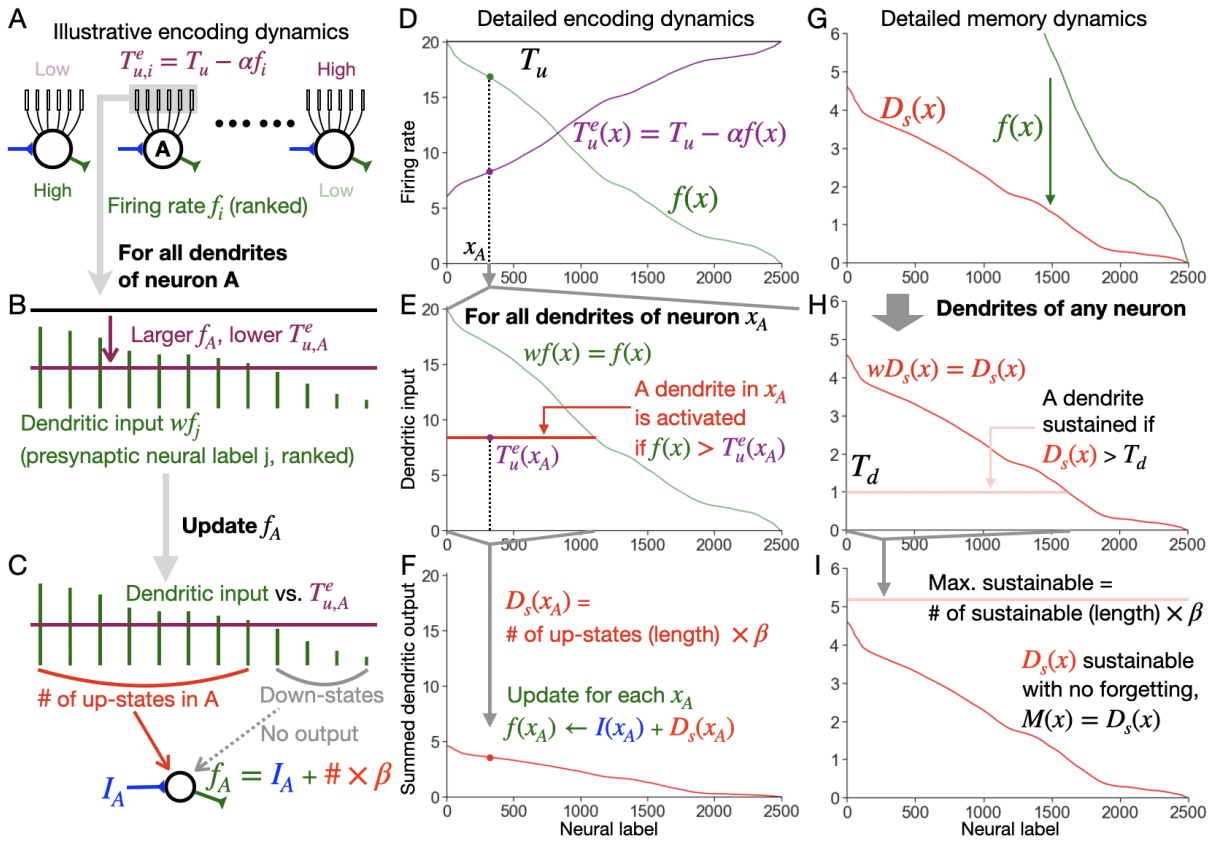

Figure 4: **Detailed dynamics during the encoding and memory periods.** A-C, Illustrative plots of encoding dynamics. **A**, Network neurons ranked by external input intensity. A neuron i with a high firing rate has a low $T^e_{u,i}$ due to the somatic effect. **B**, All dendrites of a neuron A. Dendrites are ordered by the intensity of their presynaptic neurons. Dendrites receive dendritic inputs of graded values (green sticks) through uniform weight $w = 1$. The somatic effect uniformly reduces T_u to $T^e_{u,A}$ (purple bar) of all dendrites. **C**, Dendritic activation. If a dendritic input exceeds $T^e_{u,A}$, the dendrite is activated. Each activation contributes β to the somatic firing. The soma sums an external input (blue) with all up-state dendrites. Down-state dendrites have no effect. **D-F**, Detailed encoding dynamics, drawn to illustrate Eq. 5. **D**, The green curve indicates the firing rate f(x) at a particular time step during the encoding period. f(x) sums external input I(x) and the summed dendritic output $D_s(x)$. The purple curve indicates the effective up-threshold $T^e_u(x)$. The particular f(x) plotted here is the stable activity at the end of the encoding period, while the dynamics shown are generally true. **E**, For a particular neuron x_A, all its dendritic inputs, with uniform weight $w = 1$, are plotted. Dendrites are ranked by the same input intensity order as in D. All dendrites in x_A share a common $T^e_u(x_A)$. The portion of dendrites activated is shown by the red line. **F**, $D_s(x_A)$ is updated to a new value, the number of activated dendrites times β. This new $D_s(x_A)$ is fed back to A to increase f(x) iteratively in time for each neuron A, until the system stabilizes. **G-I**, Detailed memory dynamics, drawn to illustrate Eq. 6. **G**, After I(x) is turned off, $D_s(x)$ remains. Note that the y-axis is rescaled from F. **H**, While T^e_u depends on the somatic effect, the down-threshold T_d is fixed for any neuron. The pink line shows the portion of dendrites with dendritic inputs above T_d. **I**, If the summed dendritic output $D_s(x)$ falls below the maximally sustainable activity (pink), the existing summed dendritic output $D_s(x)$ is sustainable such that $D_s(x)$ becomes M(x). The maximally sustainable activity is given by the number of sustainable dendrites times β. See details in the text.

How are graded patterns actually memorized by this network? In this section, we provide a detailed breakdown of the network dynamics during the encoding and memory periods. Fig. 4A-C is designed to help provide an intuition for the dynamics during the encoding period.

As in Fig. 3, neurons in Fig. 4A are ranked in descending order on the basis of their input intensities. All dendrites of a particular neuron A (Fig. 4B) receive graded dendritic inputs recurrently and share the same $T^e_{u,A}$ determined by the somatic firing rate f_A. A dendrite is activated if its dendritic input exceeds $T^e_{u,A}$. The summed dendritic output is given by counting all activated dendrites and multiplying by the up-state contribution factor β (Fig. 4C). The soma then sums the summed dendritic output with the external input to get an updated firing rate, which is fed back to f_A in Fig. 4A for iteration. Overall, each neuron receives identical but graded dendritic inputs due to the uniform connectivity. What differentiates the activation in each neuron is the somatic effect. For example, a neuron A with a higher firing rate has a lower $T^e_{u,A}$ for all its dendrites, allowing more dendrites to activate. This eventually leads to a stable but graded firing rate.

Next, we formally describe the encoding dynamics in the continuous limit. Because firing rates are non-decreasing in time, only the T_u branch of the dendritic bistability gets involved, and we can safely ignore T_d for now. In Fig. 4D, the firing rate (green), $f(x) = I(x) + D_s(x)$, sums the external input $I(x)$ and the summed output of all dendrites $D_s(x)$. For each neuron x, the T^e_u value (purple) is lowered by the somatic effect: $T^e_u(x) = T_u - \alpha f(x)$. Next, in Fig. 4E, all dendrites of a neuron x_A are shown, with each dendrite is specified by its presynaptic neural label. As weight w is 1, the dendritic inputs follow $wf(x) = f(x)$. If a dendritic input exceeds the effective up-threshold, $f(x) > T^e_u(x_A)$, the dendrite is activated. Because f(x) is non-increasing, the number of up-state dendrites equals the length of the solid red line:

$$f^{-1}(T^e_u(x_A)) = f^{-1}(T_u - \alpha f(x_A)) \tag{4}$$

where we assume the existence of the inverse function $f^{-1}(x)$. Next, in Fig. 4F, the updated $D_s(x)$, with $D_s(A) = \beta f^{-1}(T^e_u(x_A))$ for each choice of A, is obtained and fed back to Fig. 4D

iteratively until the system is stabilized, giving an overall dynamical equation:

$$\tau \frac{df(x)}{dt} = \beta f^{-1}(T_u - \alpha f(x)) + I(x) - f(x) \tag{5}$$

with a time constant τ to capture the slowest dynamic timescale.

Following that, we describe the dynamics during the memory period. I(x) is turned off, reducing f(x) to $D_s(x)$ (Fig. 4G). But can $D_s(x)$ self-sustain or will it decay to zero? To answer this question, in Fig. 4H, we zoom in on all dendrites of a neuron x_A. An activated dendrite can still be sustained if its dendritic input exceeds T_d, $D_s(x) > T_d$. We want to check that all activated dendrites satisfy this condition. Given the non-increasing nature of $D_s(x)$, the number of sustainable dendrites equals the length of the pink line, $D_s^{-1}(T_d)$. If the length exceeds the number of actually activated up-state dendrites for each neuron, $D_s^{-1}(T_d) > D_s(x)/\beta$, all activated up-state dendrites are preserved with the memory $M(x) = D_s(x)$. The same idea is plotted in Fig. 4I, but by comparing between the maximum sustainable activity (pink line) and the summed dendritic output (red line).

More generally, memory forgetting can happen, such that dendrites, activated during the encoding period, are no longer sustainable due to the lack of external input and fall back to the down-state. The number of actual activated up-state dendrites $D_s(x)/\beta$ should never exceed the number of sustainable dendrites (the up-limit value), $D_s^{-1}(T_d)$. If it is not the case, forgetting happens and some up-state dendrites are not sustainable. The dynamics follows:

$$\tau \frac{df(x)}{dt} = \beta \mathrm{Min}[D_s^{-1}(T_d), D_s(x)/\beta] - f(x) \tag{6}$$

where the memory M(x) is given by the stabilized f(x). We primarily focus on the performance when there is no forgetting and partial forgetting is further explained in Appendix B.

Finally, to better understand dynamics, it is important to identify two functionally separated groups of neurons. There exist some neurons whose firing rates are so low that, even with external input, they cannot produce an up-state in the dendrites they project to. They form a

noninteractive group, with no out-degrees activating dendrites of other neurons. Meanwhile, the rest form an *interactive group*, with nonzero out-degrees. These two groups of neurons can be seen in the example shown in Fig. 2C, where the out-degree of a neuron decreases from left to right along the x-axis, until it hits zero, after which neurons belong to the noninteractive group.

Perfect memory and functional separation of two groups

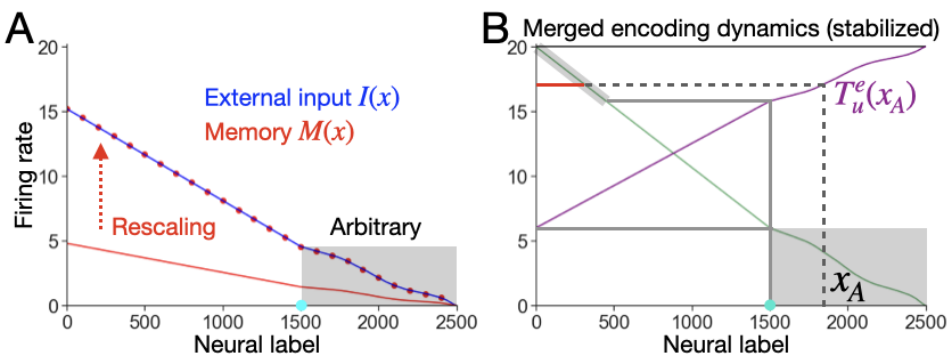

Figure 5: **Perfect memory up to a multiplicative scaling value. A**, Basic network performance. A particular form of external input I(x) gives a perfect memory up to a multiplicative scaling value. With a given cyan dot, I(x) (blue) needs to be linear for neurons at the left side of it, while I(x) can be arbitrary at the right side of it (gray box). The resulting memory (red) is proportional to I(x), as it perfectly overlaps with I(x) after rescaling (red dotted line). The cyan dot is graphically determined in B. **B**, Dynamics during the encoding period for the perfect memory. This plot merges mechanisms in Fig. 4D-F to a single one. The particular f(x) plotted is still the stable activity at the end of the encoding period. See details in the text.

With the dynamics in Equations Encoding and Memory, we can show that the network is capable of a perfect memory under a specific set of conditions. Although biological working memory systems are not typically capable of perfect memory, the present analysis is useful for understanding the nature of the present network. By perfect we mean a memory is proportional to the input by only a multiplicative scaling factor. $M(x) \propto I(x)$, or $M(x) = D_s(x) \propto I(x) \propto f(x)$. This requires a completely stable firing rate during the encoding period to satisfy:

$$\beta f^{-1}(T_u - \alpha f(x)) = D_s(x) = \frac{\beta A}{T_u} f(x), \tag{7}$$

where $\frac{\beta A}{T_u}$ is the proportionality ratio with an arbitrary neural label A between 0 and N. The choice of this particular form will be clear below. Further, to ensure no forgetting happens, all activated dendrites should be sustainable. The perfect memory solution, based of detailed algebra provided in Appendix C, is:

$$I(x) = T_u - \beta A - (\frac{T_u \alpha}{A} - \alpha\beta)x, \text{ with x between 0 and A} \tag{8}$$

$$T_u \leq (1 + \alpha)(T_u - \beta A), \tag{9}$$

$$T_d \leq (1 - \alpha)\beta A, \tag{10}$$

where the non-negativity in I(x) requires $1 \geq \alpha > 0$. The second line ensures that input is large enough to activate some dendrites, and the third line ensures that no forgetting happens during the memory period. To achieve a perfect memory, the external input I(x) needs to be partially tuned. For x between 0 and A, I(x) needs to be linear; for x between A and N, I(x) can be arbitrary. A particular I(x) is given in Fig. 5A alongside the resulting memory M(x). The chosen A is marked by the cyan dot. M(x) perfectly matches I(x) after rescaling, as indicated by the dotted red line. This holds true for arbitrary I(x) values within the gray box. See Appendix A for more simulation details.

Why does this particular form of I(x) lead to perfect memory? Recall that neurons during the memory period can be classified into two groups: an interactive group with neurons having nonzero out-degrees, and a noninteractive group with neurons having no out-degrees; these groups are separated by the particular choice of A here, as the cyan dot shown in Fig. 5. Neurons in the interactive group, from 0 to A, have high firing rates and can affect other neurons by activating their dendrites (out-degree). To prevent such unwanted interaction, inputs to this group must be well-tuned to follow a particular linear pattern. In contrast, for neurons in the noninteractive group, from A to N, their firing rates during the encoding period are lower than the lowest T_u^e at $x = 0$ (Fig. 5B). Therefore, they cannot activate any dendrites and have no

out-degree to affect others, making inputs to them unconstrained. Note that Fig. 5B merges all three panels from Fig. 4D-F into a single panel.

Different errors formed during the encoding period

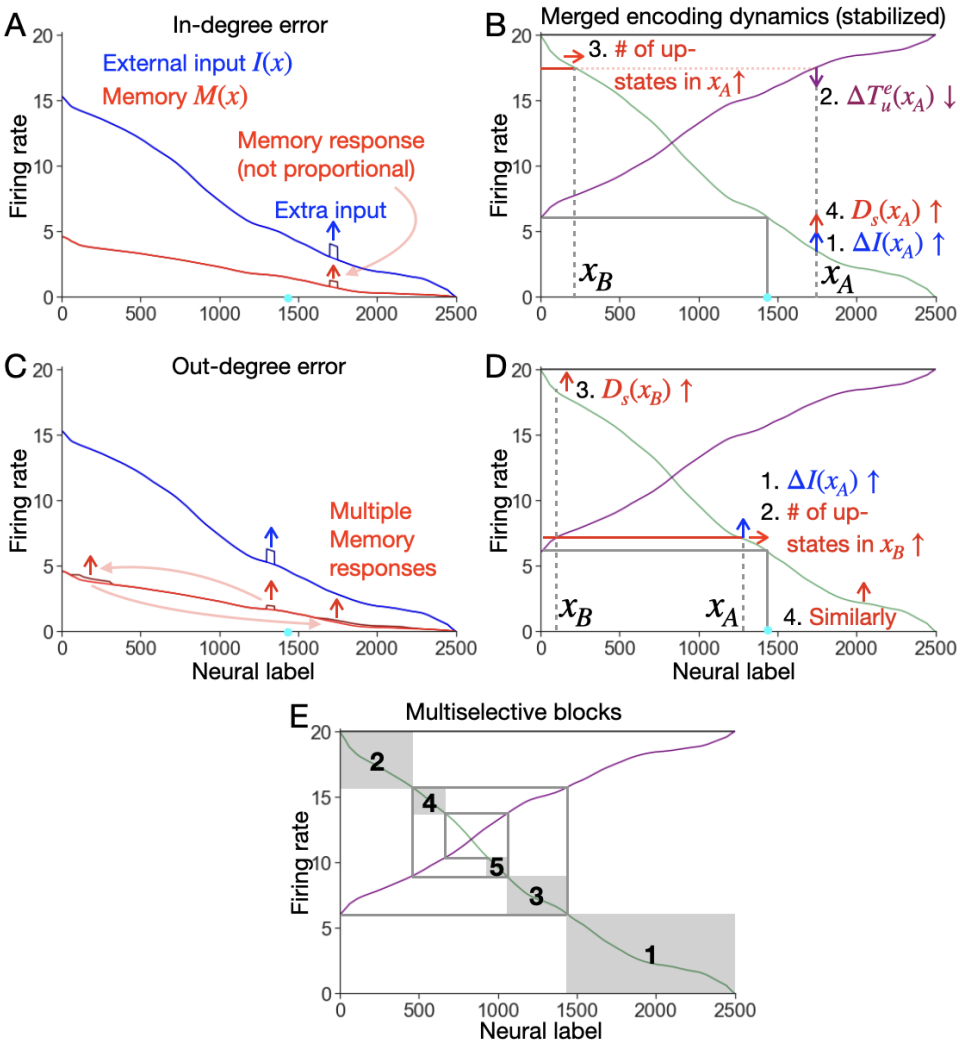

Figure 6: **Memory distortion due to in-degree and out-degree errors.** The cyan dot separates the interactive group (left) from the noninteractive group (right). **A**, With an extra local input in the noninteractive group, the memory response is local, but its magnitude is not linearly dependent on the strength of extra input, illustrating the in-degree error. **B**, Merged encoding dynamics of A. This demonstrates the mechanism underlying the in-degree error. See details in the text. **C**, With an extra local input in the interactive group, the memory response exists in three locations, as shown by the red arrows, illustrating the out-degree error. **D**, Merged encoding dynamics of C. This demonstrates the mechanism underlying the out-degree error. See details in the text. **E**, Illustration of how memory distortion leads to multiselectivity. The increase of activity in one neuron, during the encoding period, can make other neurons fire more. Neurons are functionally separated into different regions, with region 1 being just the noninteractive group. These groups are determined by the spiraling gray line starting from the point of the smallest T_u^e value. See details in the text.

If I(x) does not follow the form in Eq. 8, a graded memory can still be maintained, but with distortion. In this section, we examine several types of errors that originate from the encoding period. For errors due to forgetting during the memory period, see Appendix B.

In-degree error, out-degree error, and multiselectivity

Ideally, a local increase in the external input should induce a memory increase at the same location, with a proportional magnitude. If this is true, any graded input can be memorized perfectly, up to a multiplicative scaling factor, as it just consists of an aggregation of numerous small input increases. However, this is not always true for this network. In general, the memory increase is not proportional to the input increase (*in-degree error*), and may exist in multiple locations (*out-degree error*). We show the mechanisms causing two errors as well as their interaction below.

Let's look at in-degree error first. Fig. 6A shows an example of how a local input increase leads to extra memory response. However, this response is not proportional to the magnitude of input increase, as the mechanism explained in Fig. 6B with merged encoding dynamics. In Fig. 6B, we consider the input increase to be infinitesimal, so that we can ignore the width and strength of a specific extra local input. Based on previously shown dynamics (Fig. 4D-F), the number of activated dendrites (in-degree) in neuron x_A is proportional to the length of the red line (Fig. 6B). An extra input to x_A lowers $T_u^e(x_A)$, increases in-degree of x_A, and thus $M(x_A)$. This increase depends on $f(x_B)$, the slope of f(x) at x_B. This property is problematic, as now the memory response is determined by a nonlocal slope of firing rates of other neurons, causing the in-degree error.

While in-degree error describes how other neurons can affect the activity of a given neuron, out-degree error describes how a given neuron can affect other neurons. Different from in-degree error, which can exist for nonzero memory activity, out-degree error only exists in the

interactive group, where neurons have out-degrees. An example is shown to illustrate out-degree error in Fig. 6C, where a local extra input to the interactive group causes memory changes in three locations. The mechanism is explained in Fig. 6D. For a neuron x_B, it receives dendritic inputs, some above $T_u^e(x_B)$ and some below. There can exist a dendrite whose dendritic input is right below $T_u^e(x_B)$. We denote the presynaptic neuron of this dendrite as x_A. An increase in $f(x_A)$ can activate the corresponding dendrite in x_B, inducing a memory increase in x_B. Similarly, the increase in x_B can trigger a nonlocal memory increase in another neuron through the corresponding dendrite. It sequentially induces nonlocal memory increases until it stops at a neuron in the noninteractive group, where no out-degree exists. For all three locations illustrated, each still has the in-degree error, in parallel with the out-degree error.

In-degree and out-degree errors give a particular form of multiselectivity. With a horizontal line starting from the y-axis with the lowest T_u^e, the firing rate during the encoding dynamics in Fig. 6E can be separated into different regions, gray boxes, as guided by the spiraling gray line. Each gray box includes a portion of neurons (x-axis) depending on their firing rates. An input increase in region 3, as the example shown in Fig. 6D, gives three memory responses, located in region 3, 2, and 1 respectively. More generally, a neuron in a region R can influence in total R regions, $R, R-1, \cdots, 1$, while can also be susceptible to the input increase in region $R+1, R+2, \cdots$. Further, for a neuron with a smaller region label, it is less influential to other regions but more susceptible under a small extra input increase (e.g. from noise). For neurons with a larger region label, they are more influential but less susceptible in the sense that they can only be influenced by higher numbered regions. Therefore, multiselective neurons exhibit a level of anticorrelation in influential and susceptible levels, and the multiselectivity level depends on firing rate of each neuron. Such properties can potentially be tested in the brain by perturbing neural activities.

With in-degree and out-degree errors introduced, we revisit the perfect memory case in

Fig. 5B for more understanding. For firing rates within the noninteractive group (shaded box), the memory change of any neuron x_A, is proportional to $f'(x_B)$, with common $f(x_B)$ indicated by the thick shaded line. In the case $f(x_B)$ is linear, $f'(x_B)$ is a constant such that induced memory change is linear without nonlocal dependence of x_B, making perfect memory possible. In addition, activity in the noninteractive group causes no out-degree error. To support the linear slope region, neurons in the interactive group thus need to be tuned linearly.

Encoding failure and saturation

There exists another source of error related to the overall scaling of input. In Fig. 2B, the multiplicative scaling factor of the input pattern is tuned such that the memory is graded across the entire range of neurons. However, if the multiplicative scaling factor is reduced, with the overall shape unchanged, the high-firing portion of the input can still be encoded, but the low-firing portion becomes too weak to have activated dendrites. This gives *encoding failure* with zero memory in the low-firing neurons (Fig. S2A). If the multiplicative scaling factor is reduced further, the input pattern may be so small that no dendrites are activated, a total encoding failure. To avoid total encoding failure, we require:

$$(1 + \alpha)I(0) \geq T_u \tag{11}$$

where I(0) denotes the peak of the input pattern. This condition guarantees that the input activates at least one dendrite, the one that projects from neuron 0 back to itself, having the largest pre- and postsynaptic firing rate I(0).

On the other hand, if we scale up the input pattern used in Fig. 2B, the memory may begin to saturate. The saturation can happen in two ways. One is *out-degree saturation*, where a high-firing neuron activates all dendrites it projects to. It causes all projected neurons to have at least one activated dendrite, leading to a constant background firing (Fig. S2D). The other is in-degree saturation, where high-firing neurons have activated all possible dendrites, giving

a flat plateau of rates (Fig. S2G). For a more comprehensive treatment of encoding failure or saturation, see Appendix D.

Runaway and insensitivity

The interaction between neurons can form a positive feedback loop that may destabilize the system, causing *runaway error*. An example is shown in Fig. 3B, where a negative noise fluctuation may cause recurrent reduction in activity. Here we provide a semi-analytical treatment describing the conditions under which this may occur.

To see how the encoding dynamics lead to runaway error, we revisit Fig. 6B with a local infinitesimal increase, $\delta f(x_A)$, at x_A. This lowers $T_u^e(x_A)$ by $\alpha\delta f(x_A)$, and expands the number of up-state dendrites by $-\frac{\alpha}{f'(x_B)}\delta f(x_A)$. Therefore, the induced change in the summed dendritic output is:

$$\delta D_s(x_A) \leftarrow -\frac{\alpha\beta}{f'(x_B)}\delta f(x_A), \tag{12}$$

where $\delta D_s(x_A)$ feeds back into $\delta f(x_A)$, leading to further $D_s(x_A)$ changes. Runaway error happens when the gain factor, $-\frac{\alpha\beta}{f'(x_B)}$, is larger than one, in which case the positive feedback blows the activity up. To prevent runaway error, we need:

$$-\frac{\alpha\beta}{f'(x_B)} < 1. \tag{13}$$

As $f'(x_B) = I'(x_B) + D_s'(x_B)$, we have $\alpha\beta < |I'(x_B)| + |D_s'(x_B)|$. We get the sufficient condition:

$$\alpha\beta < |I'(x_B)|. \tag{14}$$

This condition applies to all possible x_B. As the choice of x_A includes all neurons, x_B, which is geometrically determined from x_A, includes all neurons in the interactive group. For the example in Fig. 6, the input satisfies this condition, and the memory has no runaway error. Runaway error, if it happens, does not completely destroy the graded memory, as in the example shown in Fig. S3A.

On the other hand, we want $D_s(x)$ to be sensitive to a small increase. That is, $\delta f(x_A)$ leads to nonzero $\delta D_s(x_A)$. This requires no sudden drop in $|I(x_B)|$, or $|I'(x_B)| < \infty$, such that the gain factor is nonzero. Otherwise, *insensitivity error* happens, where part of the memory gives a flat firing rate, insensitive to the graded input, as the example shown in Fig. S3C. See Appendix E for more details about runaway and insensitivity errors.

Memory surviving from a strong local inhibition

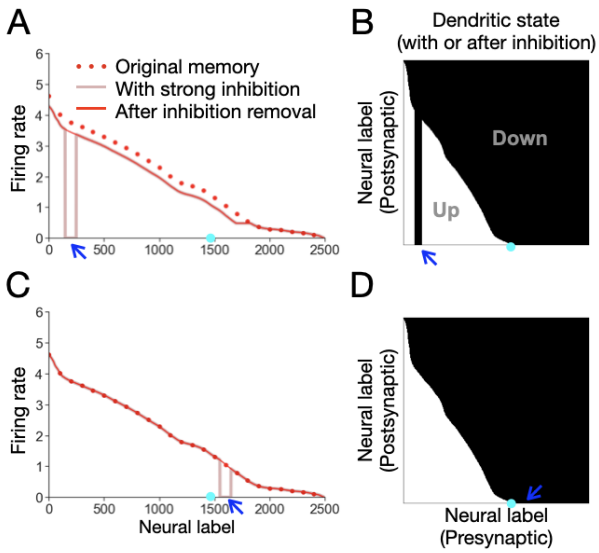

Figure 7: **Memory performance when a subset of neurons are briefly silenced by inhibition.** The cyan dot separates the interactive group (left) from the noninteractive group (right). **A**, Memory degradation under a strong local inhibition. With the originally maintained memory in Fig. 2B (dotted red line), a strong local inhibition is applied to the interactive group of neurons. The inhibition, after stabilization, kills the local activity (blue arrow) and reduces activity in other neurons (pink line). After its removal, the activity for inhibited neurons bounces back and restabilizes (red line). **B**, Dendritic state plot for the memory in A, during the presence of the inhibition or after its removal. The inhibited neurons (blue arrow) cannot support the activation of dendrites in other neurons, having no out-degrees. The memory activity in A is proportional to the summed up-states along the vertical axis. **C, D**, similar to A, B, but with the inhibition applied to the noninteractive group of neurons. The inhibition does not reduce activity in other neurons, and the memory is perfectly recovered after the inhibition removal.

In Fig. 3A, B, We have shown the stability of the network in the face of input noise and connectivity variability. Here, we consider another situation, where a subset of neurons are silenced temporally. From a physiological perspective, such an inhibition might be performed by optogenetic silencing of a portion of a neuron population. From a cognitive perspective, people may need to temporarily silence one working memory representation to focus on a dif-

ferent task, and research shows that they can reactivate the temporarily silenced memory later. Therefore, it is interesting to see how a maintained memory in our model reacts to a temporally applied strong local inhibition to the soma.

Upon the application and removal of a strong local inhibition to the soma, memory can recover. An example is shown in Fig. 7A with the original memory (dotted red line) reproduced from Fig. 2B. The inhibition silences some high-firing neurons (blue arrow) belonging to the interactive group. After stabilization, this inhibition does not erase the memory completely but reduces the activity of other neurons. This reduction is explained in the dendritic states. Because the inhibited neurons have zero activity, their out-degrees go to zero (Fig. 7B, blue arrow), and previous up-state dendrites (Fig. 2C) drop to down-states. This deactivation leads to a reduction of memory activity beyond just the inhibited neurons because the other neurons lose some of their up-state dendrites from the inhibited neurons (Fig. 7A, pink line). Notably, upon the removal of the inhibition, the memory in the inhibited neurons restabilizes to a high level (Fig. 7A, red line). This is because up-state dendrites still exist in the inhibited neurons, supported by dendritic inputs from other neurons and not directly affected by the strong somatic inhibition.

In general, memory degradation and recovery depend upon the in- and out-degrees in inhibited neurons. High firing neurons have more out-degrees, and inhibiting them reduces broader activity among the network. Meanwhile, they also have more in-degree and are more susceptible to dendritic deactivation because of silencing presynaptic neurons. In contrast, if the local inhibition is applied to neurons in the noninteractive group, with no out-degrees, it does not affect other neurons at all, and inhibited neurons are fully recovered (Fig. 7C, D).

The necessity of dendritic bistability for memorizing novel graded patterns

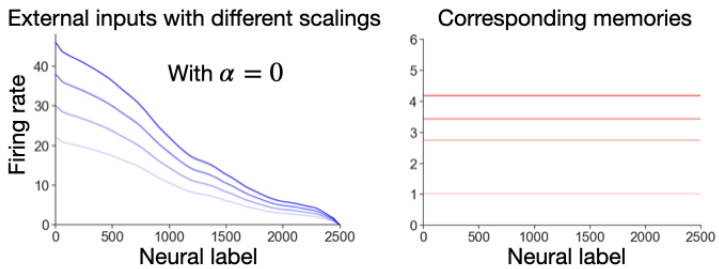

Figure 8: **Memory performance when the somatic effect is absent,** $\alpha = 0$**.** In this case, a novel graded pattern leads to a uniform memory, while different input scaling factors result in different uniform values, matched by corresponding shadeness.

Dendritic bistability is necessary to memorize novel patterns in this network. To demonstrate this, we manipulated either the bistable range or the somatic effect. First, the bistable range was set to zero, yielding a step function with a single threshold T_0. The somatic effect was unchanged. Similar to Eq. 7, the stable firing rate during the memory period in this situation is:

$$\beta M^{-1}(T_0 - \alpha M(x)) = M(x), \tag{15}$$

With similar algebra as performed for Eq. 19, we can solve the equation as $M(x) = T_0 - \alpha\beta x$, with x from 0 to T_0/β. However, this solution is unstable, as $M'(x) = -\alpha\beta$ does not satisfy the stability requirement Eq. 13. Second, if there is no somatic effect, $\alpha = 0$, all neurons are functionally identical. Therefore, the memory activity is uniform across neurons, although the activation level of the memory continues to be scaled by the activation level of the input (Fig. 8).

A spiking network to maintain novel graded patterns

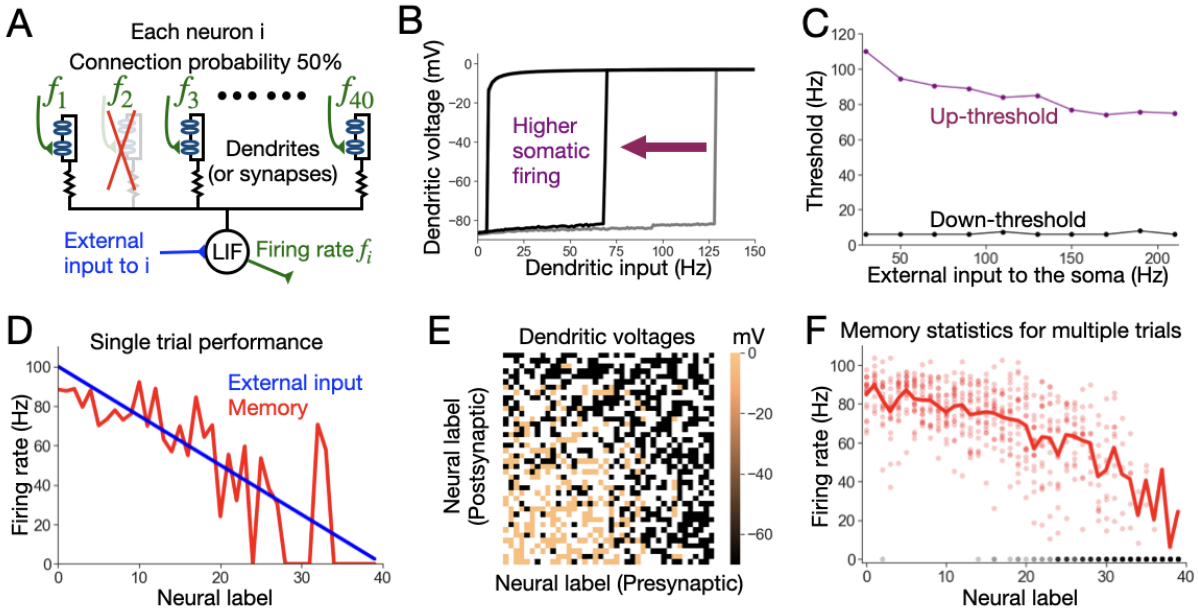

Figure 9: **Spiking network model, structure and performance.** **A**, Network structure. The network has 40 neurons. The connection probability between any two neurons is 50%. Each connection goes into a separate conductance-based dendrite or synapse. The soma operates as a leaky integrate-and-fire way. **B**, Dendritic bistable voltage. A dendrite shows bistable voltage under changing dendrite input frequency. A higher somatic firing can lower its up-threshold. Dendritic voltage is averaged over 500 ms. **C**, Somatic firing and thresholds. Threshold values are recorded as the soma fires more. Up-threshold drops (purple), while the down-threshold remains mostly unchanged (black). **D**, Performance. A linear input (blue) gives a similar but noisy memory (red). Some neurons fail to maintain any activity. **E**, Memory in dendritic voltages. The heat map shows voltages of all dendrites for the memory in D, with white pixels indicating the absence of a connection. **F**, Averaged memory performance. We ran 20 trials and recorded all memories. The red dots represent non-zero memory activity across trials, while the gray dots represent zero memory activity. Lower firing neurons have darker gray dots, indicating a higher probability to have zero memory activity. The red line represents the average of all non-zero memory activities. Graded values are encoded in these non-zero memory activities on average. See Appendix F for details.

In this section, we present a spiking network model which can memorize novel graded patterns as happens in the rate-based model. We built a network of 40 randomly connected neurons, where the probability for a connection is 50%. Each connection projects to a separate dendrite, such that, on average, a neuron has 20 dendrites. Each dendrite is conductance-based with NMDA-receptor dynamics [25, 32], which can show bistability in dendritic voltage (Fig. 9B).

In such a bistable dendrite, somatic firing can lower its up-threshold. As shown in Fig. 9B, when the somatic input of a neuron is increased from 30 Hz to 210 Hz, the up-threshold of one

of its dendrites is reduced. Fig. 9C shows the up-threshold reduction over a range of somatic inputs. Physiologically, the reduction occurs in two steps. Initially, most NMDA receptors remain closed due to the blockage of Mg^{2+}. This makes the conductance of NMDA receptors small such that a large dendritic input is required to open NMDA receptors and activate the dendrite. However, bAPs, from somatic firing, provide the depolarization to remove Mg^{2+} block from the NMDA receptors, facilitating the opening of the NMDA receptors.

The down-threshold behaves differently. The dendrite goes to the down-state below –80 mV if the openness of NMDA receptors is too low to let enough current pass through. However, as the dendritic input frequency drops, the openness of NMDA receptors can still saturate at a high value. Therefore, the down-threshold is typically quite low in our simulation. This can be beneficial as activated dendrites can still be maintained even when firing rates get small. In addition, Fig. 9B,C shows that the down-threshold is insensitive to somatic firing. This is because bAPs do not significantly affect the dendrite. When the dendritic voltage is already high, most NMDA receptors are already unblocked from Mg^{2+} even without the help of bAPs.

This network model can maintain novel graded patterns. Fig. 9D depicts the performance under an example linear input, with the memory maintaining graded values despite noise. Fig. 9E shows the corresponding dendritic voltages during the memory period, which are indeed higher for those dendrites with both their pre- and postsynaptic firing rate high, showing the associative nature as the rate-based model. As the memory in Fig. 9D is still noisy, to better show that memory can be graded, 20 trials were run with all memory activities recorded (Fig. 9F). Non-zero memory activities are shown in red dots, whose average (red line) goes gradually lower, indicating the maintenance of graded values. In addition, zero memory activity, which represents a failure to hold information, is shown in gray dots. Consistent with intuition, for neurons with lower input intensities, memory activity is more likely to be zero, with gray dots going darker. See Appendix F for more details.

Discussion

In this paper, we present a network for the storage of novel graded patterns in working memory. In its simplest form, the network has uniform connectivity. Each neuron contains multiple bistable dendrites, each of which receives an input from one other neuron in the network. Each dendrite is binary, which provides stability, but the number of dendrites in the on-state for a given neuron can vary, allowing the summed voltage in the soma and the neuron's firing rate to vary in a graded manner. The activation of each dendrite is associative, effectively showing fast Hebbian plasticity.

We analyzed the dynamics of the network, demonstrated robustness to noise without fine tuning, showed the conditions necessary for perfect memories, and characterized various errors. Notably, the maintained memory was quite robust under various conditions, including somatic noise, connectivity perturbation, or a strong temporal local inhibition.

Mechanistically, maintaining a novel graded pattern relies critically on local saturation in dendrites. In previous models [31, 12] with fast Hebbian plasticity, upon the receiving of external inputs, the overall activity can easily saturate at the maximum rate. In our model, saturation happens locally in each dendrite, where increasing dendritic input no longer increases dendritic output. This local saturation prevents the positive feedback loop that could lead to the global saturation in overall activity, allowing the system to still remain sensitive to different inputs. A similar concept is discussed in [22].

Biological plausibility and alternative implementations

Dendritic bistability is essential to the present model. One way to achieve this bistability, as shown in the spiking implementation of our model, is to take advantage of the unique dynamics of NMDA receptors, as in previous works [25, 32]. The effect of the somatic voltage on the dendritic thresholds is a second essential component of the present model. A major candidate for

this somatic effect is bAPs [34, 35, 13], as shown in the spiking implementation of our model. But this is not the only possibility, especially when action potentials are missing, or the back-propagation is weak [15]. Other mechanisms may similarly provide the required postsynaptic dependent depolarization. For example, this can arise from subthreshold voltage in the soma or another input current roughly colocalized within the target dendrite [15]. Similar ideas have been tested in [18] for long-term potentiation. Conversely, If bAPs are very strong, dendrites may be highly coupled with the soma [3] with a potential worry that the soma and dendrites will act as a single unit instead of separate units. However, a previous model [4] shows that dendrites can still maintain high independence from the soma.

The spiking implementation of our model included an average of only 20 bistable dendrites per neuron, a realistic but small estimate which gives noisy performance. There are two potential ways to greatly increase the number of bistable units such that the performance is closer to the case in the continuum limit. One way is to use bistable synapses instead of bistable dendrites. Experiments [10, 2] have suggested that each synapse can function as a separate computational unit. A second way is to use bistable neurons in place of bistable dendrites. The dendrites of each neuron in the current model would map to a cluster of neurons, which would bidirectionally connect to another neuron, mapped by the soma in the current model. This implementation would yield a small-network level realization of the currently intracellular structure, where cellular bistability [27, 39] is required instead of dendritic bistability. In all of these cases, the key is to have a set of bistable units, which provide stability, and to combine them so that a graded value can be produced.

Two interpretations of the performance

The multiselectivity shown in network dynamics is a characteristic property of the model and makes neural firing rates correlated. It is rooted in the associative nature of the dendritic

state, which is determined by both pre- and postsynaptic firing rates. During the encoding period, these associative dendrites may bring distortion to the formed memory and show an anticorrelation between susceptibility and influence among neurons. However, the association can also be beneficial, as it helps the recovery of memory after temporary inhibition of a subset of neurons. Similar to other plasticity rules, which maintain information in synapses instead of somatic activity, the memory maintains information in dendritic states, which may survive somatic inhibition.

All descriptions above by default consider the network as a single system. However, motivated by very different behavior in the interactive and noninteractive groups of neurons, especially for the case of perfect memory, we propose a second, dual-system interpretation. Suppose that the input consists of two separate sources simultaneously. One is an auxiliary internal input from a certain brain region, providing fixed high intensities to the interactive group. Another is an arbitrary external input, providing relatively low intensities to the noninteractive group. For example, the internal input may represent a general control or attentional signal of the current task setting, and the external input can represent a more explicit and flexible task-relevant sensory input. The sensory input can be memorized only if the internal input is active to provide neurons receiving the sensory input with large dendritic inputs during the encoding period. The final memory maintains activities caused by both the auxiliary input and the external input.

Adopting this dual-system interpretation reduces much memory distortion. As it targets the noninteractive group, the arbitrary external input does not induce any out-degree error to other neurons. The induced memory for each neuron in this group is affected by the same slope of high firing rates as Fig. 6B shows, which is fixed internally. If the internal input is further trained to follow a linear decay in Eq. 8 (Fig. 5B), an arbitrary external input can be memorized perfectly. The potential price paid for this interpretation is that training could be required to provide a fixed internal input.

Another way to memorizing graded patterns by parametric working memory models

There exists another approach to storing novel patterns in working memory, in which each element of the input is stored in a separate simple circuit that is capable of storing a continuous scalar value (without any connections between these simple circuits) [21, 6, 16]. In other words, the same simple memory circuit is independently replicated for each input element. There have been many models to memorize a single scalar value, also known as parametric working memory [21, 6, 16], with both cellular [27, 19, 14, 39] and network mechanisms [33, 24, 7, 17, 28].

It is unclear whether this replication approach can memorize novel working memory. Typically, a parametric working memory model forms an attractor for the memorized scalar after some parameter tuning. It is possible that such attractors are gene-encoded without prior training, especially for working memory related to evolutionary important intuition. However, it is not likely that all information, such as language, concepts, or objects, has already been in the brain as attractors by gene. In other words, this approach likely needs attractor formation and is not suitable for novel working memory.

There are also distinctions in memory performance between our model and the replication approach. Our model shows multiselectivity, which is more consistent with recordings under working memory tasks [37, 26] in regions where neurons are mutually connected to a high degree. In contrast, the replication approach has separate maintenance for each input entry, showing single selectivity with no connections or interactions between replicas.

ACKNOWLEDGMENTS

This work was supported by the Air Force Office of Scientific Research (AFOSR) under award number FA9550-22-1-0532.

Appendix

A: Simulation details of the rate network model

We include details of network implementation here. The default input pattern first appearing in Fig. 2 was generated based on a cloud image (CCSN˙v2/Cu/Cu-N083.jpg) in [38]. The original image was compressed to a resolution of 50×50, converted into gray scales and ranked by its intensity in the descending order. To get rid of intensity degeneracy, which exists as intensity is an integer value out from 0 to 255, the ranking was smoothed by using the 'nearest' mode of ndimage.convolve in Python. This gets rid of degeneracy by giving a small distortion to each intensity value. Further, the minimal intensity was subtracted from all intensities. Finally, the input pattern was scaled to have a maximum value of 15.33. Eventually, this results in discrete input values I_i with i from 1 to $N = 2500$, which are approximated to be continuous, I(x), for the analysis in the main text.

A rate network model is used except in Fig. 9. For the default setting used in Eq. 2 and Eq. 3, parameters are $N = 2500$, $T_u = 20$, $T_d = 1$, $\beta = 0.0032$, $\alpha = 0.7$, $\tau = 50$ ms and $w = 1$. To memorize a novel input pattern, the network starts with all dendrites in the down-state, an input is applied for 1000 ms during the encoding period, and it is turned off for another 1000 ms during the following memory period. Equilibrium is reached in each period. Simulations are done using forward Euler's method with $dt = 1$ ms in Python [Version: 3.9.16].

In Fig. 3A,B, we implement Gaussian independent noise throughout the memory period. The noise has a zero mean and a standard deviation of 1 (Fig. 3A) or 5 (Fig. 3B), applied with a time step of 1 ms. The dark line averages the memory activity over a period of 500 ms, after 1000 ms into the memory period. The noise-free memory in Fig. 2B is reproduced as a dotted red line for comparison. Firing rates going below zero due to noise are set to zero.

In Fig. 3C, we implement three variations to the original connectivity. For the top panels,

weight values were drawn from a Gaussian distribution with mean 1 and standard deviation 0.1, 0.5 or 1 respectively, with a minimal value set to 0. For the middle panels, the connection probability p was changed to 50%, 10% or 1% respectively. As a dendrite exists only if a connection exists, the number of dendrites for each neuron is smaller. To balance this reduction, β was normalized correspondingly $\beta \leftarrow \beta/p$. For the bottom panels, the connectivity exists only locally, within the two-dimensional space defined by horizontal and vertical pixel labels with periodic boundary conditions. If two neurons are located within a radius of 20, 10 or 5, the weight value is one, otherwise zero. To balance the reduction of dendrites, β was normalized correspondingly $\beta \leftarrow \beta N/N_{loc}$, where N_{loc} is the number of dendrites in each neuron. For all panels in Fig. 3C, we use the same graded scale bar for ease of direct comparison. However, due to changes in connectivity, the actual firing rates of some neurons may exceed the maximum scale shown. This does not change qualitative results.

For figures 5-8, some modifications were applied to the default parameter values. In Fig. 5, the external input for a perfect memory has a linear part $I(x) = 14.4 - 0.00576x$, for x between 0 and $A = 1500$. For the arbitrary part, we chose another image, **CCSN˙v2/Cu/Cu-N004.jpg** in [38], followed the same steps as mentioned above and took the 1000 lowest intensity pixels. This portion was rescaled so that its maximum value matches the minimal value of the linear part in neuron 1500. In Fig. 6, the extra input was given to neurons 1700 to 1750 (Fig. 6A) or 1300 to 1350 (Fig. 6C) during the encoding period with a magnitude of 1 adding to the existing I(x). In Fig. 7, a signal with an intensity of -20 was given to neurons 150 to 250 (Fig. 7A) or 1550 to 1650 (Fig. 7C) at the end of the memory period. It lasted for 1000 ms, and another 1000 ms was simulated after its removal. In Fig. 8, we set $\alpha = 0$ and the maximum external input intensities are $I(0) = 22$, 30, 38 or 46.

For supplementary figures, some other modifications were applied to the default parameter values. In Fig. S1A, $T_d = 2$. In Fig. S2A, D, and G, the maximum external input intensities are

I(0) = 13, 17, and 19 respectively. In Fig. S3, the derivative was estimated through 4 nearby neurons from a discrete function f(x): $f'(x) \approx \frac{-f(x+2\Delta x)+8f(x+\Delta x)-8f(x-\Delta x)+f(x-2\Delta x)}{12\Delta x}$, with step size $\Delta x = 5$. In Fig. S3A, we chose the same image as used in Fig. 5, with a maximum external input intensity I(0) = 14.17. In Fig. S3C, with the default input pattern, the input pattern was rescaled such that I(0) = 15.8, and all intensities after neuron 1200 were decreased to 80% of their original values, so that a discrete jump was created in the input pattern.

B. Forgetting during the memory period

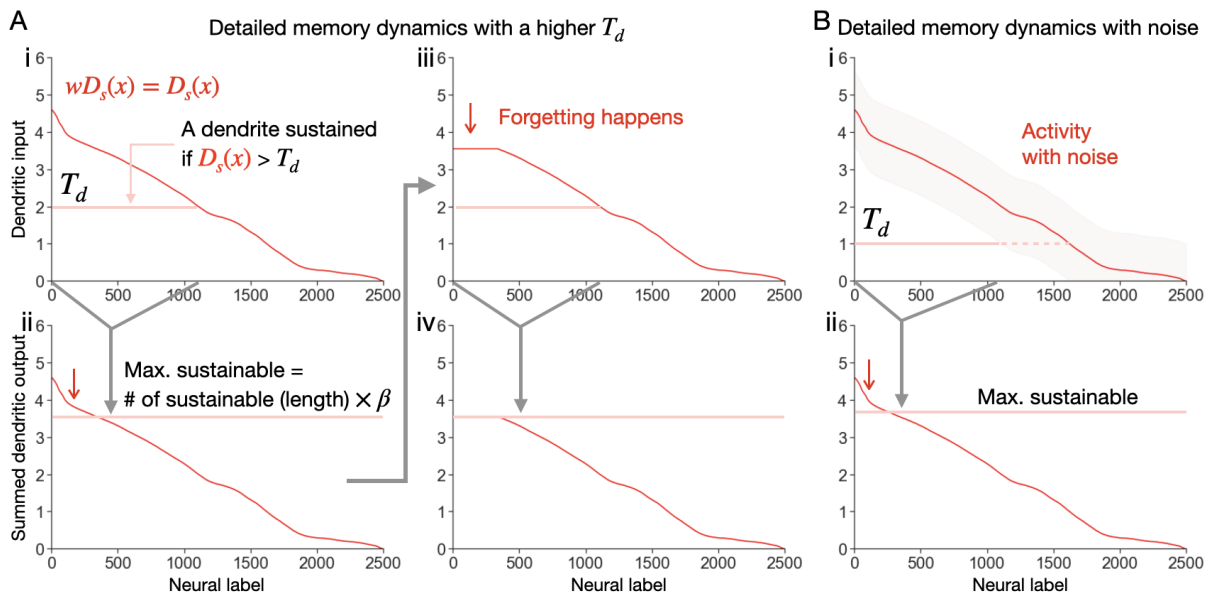

Figure S1: **Detailed memory dynamics with forgetting.** In Fig. 4G-H, dynamics during the memory period is shown with a low T_d value. Here the same dynamics is studied with a higher T_d. More explicitly, in this new situation, with Fig. 4G unchanged, Two plots in Fig. 4H, I, are replaced by two plots in each column here. **A**, Case with a higher T_d. i, The pink line shows the portion of dendrites with inputs above the new T_d with a higher value. ii, The maximum sustainable activity is given by the number of sustainable dendrites times β. It is lower than the case in Fig. 4I for the high-firing neurons. These neurons have some dendrites drop from the up-state to the down-state and reduce activity (red arrow). iii, The new $D_s(x)$ is used to determine the portion of dendrites above the new T_d, with the length of the pink line unchanged. iv, The maximally sustainable activity is unchanged, and the entire $D_s(x)$ is below it. This shows that $D_s(x)$ is now sustainable, $M(x) = D_s(x)$. **B**, Case with noise during the memory period. i, Noise makes activity fluctuate as the shaded region illustrates. The solid pink line shows the portion of dendrites with inputs constantly above T_d, but dendrites in the dashed line can go below T_d due to noise. ii, The maximally sustainable activity by the pink line. High-firing neurons will drop memory activity (red arrow) as the noisy memory example shown in Fig. 3B.

In Fig. 4G-I, we describe the memory dynamics when no forgetting happens. Here we extend the dynamics to the case with forgetting. That is, activated up-state dendrites falling

back to down-states. In this case, some dendrites are more susceptible than others. An example is shown in Fig. S1A, with all conditions identical to Fig. 4 except with a higher T_d. Fig. S1Ai, SXAii follow from Fig. 4G, replacing Fig. 4H, 4I. Compared to the previous case in Fig. 4H, in Fig. S1Ai, a higher T_d makes the number of sustainable dendrites smaller (pink line). Some up-state dendrites in high-firing neurons are not sustainable (Fig. S1Aii, red arrow) and forgetting happens in high-firing neurons (Fig. S1Aiii). Importantly, this decrease in activity may not lead to further, recursive, decrease in dendritic up-states (Fig. S1Aiv). It is because the number of sustainable dendrites is determined by the intersection of T_d and low firing rates, which may be untouched. That is, if memory forgetting happens, high-firing neurons are first affected, but not necessary for low-firing neurons.

For memory degradation under large noise (Fig. 3B), high-firing neurons are clipped due to a similar mechanism. But instead of having a higher T_d as in Fig. S1A, here activity can be temporally lower due to noise, as indicated by the shaded region in Fig. S1Bi. It also makes the number of sustainable dendrites smaller, and thus leads to forgetting. Note that in general, it is not easy to flip a down-state dendrite to the up-state, which requires noise to be comparable to the external input strength received during the encoding period.

C. Solving for a perfect memory

We derive conditions for a perfect memory here. We first simplify Eq. 7 to

$$T_u = \alpha f(x) + f(\frac{A}{T_u} f(x)). \tag{16}$$

Recall that f(x) is a non-increasing function that decays to 0, $f(N) = 0$. Substituting the relation into Eq. 16 gives: $f(0) = T_u$.

On the other hand, taking the derivative of Eq. 16 gives:

$$0 = \alpha f'(x) + f'(\frac{A}{T_u} f(x)) f'(x) \frac{A}{T_u}. \tag{17}$$

For a nontrivial solution, $f'(x) \neq 0$, we have:

$$f'(\frac{A}{T_u}f(x)) = -\frac{T_u\alpha}{A}. \tag{18}$$

This equation gives a constant slope $-\frac{T_u\alpha}{A}$ for f(x) with a range determined by $\frac{A}{T_u}f(x)$. Note that f(x) can take values from $f(N) = 0$ to $f(0) = T_u$. This means that $\frac{A}{T_u}f(x)$ takes values from 0 to A. In other words, f(x) is linear with a slope $-\frac{T_u\alpha}{A}$ for x from 0 to A.

Altogether, we have the solution for Eq. 7:

$$f(x) = T_u - \frac{T_u\alpha}{A}x, \tag{19}$$

$$I(x) = T_u - \beta A - (\frac{T_u\alpha}{A} - \alpha\beta)x, \tag{20}$$

$$D(x) = \beta A - \alpha\beta x, \tag{21}$$

where x ranges from 0 to A. We require $1 \geq \alpha > 0$ such that f(x) is not negative for $x = A$.

I(x) needs to be large enough, such that it can activate some dendrites for memory. It requires the maximum value of I(x), $I(0) = T_u - \beta A$, to satisfy:

$$T_u \leq (1 + \alpha)(T_u - \beta A) \tag{22}$$

See Eq. 11 for more details of this inequality.

On the other hand, to prevent any forgetting, the maximum sustainable line should be above $D_s(x)$, or equivalently its maximal value $D_s(0) = \beta A$, as illustrated in Fig. 4I. We need T_d to give a pink line length no less than A as illustrated in Fig. 4H, or equivalently $D_s^{-1}(T_d) \geq A$. Therefore, we have:

$$T_d \leq (1 - \alpha)\beta A. \tag{23}$$

Recall that the somatic effect can reduce T_u^e until it saturates at its lower bound, T_d. We can check if this lower bound is reached in the perfect memory case. The lowest T_u^e exists in the highest firing neuron $x = 0$ with $f(0) = T_u$, with a value $T_u - \alpha T_u$. We then have

$T_u - \alpha T_u > (1-\alpha)\beta A \geq T_d$, which shows that no saturation in T_u^e happens. Note that for the first inequality, we use the relation $T_u > \beta A$, which is derived from Eq. 22, $T_u \geq \frac{\beta A(1+\alpha)}{\alpha}$.

Altogether, we get the condition for a perfect memory as shown in the main text.

D. Further details about how input scaling leads encoding failure and saturation

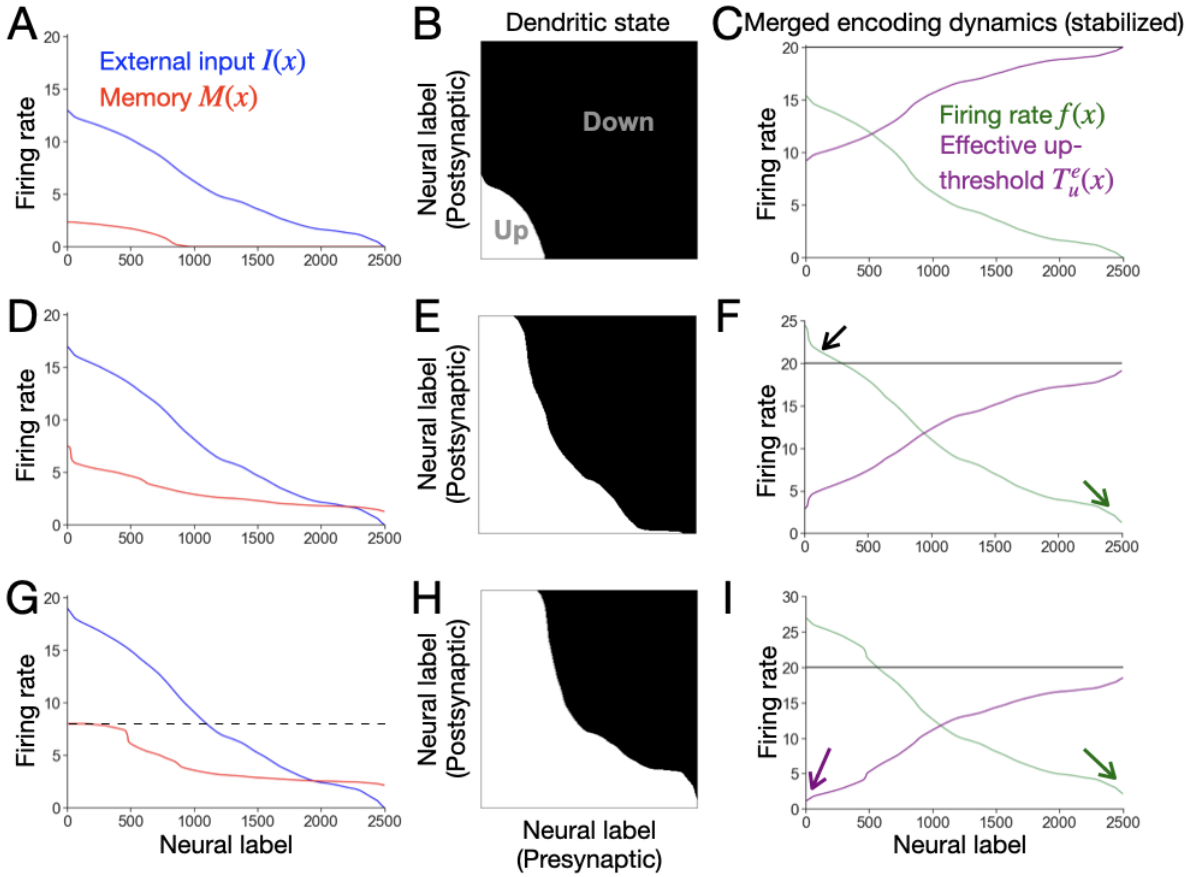

Figure S2: **Encoding failure or saturation due to a multiplicative scaling of** I(x)**. A-C**, Memory with encoding failure. **A**, An input reduced by a multiplicative scaling value, compared to the one used in Fig. 2B, results in a memory where neurons receiving low input intensities remain silent. **B**, All dendrite states of the memory. **C**, Merged encoding dynamics. For low-firing neurons, their $T_u^e(x)$ are so high that no dendrites in them can be activated, giving zero memory. **D-F**, Memory with out-degree saturation. **D**, An input pattern increased by a multiplicative scaling value leads to constant background activity in the resulting memory. **E**, All dendrite states of the memory. **F**, The high-firing neurons exceeding T_u (black arrow) activate all dendrites they project to. Therefore, every neuron has some up-state dendrites, leading to constant background activity (dark green arrow). **G-I**, Memory with in-degree saturation. **G**, An input pattern increased by a larger multiplicative scaling value than D leads to further saturation, with the memory reaching a maximum firing rate (dashed horizontal line). **H**, All dendrite states of the memory. **I**, Some effective up-thresholds of high-firing neurons (purple arrow) are lower than the smallest firing rate (dark green arrow) making all dendrites in these neurons activate.

First, we examine what is required for an input pattern to avoid encoding failure. For any neuron receiving a nonzero input activity, we want a nonzero memory activity. This requires that a neuron receiving an infinitesimal input (i.e. neuron N) to have at least one dendrite activated. The most activatable dendrite in this neuron is the one that receives dendritic input from neuron 0, with strength f(0). Therefore we need $f(0) \geq T_u$. The exemplary input in Fig. 2B is tuned to make $f(0) = T_u$ such that no encoding failure happens.

On the other hand, Fig. S2D shows an example of out-degree saturation. The high firing rates exceed T_u, such that out-degree saturation happens (Fig. S2F, black arrow), with all dendrites they project to activate (Fig. S2E). These up-state dendrites, distributed among all neurons, provide a constant background activity (Fig. S2F, green arrow). Out-degree saturation is benign, as a constant background term does not destroy graded memories.

In contrast, in-degree saturation can be more problematic, with high memory activity saturating at a certain level (Fig. S2G). One potential way it happens is when high-firing neurons have their respective T_u^e hitting the same minimal value, $T_u^e = T_d$, such that they activate the same number of up-state dendrites. Another way, as shown in Fig. S2G-I, is that, for high-firing neurons, their respective T_u^e (Fig. S2I, purple arrow) is low enough that all dendritic inputs, even from the lowest firing neuron N (green arrow), exceed it. This leads to a memory with saturation (Fig. S2G, dashed line).

To avoid encoding failure and saturation, the parameter range $0 < \alpha \leq 1$ is necessary. $f(0) \geq T_u$ is required to avoid encoding failure. Then, if $\alpha > 1$, $T_u^e(0)$ reaches its minimum value T_d, as $T_u - \alpha f(0) < T_d$. This leads to in-degree saturation.

E. Further details about how input slope leads to runaway and insensitivity

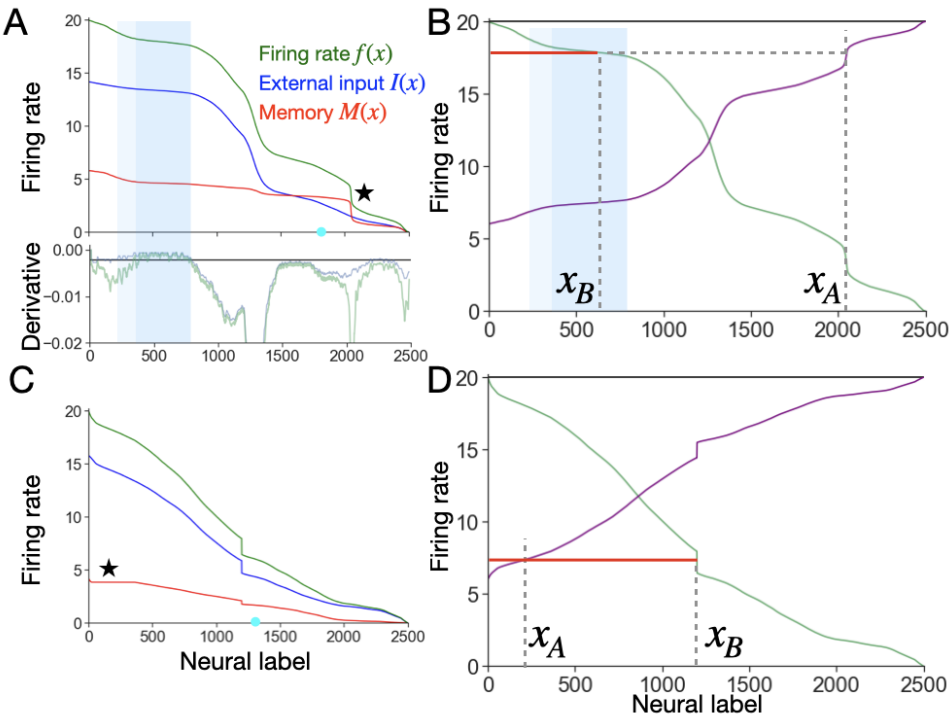

Figure S3: **Runaway and insensitivity errors.**The green curve indicates the stabilized firing rate f(x) during the encoding period. The cyan dot on the horizontal axis separates the interactive group (left) from the noninteractive group (right). **A, B,** Runaway error. **A** different input pattern from Fig. 6 is used. A, Top: Basic performance. Bottom: Derivatives of f(x) (green) and I(x) (blue). A horizontal line indicates the value of $-\alpha\beta$. The lighter blue region has a slope of I(x) which falls below $-\alpha\beta$. The darker blue region has a slope of f(x) above $-\alpha\beta$. The star indicates a local discrete jump in the memory, even though I(x) has no jump. **B**, Merged encoding dynamics. See details in the text. **C, D,** Insensitivity error. Another input pattern is used. **C**, Discrete jump in the input makes the memory discrete at the same position. In addition, the memory has a flat region as the star indicates. **D**, Merged encoding dynamics. See details in the text.

Fig. S3A shows an example of runaway error where I(x) and f(x) fail to satisfy Eq. 13 and Eq. 14. Fig. S3A, bottom, shows the slopes of I(x) and f(x), with the horizontal line showing $-\alpha\beta$. The violation of the sufficient condition, Eq. 14, may not lead to runaway error as long as Eq. 13 is still satisfied (light blue region). However, the violation of Eq. 13 (dark blue region) gives the runaway error as the star indicates. As the dynamics shown during the encoding period (Fis. S2B), the dendritic inputs from neurons in the dark blue region to neuron x_A are very similar. That is, a small increase $\delta f(x_A)$ can lead to the activation of all these dendrites through a positive feedback loop, resulting in a discontinuous jump in $D_s(x_A)$.

Fig. S3C shows an example of insensitivity error, with a discreteness in I(x). As the dynamics shown during the encoding period (Fig. S3D), due to the discontinuity in x_B, a slightly larger activity in neuron x_A does not change the in-degree of x_A, which is the length of the red line. This leads to a flat memory as the star indicates in Fig. S3C.

For runaway and insensitivity errors, we only consider how a local increase to neuron x_A directly causes x_A to gain more up-state dendrites. Technically, the local increase may cause more activity in other neurons first, which acts back to x_A for more activation. This indirect path can be ignored because an increase in x_A maximally activates one dendrite in other neurons. In the continuum limit, the contribution of a single dendrite is negligible.

F. Simulation details about the spiking network

We constructed a single neuron model with $n = 20$ dendrites to show the dendritic bistability based on dynamics of NMDA receptors. This model is based on dendritic bistability models with more detailed gating kinetics [25, 32]. Euler's method is used for integrating the differential equations in the simulation, with step size 0.25 ms, which was performed in python [Version: 3.9.16].

The somatic voltage V_s follows dynamics:

$$C\frac{dV_s}{dt} = -I_L - I_d - I_{ext} - I_{ton} - I_{s,noise}, \tag{24}$$

$$= -g_L(V_s - E_L) - \sum_i^n g_{ds}(V_s - V_{d,i}) \tag{25}$$

$$- g_{ext}s_{ext}(V_s - E_{ext}) - I_{ton} - I_{s,noise}, \tag{26}$$

where $C = 10$ nF/mm^2, $g_L = 1$ μS/mm^2, $E_L = -80$ mV, $g_{ds} = 0.01$ μS/mm^2, $g_{ext} = 30$ μS/mm^2, $E_{ext} = 0$ mV and $I_{ton} = -34$ nA/mm^2. $I_{s,noise}$ is independent Gaussian noise with a mean zero and a standard deviation of one applied at every time step 1 ms. The synaptic activation s_{ext} was driven by an external spike train with dynamics specified below. Once V_s

crossed the firing threshold –50 mV, a spike was triggered and the somatic voltage was set to 30 mV for 3 ms before resetting to –55 mV for another 3 ms.

The dendritic voltage $V_{d,i}$ follows dynamics:

$$C\frac{dV_{d,i}}{dt} = -I_{L,i} - I_{s,i} - I_{NMDA,i} - I_{KIR,i} - I_{d,noise,i}, \tag{27}$$

and each term on the right hand side is:

$$I_{L,i} = g_L(V_{d,i} - E_L), \tag{28}$$

$$I_{s,i} = \varkappa g_{ds}(V_{d,i} - V_s), \tag{29}$$

$$I_{NMDA,i} = g_{NMDA,i} s_{NMDA,i} \frac{V_{d,i} - E_{NMDA}}{1 + 0.15e^{-0.08V_{d,i}}}, \tag{30}$$

$$I_{KIR,i} = g_{KIR,i} \frac{V_{d,i} - E_{KIR}}{1 + e^{0.1(V_{d,i} - E_{KIR} + 10)}}, \tag{31}$$

where $E_{NMDA} = 0$ mV, $E_{KIR} = -90$ mV and the area ratio of the soma to a dendrite $\varkappa = 37$. $I_{d,noise,i}$ is independent Gaussian noise with a mean zero and a standard deviation 1. $g_{NMDA,i}$ was generated from a Gaussian random variable with mean 34.5 μS/mm^2 and standard deviation 1. Similarly, $g_{KIR,i}$ with mean 41 μS/mm^2 and standard deviation 1.

The synaptic activation s_α, with $\alpha =$ ext or NMDA, i, follows dynamics:

$$\frac{ds_\alpha}{dt} = -s_\alpha/\tau_\alpha,$$

$$\Delta s_\alpha = 0.5(1 - s_\alpha), \text{ each time a spike is received,}$$

where $\tau_{ext} = 2$ ms and $\tau_{NMDA,i} = 100$ ms respectively. s_{ext} receives an external spike train, and $s_{NMDA,i}$ receives the spike train generated by neuron i.

Such a single neuron model shows dendritic bistability. To generate a bistable curve in Fig. 9B, a Poisson spike train with a frequency of 30 Hz is delivered to the soma, and a fixed period spike train with changing frequencies is delivered to all 20 dendrites. The latter mimics dendritic inputs from other neurons, which has a relatively fixed period due to the integrate-and-fire soma. The dendritic input frequency starts from 0 Hz and increases by 1 Hz every 1400 ms

until it reaches 150 Hz. Afterwards, it decreases by 1 Hz every 1400 ms until it goes back to 0 Hz. The dendritic voltage shown is averaged over the last 500 ms before the frequency change happens. The black curve in Fig. 9B is generated with another somatic input of 210 Hz. For Fig. 9C, we ran 10 trials with the somatic input ranges from 30 Hz to 210 Hz and recorded the averaged threshold values across all trials and all 20 dendrites. Thresholds are estimated by the largest slopes.

We extended the neural model to a homogeneous network of $N = 40$ neurons. The probability a connection exists from one neuron to another is $p = 0.5$, which makes n a variable with mean 20. Voltages are stabilized at about -80mV initially. The input to the soma of each neuron is a Poisson spike train with a linear frequency from 100 to 0 Hz (Fig. 9D). The encoding period lasts for 1200 ms, the memory period lasts for 800 ms. The dendritic voltage and firing rate are averaged over the last 300 ms (Fig. 9E). We further ran 20 trials to get the statistical performance (Fig. 9F).

References

[1] Shun-ichi Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological cybernetics*, 27(2):77–87, 1977.

[2] Roberto Araya. Dendritic morphology and function. *Neuroscience in the 21st Century: From Basic to Clinical*, pages 571–606, 2022.

[3] Lou Beaulieu-Laroche, Enrique HS Toloza, Norma J Brown, and Mark T Harnett. Widespread and highly correlated somato-dendritic activity in cortical layer 5 neurons. *Neuron*, 103(2):235–241, 2019.

[4] Bardia F Behabadi and Bartlett W Mel. Mechanisms underlying subunit independence in pyramidal neuron dendrites. *Proceedings of the National Academy of Sciences*, 111(1):498–503, 2014.

[5] Flora Bouchacourt and Timothy J Buschman. A flexible model of working memory. *Neuron*, 103(1):147–160, 2019.

[6] Carlos D Brody, Ranulfo Romo, and Adam Kepecs. Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Current opinion in neurobiology*, 13(2):204–211, 2003.

[7] Kathleen P Champion, Olivia Gozel, Benjamin S Lankow, G Bard Ermentrout, and Mark S Goldman. An oscillatory mechanism for multi-level storage in short-term memory. *Communications biology*, 6(1):829, 2023.

[8] Albert Compte, Nicolas Brunel, Patricia S Goldman-Rakic, and Xiao-Jing Wang. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral cortex*, 10(9):910–923, 2000.

[9] Christos Constantinidis and Xiao-Jing Wang. A neural circuit basis for spatial working memory. *The Neuroscientist*, 10(6):553–565, 2004.

[10] Victor Hugo Cornejo, Netanel Ofer, and Rafael Yuste. Voltage compartmentalization in dendritic spines in vivo. *Science*, 375(6576):82–86, 2022.

[11] Martha A Erickson, Lauren A Maramara, and John Lisman. A single brief burst induces glur1-dependent associative short-term potentiation: a potential mechanism for short-term memory. *Journal of cognitive neuroscience*, 22(11):2530–2540, 2010.

[12] Florian Fiebig and Anders Lansner. A spiking working memory model based on hebbian short-term potentiation. *Journal of Neuroscience*, 37(1):83–96, 2017.

[13] Valerio Francioni and Mark T Harnett. Rethinking single neuron electrical compartmentalization: dendritic contributions to network computation in vivo. *Neuroscience*, 489:185–199, 2022.

[14] Erik Fransén, Babak Tahvildari, Alexei V Egorov, Michael E Hasselmo, and Angel A Alonso. Mechanism of graded persistent cellular activity of entorhinal cortex layer v neurons. *Neuron*, 49(5):735–746, 2006.

[15] Nace L Golding, Nathan P Staff, and Nelson Spruston. Dendritic spikes as a mechanism for cooperative long-term potentiation. *Nature*, 418(6895):326–331, 2002.

[16] Mark S Goldman, Albert Compte, and Xiao-Jing Wang. Neural integrator models. *Encyclopedia of neuroscience*, pages 165–178, 2009.

[17] Mark S Goldman, Joseph H Levine, Guy Major, David W Tank, and H Sebastian Seung. Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cerebral cortex*, 13(11):1185–1195, 2003.

[18] Jason Hardie and Nelson Spruston. Synaptic depolarization is more effective than back-propagating action potentials during induction of associative long-term potentiation in hippocampal pyramidal neurons. *Journal of Neuroscience*, 29(10):3233–3241, 2009.

[19] Michael E Hasselmo and Chantal E Stern. Mechanisms underlying working memory for novel information. *Trends in cognitive sciences*, 10(11):487–493, 2006.

[20] Donald Olding Hebb. *The organization of behavior: A neuropsychological theory*. Psychology press, 2005.

[21] Alexei A Koulakov, Sridhar Raghavachari, Adam Kepecs, and John E Lisman. Model for a robust neural integrator. *Nature neuroscience*, 5(8):775–782, 2002.

[22] Kamesh Krishnamurthy, Tankut Can, and David J Schwab. Theory of gating in recurrent neural networks. *Physical Review X*, 12(1):011011, 2022.

[23] Anders Lansner, Florian Fiebig, and Pawel Herman. Fast hebbian plasticity and working memory. *Current opinion in neurobiology*, 83:102809, 2023.

[24] Sukbin Lim and Mark S Goldman. Balanced cortical microcircuitry for maintaining information in working memory. *Nature neuroscience*, 16(9):1306–1314, 2013.

[25] John E Lisman, Jean-Marc Fellous, and Xiao-Jing Wang. A role for nmda-receptor channels in working memory. *Nature neuroscience*, 1(4):273–275, 1998.

[26] Mikael Lundqvist, Scott L Brincat, Jonas Rose, Melissa R Warden, Timothy J Buschman, Earl K Miller, and Pawel Herman. Working memory control dynamics follow principles of spatial computing. *Nature Communications*, 14(1):1429, 2023.

[27] Guy Major and David Tank. Persistent neural activity: prevalence and mechanisms. *Current opinion in neurobiology*, 14(6):675–684, 2004.

[28] Maxim Nikitchenko and Alexei Koulakov. Neural integrator: A sandpile model. *Neural computation*, 20(10):2379–2417, 2008.

[29] Pojeong Park, Arturas Volianskis, Thomas M Sanderson, Zuner A Bortolotto, David E Jane, Min Zhuo, Bong-Kiun Kaang, and Graham L Collingridge. Nmda receptor-dependent long-term potentiation comprises a family of temporally overlapping forms of synaptic plasticity that are induced by different patterns of stimulation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1633):20130131, 2014.

[30] Stefan Remy and Nelson Spruston. Dendritic spikes induce single-burst long-term potentiation. *Proceedings of the National Academy of Sciences*, 104(43):17192–17197, 2007.

[31] A Sandberg, Jesper Tegnér, and Anders Lansner. A working memory model based on fast hebbian learning. *Network: Computation in Neural Systems*, 14(4):789, 2003.

[32] Honi Sanders, Michiel Berends, Guy Major, Mark S Goldman, and John E Lisman. Nmda and gabab (kir) conductances: the "perfect couple" for bistability. *Journal of Neuroscience*, 33(2):424–429, 2013.

[33] H Sebastian Seung, Daniel D Lee, Ben Y Reis, and David W Tank. Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron*, 26(1):259–271, 2000.

[34] Nelson Spruston. Pyramidal neurons: dendritic structure and synaptic integration. *Nature Reviews Neuroscience*, 9(3):206–221, 2008.

[35] Greg J Stuart and Nelson Spruston. Dendritic integration: 60 years of progress. *Nature neuroscience*, 18(12):1713–1721, 2015.

[36] Botond Szatmáry and Eugene M Izhikevich. Spike-timing theory of working memory. *PLoS computational biology*, 6(8):e1000879, 2010.

[37] Melissa R Warden and Earl K Miller. Task-dependent changes in short-term memory in the prefrontal cortex. *Journal of Neuroscience*, 30(47):15801–15810, 2010.

[38] Jinglin Zhang, Pu Liu, Feng Zhang, and Qianqian Song. Cloudnet: Ground-based cloud classification with deep convolutional neural network. *Geophysical Research Letters*, 45(16):8665–8672, 2018.

[39] Joel Zylberberg and Ben W Strowbridge. Mechanisms of persistent activity in cortical circuits: possible neural substrates for working memory. *Annual review of neuroscience*, 40:603–627, 2017.