High Capacity and Dynamic Accessibility in Associative Memory Networks with Context-Dependent Neuronal and Synaptic Gating

William F. Podlaski,[1,2,*] Everton J. Agnes,[1,3,†,‡] and Tim P. Vogels[1,4,‡]

[1]Centre for Neural Circuits and Behaviour, University of Oxford, OX1 3SR Oxford, United Kingdom
[2]Champalimaud Neuroscience Programme, Champalimaud Foundation, 1400-038 Lisbon, Portugal
[3]Biozentrum, University of Basel, 4056 Basel, Switzerland
[4]Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria

(Received 19 December 2022; revised 8 August 2024; accepted 12 December 2024; published 13 March 2025)

Biological memory is known to be flexible—memory formation and recall depend on factors such as the behavioral context of the organism. However, this property is often ignored in associative memory models, leaving it unclear how memories can be organized and recalled when subject to contextual control. Because of the lack of a rigorous analytical framework, it is also unknown how contextual control affects memory stability, storage capacity, and information content. Here, we bring the dynamic nature of memory to the fore by introducing a novel model of associative memory, which we refer to as the context-modular memory network. In our model, stored memory patterns are associated to one of several background network states, or contexts. Memories are accessible when their corresponding context is active, and are otherwise inaccessible. Context modulates the effective network connectivity by imposing a specific configuration of neuronal and synaptic gating—gated neurons (synapses) have their activity (weights) momentarily silenced, thereby reducing interference from memories belonging to other contexts. Memory patterns are randomly and independently chosen, while neuronal and synaptic gates may be selected randomly or optimized through a process of contextual synaptic refinement. Through analytic and numerical results, we show that context-modular memory networks can exhibit both improved memory capacity and differential control of memory stability with random gating (especially for neuronal gating). For contextual synaptic refinement, we devise a method in which synapses are gated off for a given context if they destabilize the memory patterns in that context, drastically improving memory capacity and enabling even more precise control over memory stability. Notably, synaptic refinement allows for patterns to be accessible in multiple contexts, stabilizing memory patterns even for weight matrices that alone do not contain any information about the memory patterns, such as Gaussian random matrices. Overall, our model integrates recent ideas about context-dependent memory organization with classic associative memory models and proposes a rigorous theory which can act as a framework for future work. Furthermore, our work carries important implications for the understanding of biological memory storage and recall in the brain, such as highlighting an intriguing trade-off between memory capacity and accessibility.

DOI: 10.1103/PhysRevX.15.011057

Subject Areas: Biological Physics, Interdisciplinary Physics, Statistical Physics

I. INTRODUCTION

Discrete attractor networks offer a simple yet powerful model for autoassociative memory, enabling the retrieval of stored, content-addressable memory patterns from noisy external cues [1–3]. The biological and computational relevance of such associative memory models depends on the statistics of the memory patterns and the choice of connectivity architecture [3–5], among other factors. While classic models typically feature uncorrelated patterns (i.e., half of the network units are active for each pattern, chosen uniformly at random) and all-to-all connectivity [1], many other variants have been proposed and explored [3]. For example, sparse [6], hierarchical [7,8], and structured [9] patterns have been shown to confer computational advantages to storage capacity and recall, and are more biologically plausible [10]. Furthermore, networks with sparse [11] or modular [12–15] connectivity exhibit improved storage ability and their structure may resemble brain circuits more closely [16,17].

*Contact author: william.podlaski@gmail.com
†Contact author: everton.agnes@gmail.com
‡These authors contributed equally to this work.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

A common feature of much theoretical work on associative memory is the focus on memory capacity—the maximum number of stable patterns that a network is able to store [3,6,18,19]. While capacity is certainly a crucial quantity when considering the function of associative memory models, the psychology and neuroscience literature has long highlighted the dynamic and context-dependent nature of memory as an equally significant property [20–27]. As early as the 1960s, for example, it was observed in free-recall experiments, in which subjects were asked to memorize and later recall lists of words, that sequential cued recall of lists by contextual or conceptual category was often more successful than unconstrained recall of all lists simultaneously [20]. This and related work has led to the proposal that memory *accessibility*—the relative ease by which a particular memory can be retrieved—is not constant, but can vary over time depending upon additional cues and background context [20,28]. This concept provides an explanation for some memory recall failures, such as the "tip of the tongue" phenomenon in which a memory or a piece of information is momentarily unable to be recalled, suggesting temporary inaccessibility [25,29].

Though sometimes inconvenient (such as in the above example), this dynamic nature of memory has the potential to enable powerful flexibility over when and how memories are accessed, which biological memory systems may employ to influence behavior [26,27,30]. Given the vast number of memories that humans and other animals may acquire over a lifetime, the ability to dynamically search for and recall specific subsets is likely just as important as the ability to store them in the first place [31]. Furthermore, similar or identical cues often require very different behavioral responses in different contexts—e.g., encountering a dangerous animal at the zoo versus in the wild [23]. Studies of contextual fear conditioning confirm that such fear memories are often linked to particular contexts [23,25], and that overgeneralization across contexts may explain pathological memory conditions such as post-traumatic stress disorder [23]. In sum, neuroscience and psychology have reached the consensus that the ability to dynamically store [24] and recall [25] memories depending upon the behavioral context is a crucial aspect of memory's functional role. However, despite several algorithmic models of context-dependent memory [28,32], classic network models of associative memory typically feature static and uniform memory stability (i.e., all memories can be recalled at all times with similar accessibility), and thus not only fail in capturing this crucial aspect of biological memory, but also miss out on possible computational advantages of such a flexible approach.

The experimental evidence for dynamic memory accessibility [20–23,25,26], as well as the potential computational advantages of such flexibility [24,25], raises fundamental questions about the circuit-level organization of memory and its theoretical study. First, how can a memory system reliably store a large number of memory patterns while also flexibly regulating their accessibility? Obtaining an answer to this question is crucial for a complete understanding of biological memory and its role in complex behavior. Doing so requires specific hypotheses for the mechanistic basis of dynamic memory stability control in neural circuits. Such mechanisms should enable a network to modulate the stability of particular attractor states on demand in order to read out memories when necessary, but avoid recall when undesired. One promising theory, backed up by a host of experimental evidence, suggests that memories may be bound to specific contexts, requiring reactivation of the encoding context (i.e., reinstatement) for successful recall [28,31,33,34]. Mechanistically, top-down contextual signals from distinct brain areas may act as control signals for memory storage and recall [26,35–37]. Previous models have shown that contextual input or gating is a viable option for flexible information routing [38,39], as well as continual learning in artificial neural networks [40]. However, it is unclear how such schemes may be included, and rigorously studied, in associative memory models.

This leads to a second important question: How might such a dynamic memory system be studied analytically? Without a theoretical framework and analytical insights, it will be virtually impossible to unveil possible mechanistic, biological implementations of such dynamic accessibility, to understand what the strengths and weaknesses of flexible memory implementations are, and to eventually elaborate predictions for future experiments. Previous contextual gating models have primarily relied upon numerical simulations, making broad computational claims difficult. The large body of theoretical analysis tools based on associative memory models [3] provides an ideal starting point for the study of contextual control of memory recall, and may serve as the basis to explore whether rigorous theory of flexible memory might have implications for other domains, such as sensory processing and information routing. However, up to this point, it remains unclear how previous mathematical analysis of associative memory may be adapted to such a setting.

Here, we aim to answer the fundamental questions mentioned above by introducing a new model of associative memory, which we refer to as the *context-modular memory network*. In the model, retrieval dynamics are controlled by an externally imposed background state, which we refer to as *context*. This control is exerted onto the network by silencing neurons and synapses, termed *neuronal gating* and *synaptic gating*, respectively. These gating schemes, which are defined in Sec. II, modify the effective network connectivity, and thereby also the accessibility of memories, i.e., their stability. We show in Sec. III that Hebbian learning in concert with random gating can considerably improve memory capacity when compared

to the classic Hopfield model, especially for random neuronal gating (e.g., up to 7 times more for a network with 100 contexts). We then show that both random neuronal and synaptic gating can effectively stabilize memories in the active context and destabilize memories in the inactive context, thereby modulating memory accessibility. Interestingly, these results may indicate that low levels of neuronal activity found in the brain (and their allocation to particular memories [24]) might not only serve to increase memory capacity, but also to flexibly control memory recall. We then introduce the concept of *refinement* in Sec. IV, in which synaptic gating (but not neuronal gating) can be optimized to improve memory stability for specific groups of previously learned patterns. This scheme results in a drastic increase in memory capacity (e.g., up to 40 times more for a network with 100 contexts), plus an even more powerful ability to modulate memory accessibility. Finally, we show that overlapping sets of memories across different contexts can be stabilized through synaptic refinement, demonstrating how contexts may share subsets of common memories. The precise control needed to implement synaptic refinement suggests novel hypotheses for ongoing plasticity following initial memory storage, which can be explored experimentally. Overall, this work incorporates novel ideas about the dynamic and behaviorally relevant use of memory into classic associative memory models, and may serve as a novel framework for the study of flexible associative memory.

II. MODEL

A. Model description

We introduce a context-dependent autoassociative memory network, the context-modular memory network, whose primary functionality is to have dynamic memory accessibility—that is, depending on a background context signal, the model stabilizes a subset of memory patterns while keeping others hidden. This is achieved by temporarily modifying the network such that there is a different effective connectivity matrix for each contextual state [Fig. 1(a)]. To implement such a model, we modify a standard, fully connected Hopfield network with N neurons [1] to include s contextual states, i.e., contexts, which control two types of gating. First, in neuronal (nrn) gating, each context (cxt) has a corresponding set of allocated neurons N_{cxt}, while the other $N - N_{\mathrm{cxt}}$ neurons are "gated off" (defined more precisely below). We use the parameter $a = N_{\mathrm{cxt}}/N$ to denote the proportion of allocated neurons per context. Second, in synaptic (syn) gating, each contextual state defines a particular set of allocated connections (synapses) per allocated neuron K, with the other $N_{\mathrm{cxt}} - K$ connections being gated off. We use the parameter $c = K/N_{\mathrm{cxt}}$ to denote the connection density.

We impose that only one context is active at any given time, and we consider the recall dynamics of the network

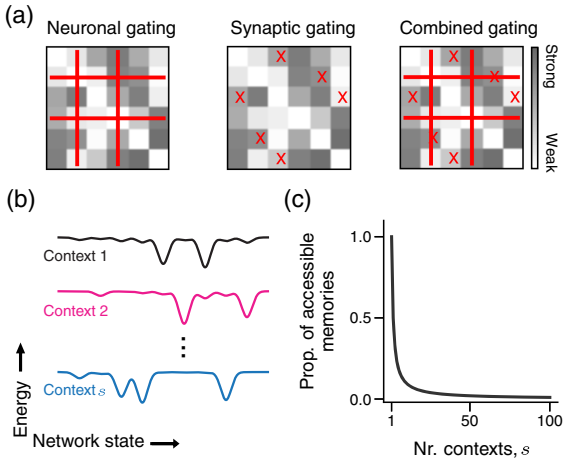


FIG. 1. The context-modular memory network. (a) Contextual configurations alter the effective connectivity matrix of the associative memory network: Neuronal gating removes particular columns and rows (left), synaptic gating removes individual elements (center), and together, they implement both effects (right). (b) Each contextual state produces a distinct energy landscape with a unique set of attractor memory patterns, with others remaining hidden. (c) Proportion of accessible memories, i.e., ratio of memories stored per context divided by the total number of memories as a function of the number of contexts s.

from the perspective of this active context, which we denote by k. The state of neuron i is described by $S_i^{(k)} \in \{0, \pm 1\}$ when context k is active, with dynamics

$$S_i^{(k)} = \mathrm{sgn}\left(a_i^k \sum_{j=1}^{N} a_j^k c_{ij}^k J_{ij} S_j^{(k)} \right), \tag{1}$$

where J_{ij} is the synaptic weight between neurons j and i (independent of context), and $a_j^k \in \{0, 1\}$ and $c_{ij}^k \in \{0, 1\}$ represent the effects of neuronal and synaptic gating for context k, respectively. Note that we use the convention $\mathrm{sgn}(0) = 0$ in Eq. (1). We thus see that "gating off" refers to the setting of $a_j^k = 0$ or $c_{ij}^k = 0$, which effectively removes the neuron or connection, respectively, from the network dynamics. As a consequence, neuronal gating does not set neurons to the typical inactive state of the Hopfield network ($S_i^{(k)} = -1$), but instead to zero. This can be interpreted as setting all the output connections of the gated neuron to zero, but is equivalent to inactivation of the neuron in a more biologically plausible $\{0, 1\}$ formulation of the network (Appendix A). We note that standard Hopfield dynamics correspond to setting $a_j^k = 1$ and $c_{ij}^k = 1$ for all i and j.

The network stores p memory patterns per context, making $P = sp$ total memories (except for the case in which memories are randomly allocated to multiple contexts, explored in Sec. IV B). Memory patterns are defined

as $\xi_i^{\sigma^\mu} \in \{\pm 1\}$ for neuron i and memory μ of context σ, and we consider random uncorrelated patterns with $\Pr(\xi_i^{\sigma^\mu} = \pm 1) = 1/2$.

B. Model interpretation

The gating schemes defined above can be interpreted as mechanisms that impose a distinct effective connectivity matrix for each contextual state. The network dynamics can then be viewed as following a context-dependent energy function [Fig. 1(b)],

$$H^{(k)} = -\sum_{i=1}^{N}\sum_{j=1}^{N} a_i^k a_j^k c_{ij}^k J_{ij} S_i^{(k)} S_j^{(k)}, \tag{2}$$

with context k being active. Neuronal gating temporarily removes particular rows and columns from the connectivity matrix [Fig. 1(a), left], whereas synaptic gating temporarily removes individual entries [Fig. 1(a), middle]. The combination of the two produces a smaller and sparser connectivity matrix [Fig. 1(a), right], with a potentially large number of synaptic connections remaining hidden, to be used in other contexts. This enables the fully connected network to exhibit an effectively modular structure, in which the active context k imposes an *inhibitory mask*, given by $a_i^k a_j^k c_{ij}^k$, onto the connectivity matrix. It will be useful to analyze the model in terms of the proportion of gated neurons, i.e., neuronal gating ratio, $1-a$, and proportion of gated synapses, i.e., synaptic gating ratio, $1-c$. The neuronal gating ratio indicates the proportion of neurons that are gated off while a given context is active. The synaptic gating ratio, usually referred to as dilution [41], indicates the proportion of synapses that are gated off during recall of memories in the active context.

A crucial feature of the model is the assumption that only a subset of memories needs to be stable in each context. As the number of contexts s increases, the proportion of accessible memories decreases as $1/s$ [Fig. 1(c)], leading to a clear trade-off between the storage capacity gains (which increases with s) and memory accessibility across all patterns at any given time. This also means that the proportion of inaccessible memories grows with the number of contexts, making it just as important to ensure that the hidden memories are unstable as it is to ensure that the accessible memories are stable.

C. Memory capacity definitions

We define parameters related to pattern storage to study the maximum number of memories that can be successfully retrieved. The network *load* refers to the ratio of stored patterns to network size. Here, we have two notions of load. First, the *single-context load* is defined as

$$\alpha_{\mathrm{cxt}} = p/N_{\mathrm{cxt}}, \tag{3}$$

which is the number of patterns stored in each individual context divided by the number of allocated neurons per context. For simplicity, we impose that all contexts store the same number of patterns and have the same number of allocated neurons, and so the single-context load is identical across all contexts. Second, the *total load* is

$$\alpha = P/N = sp/N = \alpha_{\mathrm{cxt}} sa, \tag{4}$$

which is the total number of patterns stored in the network across all contexts divided by the network size. The maximum load for which all patterns are stable attractors is known as the memory capacity. We refer to the single-context and total capacities as α^*_{cxt} and α^*, respectively [42]. Because of analytical tractability, memory capacity is probed conditioned on an externally imposed context (see Sec. V).

D. Memory capacity and overlap estimations

We study the memory capacity of the context-modular memory networks using a mean-field analysis [3,18,41], whose details are given in Appendix B. In brief, we consider an effective weight matrix $\tilde{J}_{ij}^k$ in context k by writing the gated weight matrix $a_i^k a_j^k c_{ij}^k J_{ij}$ from Eq. (1) considering only the neurons allocated to context k. We rewrite the effect of gating (i.e., effective weight matrix) as the weight matrix of an ungated equivalent classic Hopfield network plus a noise term as

$$\begin{aligned}\tilde{J}_{ij}^k &= a_i^k a_j^k c_{ij}^k J_{ij} \\ &= a_i^k a_j^k \left(\frac{1}{aN}\sum_{\mu=1}^{p} \xi_i^{k^\mu}\xi_j^{k^\mu} + \frac{1}{aN}\sum_{\mu=p+1}^{\mathcal{P}} \tilde{\xi}_i^{k^\mu}\tilde{\xi}_j^{k^\mu} + \delta_{ij}^k\right) \\ &= a_i^k a_j^k \left(\frac{1}{aN}\sum_{\mu=1}^{\mathcal{P}} \tilde{\xi}_i^{k^\mu}\tilde{\xi}_j^{k^\mu} + \delta_{ij}^k\right).\end{aligned} \tag{5}$$

Following Refs. [41,43,44], we can then estimate both memory capacity and overlap as a function of the variance of the noise term, δ_{ij}^k. In Eq. (5), $\mathcal{P}$ is the number of stable memory pattern terms. It includes the p patterns from the active context, $\tilde{\xi}_i^{k^\mu} = \xi_i^{k^\mu}$ for $i = 1,\ldots,N$ and $\mu = 1,\ldots,p$, and a subset of spurious patterns generated by the contribution of memory pattern terms from the inactive contexts, $\tilde{\xi}_i^{k^\mu}$ for $i = 1,\ldots,N$ and $\mu = p+1,\ldots,\mathcal{P}$. For simplicity, and without loss of generality, we assume that $\mathrm{Prob}(\tilde{\xi}_i^{k^\mu} = \pm 1) = 1/2$ for $i = 1,\ldots,N$ and $\mu = p+1,\ldots,\mathcal{P}$. We estimate $\mathcal{P}$ based on the learning rule, gating variables, and the parameter choices below.

The noise term δ_{ij}^k is well approximated as Gaussian noise with zero mean and variance $\Delta^2/(aN)$ [41], with $\Delta^2 = \alpha_{\mathcal{P}}\Delta_0^2$, where Δ_0^2 depends on the gating method, and

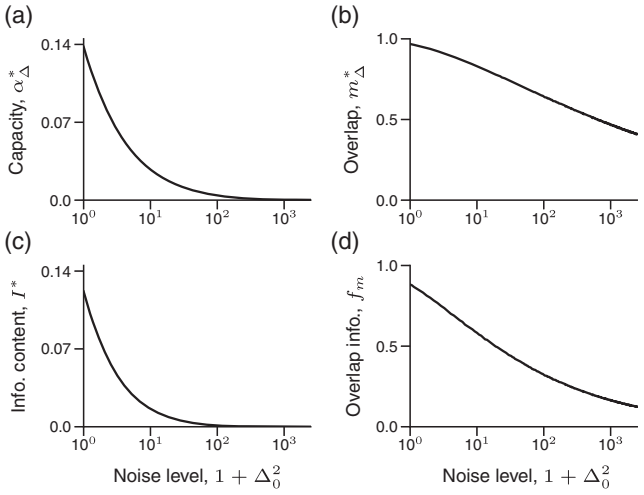


FIG. 2. Mean-field solution for a Hopfield network with random Gaussian noise following Ref. [41], shown as a function of the noise level, $1+\Delta_0^2$. (a) Maximum capacity. (b) Overlap at maximum capacity. (c) Information content per synapse. (d) Overlap information $f_{\mathrm{m}}(m_\Delta)$.

$\alpha_{\mathcal{P}} = \mathcal{P}/N_{\mathrm{cxt}}$. In general terms, we can therefore estimate $\alpha^*_{\mathcal{P}}$ as a function of $1+\Delta_0^2$,

$$\alpha^*_{\mathcal{P}} = \alpha^*_\Delta(1+\Delta_0^2), \tag{6}$$

where $\alpha^*_\Delta(1+\Delta_0^2)$ is the memory capacity of a Hopfield network with elements of the weight matrix disrupted by Gaussian noise with zero mean and variance $\alpha^*_\Delta\Delta_0^2$ [Fig. 2(a); see Appendix B and Refs. [3,18,45]].

Importantly, the average overlap between the memory patterns and their corresponding attractors at capacity m^* depends on Δ_0^2 as well [Fig. 2(b)]. We define the overlap between the memory pattern μ of context k with the network state as

$$m^{k\mu} = \frac{1}{aN}\sum_{i=1}^{N} a_i^k \xi_i^{k\mu} S_i. \tag{7}$$

Similarly to the memory capacity, we also describe the average overlap (at capacity) as a function of Δ_0^2:

$$m^* = m^*_\Delta(1+\Delta_0^2). \tag{8}$$

From Eq. (6), we can estimate the context-specific memory capacity α^*_{cxt} as

$$\alpha^*_{\mathrm{cxt}} = \frac{p}{\mathcal{P}}\alpha^*_{\mathcal{P}} = \frac{p}{\mathcal{P}}\alpha^*_\Delta(1+\Delta_0^2), \tag{9}$$

which considers only the subset of p explicitly stored memory patterns.

E. Information content definition

We calculate the information content per synapse as the total entropy across all stored patterns [19] considering the original memory patterns minus the missing information due to inaccurate recall divided by the total number of synapses [46]. For a Hopfield network with Gaussian noise, the information content at capacity can be calculated as a function of the noise [Fig. 2(c)], which is

$$I_\Delta = pN[S_{1/2} - S(m)]/N^2 = \alpha_\Delta f_{\mathrm{m}}(m_\Delta), \tag{10}$$

where $S_{1/2}$ is the entropy per neuron given that $\Pr(\xi_i^{\sigma^\mu} = \pm 1) = 1/2$, $S(m)$ is the missing information given that the average overlap between original memory patterns and corresponding attractors is m [46], and thus the overlap information [Fig. 2(d)] is

$$f_{\mathrm{m}}(m) = \left[\frac{1+m}{2}\log_2(1+m) + \frac{1-m}{2}\log_2(1-m)\right]. \tag{11}$$

Note that $f_{\mathrm{m}}(m)$ is necessary because m^* decreases considerably as a function of $1+\Delta_0^2$ [Figs. 2(b) and 2(d)]. Following the same logic for the context-modular Hopfield network, the information content at capacity can be written as

$$I^* = a\alpha^* f_{\mathrm{m}}(m^*). \tag{12}$$

F. Memory accessibility definitions

To measure memory accessibility, we consider the so-called *stability parameter* of each memory pattern [19,47], which provides an estimate of the size of each pattern's basin of attraction. For active context k, we write the stability of a particular pattern ν of context σ for an arbitrary neuron i as

$$\kappa_i^{k,\sigma^\nu} = \frac{a_i^k a_i^\sigma \xi_i^{\sigma^\nu} \sum_{j=1}^{N} a_j^k c_{ij}^k J_{ij} \xi_j^{\sigma^\nu}}{|J_i^k|}, \tag{13}$$

with

$$|J_i^k| = \sqrt{\sum_{j=1}^{N} a_j^k c_{ij}^k J_{ij}^2}. \tag{14}$$

Generally, we will consider the average stability parameter of accessible (acc) memories when setting $\sigma = k$ in Eq. (13),

$$\bar{\kappa}_{\mathrm{acc}} = \left\langle \kappa_i^{k,k^\nu} \right\rangle_{i,k,k^\nu}, \tag{15}$$

and the stability of inaccessible (inacc) memory patterns by setting $\sigma \neq k$ in Eq. (13),

$$\bar{\kappa}_{\text{inacc}} = \left\langle \kappa_i^{k,\sigma^\nu} \right\rangle_{i,k,\sigma,\sigma^\nu}, \tag{16}$$

where $\langle\cdot\rangle_{i,k,k^\nu}$ represents the average over all neurons (i), contexts (k), and patterns of context k (k^ν), and $\langle\cdot\rangle_{i,k,\sigma,\sigma^\nu}$ represents the average over all neurons (i), combinations of contexts (k and σ, with $\sigma \neq k$), and patterns of context σ (σ^ν).

III. RANDOM NEURONAL AND SYNAPTIC GATING

In this section, we consider the case in which neurons and synapses are allocated uniformly at random to each context. That is, for each context σ, variables a_i^σ and c_{ij}^σ [Eq. (1)] are set to 1 or 0 with probabilities $\Pr(a_i^\sigma = 1) = a$ and $\Pr(a_i^\sigma = 0) = 1 - a$ for neuronal gating and $\Pr(c_{ij}^\sigma = 1) = c$ and $\Pr(c_{ij}^\sigma = 0) = 1 - c$ for synaptic gating. For simplicity and mathematical tractability, the synaptic gating matrix is symmetric; $c_{ij}^\sigma = c_{ji}^\sigma$. These quantities are used to define the weight matrix through a modified Hebbian rule as

$$J_{ij} = \frac{1}{acN} \sum_{\sigma=1}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu} \xi_j^{\sigma^\mu}. \tag{17}$$

Thus, synaptic weights contain only information about memory patterns for contexts in which that synapse, and the corresponding pair of neurons, are allocated.

The effective weight matrix [cf. Eq. (5)] is given by

$$\tilde{J}_{ij}^k = \frac{a_i^k a_j^k c_{ij}^k}{acN} \sum_{\sigma=1}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu} \xi_j^{\sigma^\mu} \tag{18a}$$

$$= a_i^k a_j^k \left[\frac{c_{ij}^k}{acN} \left(\sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu} + \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu} \xi_j^{\sigma^\mu} \right) \right] \tag{18b}$$

$$= a_i^k a_j^k \left(\frac{1}{aN} \sum_{\mu=1}^{\mathcal{P}} \tilde{\xi}_i^{k^\mu} \tilde{\xi}_j^{k^\mu} + \delta_{ij}^k \right). \tag{18c}$$

By calculating the number of additional memory patterns and taking the average of the gating variables (c_{ij}^k, a_i^σ, a_j^σ, and c_{ij}^σ) in Eq. (18b), we find (Appendix C)

$$\mathcal{P} = p[1 + a^4 c^4 (s-1)]. \tag{19}$$

Isolating δ_{ij}^k in Eq. (18) results in

$$\delta_{ij}^k = \frac{c_{ij}^k}{acN} \left(\sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu} + \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu} \xi_j^{\sigma^\mu} \right) - \frac{1}{aN} \sum_{\mu=1}^{\mathcal{P}} \tilde{\xi}_i^{k^\mu} \tilde{\xi}_j^{k^\mu}. \tag{20}$$

From Eq. (20) [using $p = \mathcal{P}/[1 + a^4 c^4 (s-1)]$, cf. Eq. (19)], we can write the variance of δ_{ij}^k as

$$\frac{\Delta^2}{aN} = \left\{ \frac{1 + a^2 c (s-1)}{c[1 + a^4 c^4 (s-1)]} - 1 \right\} \frac{\mathcal{P}}{a^2 N^2} = \Delta_0^2 \frac{\alpha_{\mathcal{P}}}{aN}, \tag{21}$$

arriving at

$$1 + \Delta_0^2 = \frac{1 + a^2 c (s-1)}{c[1 + a^4 c^4 (s-1)]}. \tag{22}$$

Following Eq. (8), we obtain the overlap between the attractor states and memory patterns at capacity as given by

$$m^* = m_\Delta^* \left(\frac{1 + a^2 c (s-1)}{c[1 + a^4 c^4 (s-1)]} \right). \tag{23}$$

Similarly, by following Eq. (9), we obtain a single-context capacity (in the zero-temperature limit) of

$$\alpha_{\text{cxt}}^* = \frac{\alpha_\Delta^* \left(\frac{1 + a^2 c (s-1)}{c[1 + a^4 c^4 (s-1)]} \right)}{1 + a^4 c^4 (s-1)}. \tag{24}$$

We visualized the overlap m^* and single-context capacity α_{cxt}^* first for random neuronal gating alone ($a \leq 1$, $c = 1$) and observed that both depend on the neuronal gating level and the number of contexts [Figs. 3(a)–3(d)]. As expected, the overlap is equal to that of the standard Hopfield network ($m_{\text{H}}^* = 0.97$) for $a = 1$ and $a = 0$, decreasing as the number of contexts increases for intermediate to high levels of neuronal gating [Figs. 3(a) and 3(b)]. Similarly, single-context capacity also decreases as the number of contexts increases, but this decrease becomes gradual as the level of neuronal gating increases [Figs. 3(c) and 3(d)]. To validate the mean-field analysis, we compared the theoretical results with numerical simulations. Specifically, we simulated the dynamics of networks (of size $N_{\text{cxt}} = 4000$) with increasing memory load until the average overlap approached the theoretical value [cf. Fig. 3(b); simulation details in Appendix G]. We found that simulations and theory behave similarly [Fig. 3(d)], with small discrepancies likely due to finite-size effects and other difficulties with simulating Hopfield networks [48].

From Eqs. (4), (12), and (24), we then obtained estimates for the total capacity α^* and information content I^* of networks with random neuronal gating. We found that total memory capacity and information content per synapse

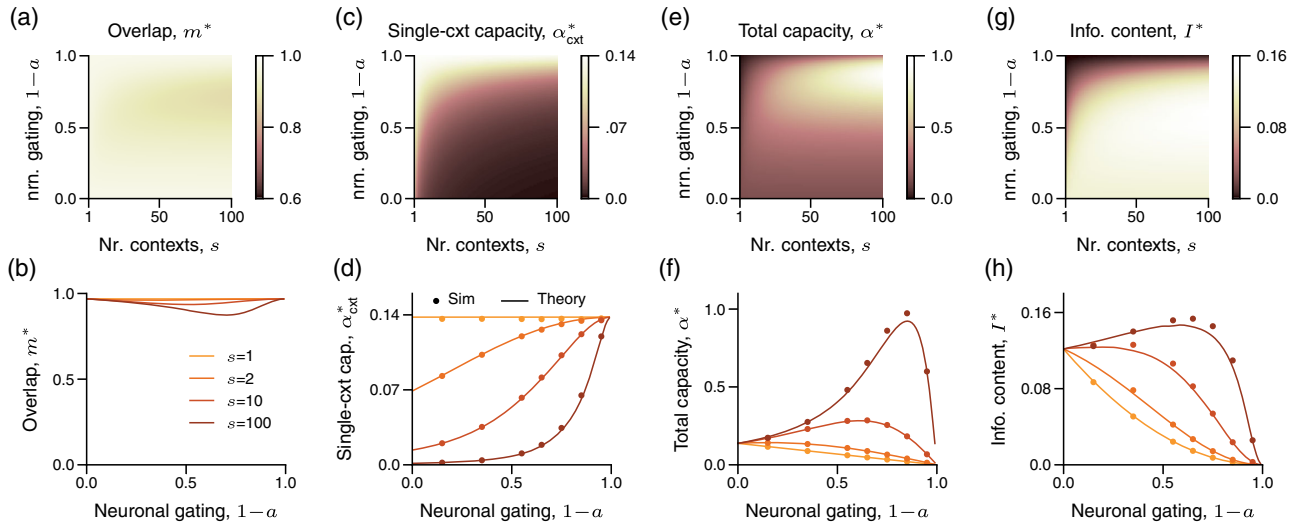


FIG. 3. Overlap, memory capacity, and information content of the context-modular memory network with random neuronal gating alone ($c = 1$). (a) Overlap estimation [color; Eq. (23)] as a function of the number of contexts s (x axis) and neuronal gating, $1 - a$ (y axis). (b) Overlap estimation [Eq. (23)] as a function of neuronal gating, $1 - a$, for four different values for the number of contexts: $s = \{1, 2, 10, 100\}$. Note that $s = 1$ corresponds to the standard Hopfield network. (c),(d) Same as (a) and (b) for single-context capacity estimation [Eq. (24)]. Dots correspond to the estimation of single-context capacity from simulations with $N_{cxt} = 4000$ (Appendix G). (e),(f) Same as (c) and (d) for total capacity estimation [Eq. (4)]. (g),(h) Same as (c) and (d) for information content per synapse [Eq. (12)]. Note that (b) does not contain numerical simulation dots, as the theoretical overlap was used as the cutoff value by which to estimate capacity and information content in (d), (f), and (h).

depended on both the gating level and the number of contexts [Figs. 3(e)–3(h)]. For high amounts of neuronal gating ($1 - a \approx 0.75$), we found that the gradual degradation of single-context capacity [Figs. 3(c) and 3(d)] leads to large increases in the total capacity, close to an order of magnitude larger than that of the standard Hopfield network for $s = 100$ contexts [Figs. 3(e) and 3(f)]. However, because the patterns themselves become relatively smaller as neuronal gating increases, the information content is generally lower than that of the standard Hopfield network [Figs. 3(g) and 3(h)], showing only modest gains for some parameter settings (e.g., $1 - a = 0.75$ and $s = 100$). We again verified the theoretical results against numerical simulations and found a close match [Figs. 3(f) and 3(h)].

We then considered overlap and capacity for the case of random synaptic gating alone ($a = 1, c \leq 1$). In contrast to random neuronal gating, we found that the overlap was modulated much more strongly by both number of contexts and synaptic gating level [Figs. 4(a) and 4(b)]. Furthermore, increasing the synaptic gating level further decreases overlap and single-context capacity [Figs. 4(c) and 4(d)] opposite to the neuronal gating case. Single-context capacity quickly drops close to zero as the number of contexts increases as well [Figs. 4(c) and 4(d)]. Overall, we found that while random synaptic gating can successfully increase the total capacity in the network, it is much less effective than random neuronal gating in doing so [Figs. 4(e) and 4(f)]. As for information content, we again found that despite some gains in capacity, the information content with random synaptic gating was similar to or less than that of the standard Hopfield network across most parameter ranges [Figs. 4(g) and 4(h)]. In this case, however, it is not due to the smaller size of the patterns, but rather the loss in information due to the low overlap values at high capacity [Figs. 4(a) and 4(b)]. As with random neuronal gating, we verified the mean-field theory for random synaptic gating by comparing to numerical simulations, and found a good correspondence [Figs. 4(d), 4(f), and 4(h)].

To understand when contextual gating was most effective, we numerically estimated the optimal gating ratio for neuronal gating alone [$c = 1$; Fig. 5(a)] and synaptic gating alone [$a = 1$; Fig. 5(b)] as a function of the number of contexts s. For both total capacity α^* and information content per synapse I^*, the optimal neuronal gating is well approximated by $1 - a_{opt} \approx 1 - (\sqrt{2s-1} - 1)/(s-1)$ [Fig. 5(a), dashed line]. This means that the larger the number of contexts s, the more neurons need to be gated for maximal memory capacity and information content per synapse. For total capacity α^* the optimal synaptic gating ratio is well approximated by $1 - c_{opt} \approx 1 - \sqrt{1/s}$ [Fig. 5(b), dashed line]. Similar to neuronal gating, this means that larger number of contexts s requires more synapses to be gated. For information content per synapse I^* the optimal synaptic gating ratio saturates at $1 - c_{opt} \approx 1/2$ [Fig. 5(b), light purple]. The more synapses are gated, the smaller the overlap between attractors and memory patterns [cf. Figs. 4(a) and 4(b)], and this decreases information content per synapse considerably

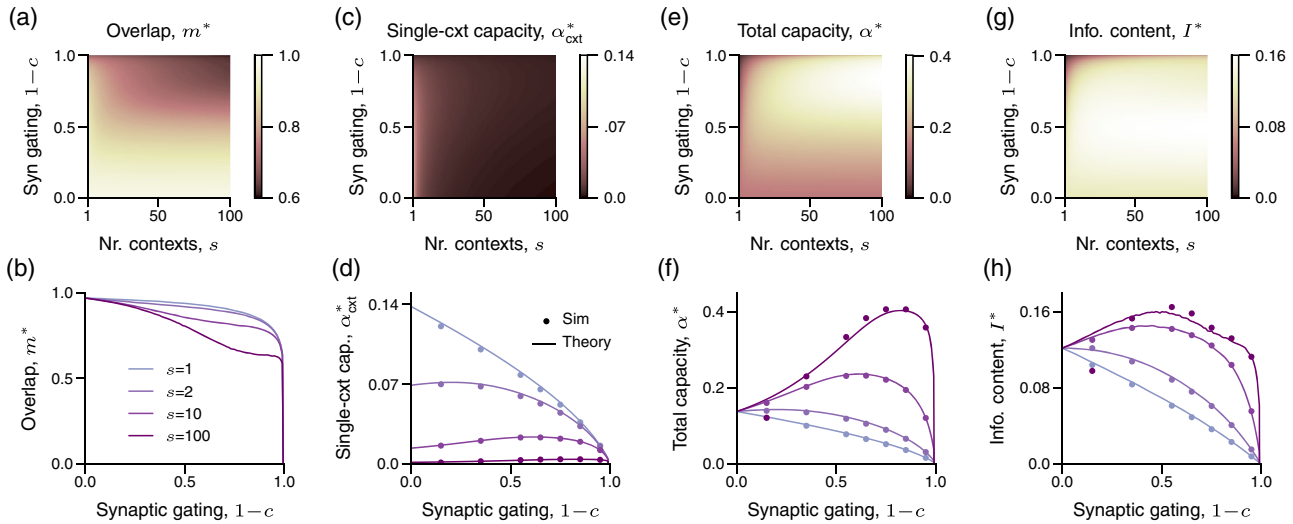


FIG. 4. Overlap, memory capacity, and information content of the context-modular memory network with random synaptic gating alone ($a = 1$). (a) Overlap estimation [color; Eq. (23)] as a function of the number of contexts s (x axis) and synaptic gating, $1 - c$ (y axis). (b) Overlap estimation [Eq. (23)] as a function of synaptic gating, $1 - c$, for four different values for the number of contexts: $s = \{1, 2, 10, 100\}$. Note that $s = 1$ corresponds to the standard Hopfield network. (c),(d) Same as (a) and (b) for single-context capacity estimation [Eq. (24)]. Dots correspond to the estimation of single-context capacity from simulations with $N_{\rm cxt} = 4000$ (Appendix G). (e),(f) Same as (c) and (d) for total capacity estimation [Eq. (4)]. (g),(h) Same as (c) and (d) for information content per synapse [Eq. (12)]. Note that (b) does not contain numerical simulation dots, as the theoretical overlap was used as the cutoff value by which to estimate capacity and information content in (d), (f), and (h).

[cf. Eqs. (11) and (12)]. This indicates the existence of a trade-off between the number of synapses gated via synaptic gating ratio, $1 - c$, and the maximal information content per synapse I^* at capacity.

Given that high capacity is achieved with substantial amounts of gating, we next checked the extent to which contexts have overlapping connectivity, which we refer to as the number of shared contexts per synapse. Since gating is defined randomly and independently for each context, the probability that a given synapse will participate in n contexts is defined by a binomial distribution $\mathcal{B}(n; s, a^2c)$ [49]. We can thus define the expected value of the shared contexts per synapse as sa^2 for neuronal gating alone, and sc for synaptic gating alone, which we visualize for several values of s ($s = \{1, 2, 10, 100\}$) in Figs. 5(c) and 5(d). For $s = 10$ and $s = 100$, we found a substantial amount of sharing between contexts both for neuronal and synaptic gating. Using the optimal gating values [cf. Figs. 5(a) and 5(b)], we then considered the amount of shared contexts per synapse at optimal gating [Figs. 5(c) and 5(d), square points], which could be well approximated as $1 + (1 - a)^2$ and $1/c$ for neuronal and synaptic gating, respectively. We thus found that the number of shared contexts was low, around 1–2, for optimal neuronal gating, but grew to much larger values, above 10 for $s = 100$, for optimal synaptic gating. This suggests that the effectiveness of random neuronal gating is in part due to the sparsity of the memory representations, which prevent interference between different contexts, similar to previous models [6] (see Sec. V). However, considering the full binomial distribution of shared contexts [Fig. 5(c), inset; $s = 10$, $a = 0.4$] revealed that over 50% of synapses still shared more than one context (2–5 contexts). This makes the model distinct from the case in which each context would use a separate set of synapses, which would yield no gains in capacity. Furthermore, we also see that nearly 20% of synapses remain unused (i.e., store 0 contexts), suggesting that random neuronal gating may still not be using all synapses in the most effective way (see Sec. V).

In addition to memory capacity and information content, we also calculated the memory accessibility of memories from both active (i.e., accessible memories) and inactive contexts (i.e., inaccessible memories). Following Eqs. (15) and (16), we arrive at mean stability of accessible and inaccessible memory patterns given by (Appendix F),

$$\bar{\kappa}_{\rm acc} = \sqrt{\frac{c}{[1 + a^2c(s-1)]\alpha^*_{\rm cxt}}} \tag{25}$$

and

$$\bar{\kappa}_{\rm inacc} = a^2c\sqrt{\frac{c}{[1 + a^2c(s-1)]\alpha^*_{\rm cxt}}} = a^2c\bar{\kappa}_{\rm acc}, \tag{26}$$

respectively. For neuronal gating alone ($c = 1$), the mean accessibility $\bar{\kappa}_{\rm acc}$ increases as the proportion of gated neurons, $1 - a$, increases, slightly decreasing with the number of contexts s [Fig. 6(a)]. Conversely, the mean inaccessibility decreases with $1 - a$ and s [Fig. 6(b)]. The effect is very similar for synaptic gating alone [$a = 1$;

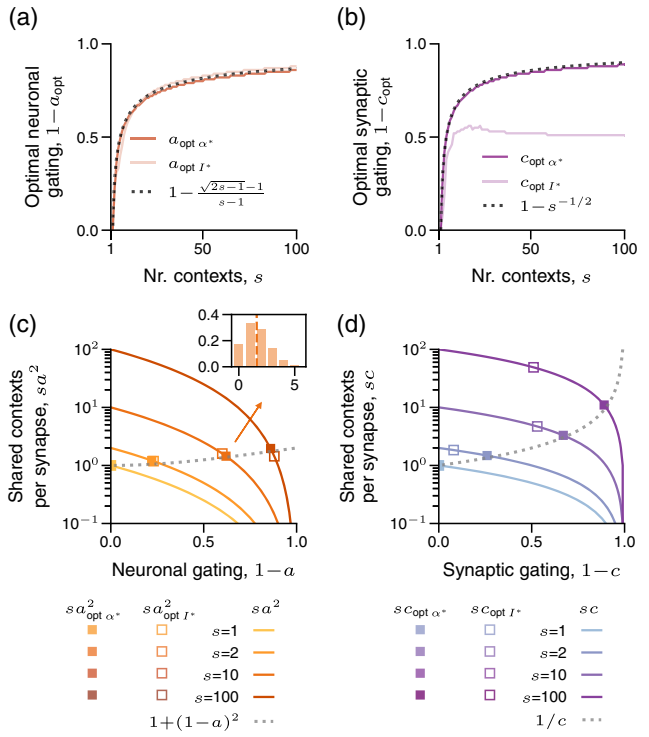


FIG. 5. Optimal gating values and number of shared contexts per synapse for random neuronal and synaptic gating. (a) Optimal neuronal gating ratio for maximal capacity, $1-a_{\mathrm{opt},a^*}$, or for maximal information content, $1-a_{\mathrm{opt},I^*}$, as a function of the number of contexts s for neuronal gating alone ($c=1$). Full lines are numerically estimated from Figs. 3(e)–3(h). Both curves are well approximated by the dotted black line, corresponding to $1-(\sqrt{2s-1}-1)/(s-1)$. (b) Same as (a) but for optimal synaptic gating ratios $1-c_{\mathrm{opt},a^*}$ and $1-c_{\mathrm{opt},I^*}$ as estimated from Figs. 4(e)–4(h) for synaptic gating alone ($a=1$). In this case $1-c_{\mathrm{opt},a^*}$ is well approximated by the dotted black curve, $1-\sqrt{1/s}$. (c),(d) Shared contexts per synapse for neuronal gating alone [(c); defined as sa^2] and for synaptic gating alone [(d); defined as sc] for four different values of s. Closed and open square points correspond to the amount of shared contexts per synapse at the optimal gating levels from (a) and (b), and the gray dotted curves correspond to $sa^2_{\mathrm{opt}}=1+(1-a)^2$ and $sc_{\mathrm{opt}}=1/c$. (c) Inset: an example of the full probability distribution of the number of shared contexts per synapse, with the mean shown as the vertical dashed line, for $s=10$ contexts and optimal neuronal gating at $a=0.4$, defined as a binomial distribution $\mathcal{B}(s,a^2)$.

Figs. 6(c) and 6(d)]. This indicates that, when using random neuronal and synaptic gating, to properly separate accessible and inaccessible memories, a considerable amount of neurons and/or synapses have to be gated, as clearly indicated by the ratio between the stability of accessible and inaccessible memories:

$$\frac{\bar{\kappa}_{\mathrm{acc}}}{\bar{\kappa}_{\mathrm{inacc}}}=\frac{1}{a^2c}. \tag{27}$$

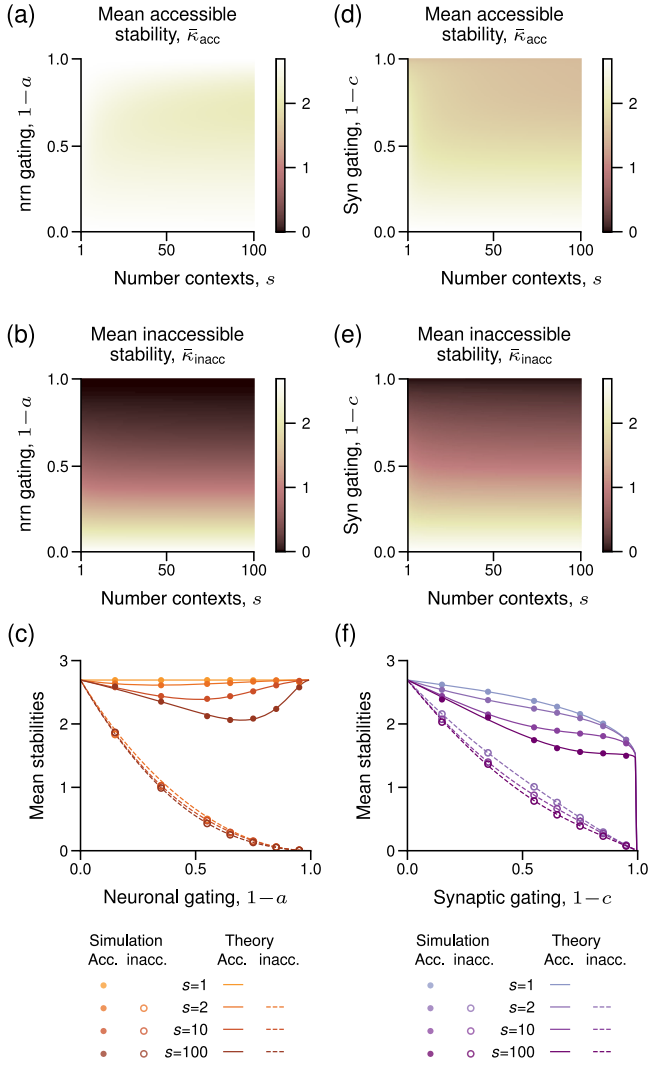


FIG. 6. Stability of accessible and inaccessible memories with random neuronal and synaptic gating. (a) Average memory stability for accessible memories $\bar{\kappa}_{\mathrm{acc}}$ [color; Eq. (25)], as a function of the number of contexts s (x axis) and neuronal gating ratio, $1-a$ (y axis), for neuronal gating alone ($c=1$). (b) Average memory stability for inaccessible memories $\bar{\kappa}_{\mathrm{inacc}}$ [color; Eq. (26)], as a function of the number of contexts s (x axis), and neuronal gating ratio, $1-a$ (y axis), for neuronal gating alone ($c=1$). (c) Average memory stability for accessible (full lines and closed circles) and inaccessible (dashed lines and open circles) memories as a function of neuronal gating, $1-a$, for neuronal gating alone ($c=1$) and four different values for the number of contexts: $s=\{1,2,10,100\}$. Note that $s=1$ corresponds to the standard Hopfield network. Dots correspond to the estimation of single-context capacity from simulations with $N_{\mathrm{cxt}}=4000$ (Appendix G). (d)–(f) Same as (a)–(c) for synaptic gating alone ($a=1$): y axis in (d),(e) and x axis in (f) show the synaptic gating ratio, $1-c$.

We numerically confirmed the theoretical values for mean stability of accessible and inaccessible memories from Eqs. (25) and (26) at full capacity [Figs. 6(c) and 6(f); see also Appendix G]. Overall, we have thus demonstrated

that memory patterns in the context-modular memory network do indeed have dynamic memory stability—depending on the active context, accessible memories will be highly accessible (high stability), and inaccessible memories will be much less accessible (low stability), and may be altogether hidden for large neuronal and synaptic gating ratios.

IV. SYNAPTIC REFINEMENT

We now present a more selective way of imposing synaptic control which we refer to as synaptic *refinement*. In this framework, the synaptic gating mask is determined only after the memory patterns are stored, so we consider that synaptic gating can act only on the recall dynamics [Eq. (1)]. More precisely, we find an optimal set of synaptic gates c_{ij}^{σ} to stabilize the subset of memory patterns in each context.

To start with, we define a hypothetical "isolated" weight matrix for each individual context, e.g., context σ, with elements J_{ij}^{σ},

$$J_{ij}^{\sigma} = \frac{a_i^{\sigma} a_j^{\sigma}}{aN} \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu} \xi_j^{\sigma^\mu}, \tag{28}$$

where $\xi_i^{\sigma^\mu}$ defines the memory patterns and a_i^{σ} defines neuronal gating variables, both being randomly drawn from a uniform distribution as in the random neuronal gating case above. Based on J_{ij}^{σ} we can thus optimize the gating state of each synapse in context σ as

$$c_{ij}^{\sigma} = \begin{cases} 0 & \text{if } \left(J_{ij} J_{ij}^{\sigma}\right) < 0 \\ 1 & \text{otherwise.} \end{cases} \tag{29}$$

This method ensures that the sign of each ungated weight reflects the sign of the correlation between neurons i and j over patterns in context σ. Importantly, the synaptic gating ratio, $1 - c$, is no longer a predetermined parameter of the network. Moreover, the optimization is agnostic about the structure of the weight matrix J_{ij}, and we consider stabilization of memory patterns in three scenarios with distinct construction of J_{ij}. First, we consider the case in which each set of memories is assigned to a different context (Sec. IV A). Second, we then consider that memories can be shared across different contexts (Sec. IV B). Finally, we explore the scenario in which the weight matrix has random symmetric entries taken from a Gaussian distribution (Sec. IV C).

A. Stabilization of a Hebbian weight matrix

Following the random neuronal and synaptic gating methods, we build a weight matrix (including random neuronal gating) as

$$J_{ij} = \sum_{\sigma=1}^{s} J_{ij}^{\sigma} = \frac{1}{aN} \sum_{\sigma=1}^{s} a_i^{\sigma} a_j^{\sigma} \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu} \xi_j^{\sigma^\mu}, \tag{30}$$

where J_{ij}^{σ} is defined in Eq. (28). Note that we did not add a normalizing constant that reflects the effective dilution when synaptic gating is applied in the dynamics [e.g., the factor c in the denominator of Eq. (17)] allowing a straightforward numerical implementation of the network dynamics. Instead, we only use a normalizing constant when estimating both memory overlap and capacity based on the effective weight matrix (details below). To theoretically analyze this model, we separate the overall weight matrix into two terms corresponding to the active context k and the inactive contexts, $\sigma \neq k$, such that

$$\begin{aligned} a_i^k a_j^k J_{ij} &= a_i^k a_j^k \left(J_{ij}^k + \sum_{\substack{\sigma=1, \\ \sigma \neq k}}^{s} J_{ij}^{\sigma} \right) \\ &= \underbrace{\frac{a_i^k a_j^k}{aN} \sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu}}_{} + \underbrace{\frac{a_i^k a_j^k}{aN} \sum_{\substack{\sigma=1, \\ \sigma \neq k}}^{s} a_i^{\sigma} a_j^{\sigma} \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu} \xi_j^{\sigma^\mu}}_{} \\ &= \qquad \mathcal{X}_{ij}^k \qquad + \qquad \mathcal{Y}_{ij}^k. \end{aligned} \tag{31}$$

In Eq. (31), we rewrite the sum over patterns as two Gaussian noise terms, $\mathcal{X}_{ij}^k \sim \mathcal{N}(0, \sigma_x^2)$ and $\mathcal{Y}_{ij}^k \sim \mathcal{N}(0, \sigma_y^2)$. Considering only the active neurons in context k, these terms have variance

$$\sigma_x^2 = \frac{p}{a^2 N^2} = \frac{\alpha_{\text{cxt}}}{aN} \quad \text{and} \quad \sigma_y^2 = a^2 (s-1) \sigma_x^2. \tag{32}$$

Following the criterion in Eq. (29), we can estimate the overall proportion of gated synapses by calculating the probability that $|\mathcal{Y}_{ij}^k| > |\mathcal{X}_{ij}^k|$ and $\mathcal{Y}_{ij}^k \mathcal{X}_{ij}^k < 0$. We thus find the proportion of gated synapses as a function of the neuronal gating parameter a and number of contexts s as

$$\begin{aligned} 1 - c &= \frac{1}{2\pi\sigma_x\sigma_y} \left[\int_0^{\infty} \mathrm{e}^{-x^2/(2\sigma_x^2)} \int_{-\infty}^{x} \mathrm{e}^{-y^2/(2\sigma_y^2)} \mathrm{d}y \mathrm{d}x \right. \\ &\quad \left. + \int_{-\infty}^{0} \mathrm{e}^{-x^2/(2\sigma_x^2)} \int_{x}^{\infty} \mathrm{e}^{-y^2/(2\sigma_y^2)} \mathrm{d}y \mathrm{d}x \right] \\ &= \frac{1}{\pi} \arctan\left(\frac{\sigma_y}{\sigma_x}\right) = \frac{1}{\pi} \arctan\left(a\sqrt{s-1}\right). \end{aligned} \tag{33}$$

Importantly, in the limit of infinite contexts, the proportion of gated synapses is $1/2$ (i.e., 50% of synapses are gated):

$$\lim_{a^2 s \to \infty} (1 - c) = \frac{1}{2}. \tag{34}$$

We observe that the synaptic gating ratio approaches this limit as the number of contexts increases and the number of

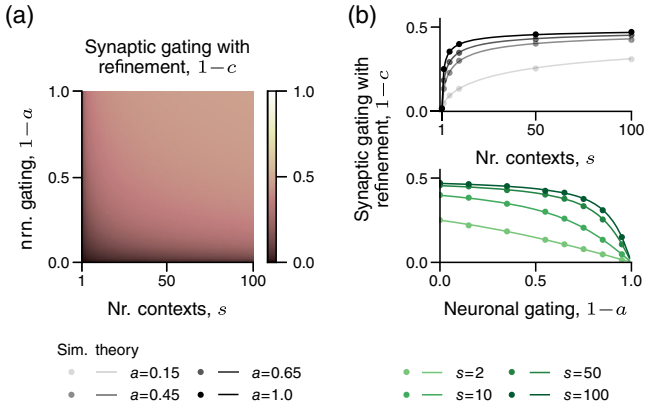


FIG. 7. Synaptic gating ratio following synaptic refinement. (a) Estimation of the resulting synaptic gating ratio, $1-c$ [color; Eq. (33)], following synaptic refinement as a function of the number of contexts s (x axis) and neuronal gating, $1-a$ (y axis). (b) Synaptic gating ratio as a function of number of contexts s, and four different values for neuronal gating, $1-a=\{0, 0.35, 0.55, 0.85\}$ (top), and as a function of neuronal gating, $1-a$, and four different values for the number of contexts, $s=\{2, 10, 50, 100\}$ (bottom). Dots correspond to the estimation of synaptic gating ratio from simulations with $N_{\mathrm{cxt}}=4000$ (Appendix G).

gated neuron decreases (Fig. 7). The intuition for this is straightforward—when enough noise has been added to the weight matrix, each element has the desired sign for approximately half of the connections. Notably, this also implies that the number of contexts (and thus patterns) that the network stores can potentially be increased indefinitely, without affecting the fact that approximately half of the synapses will still reflect the appropriate correlation structure of the accessible memory patterns. We compare our results with simulations of weight matrices of $N=4000$ and confirm that our calculation for the overall connectivity holds [Fig. 7(b)].

To calculate memory capacity, we first write the effective weight matrix [cf. Eq. (5)] as

$$\tilde{J}_{ij}^{k}=\frac{a_i^k a_j^k c_{ij}^k}{\mathcal{J}^k}J_{ij}=\frac{a_i^k a_j^k c_{ij}^k}{\mathcal{J}^k}(\mathcal{X}_{ij}^k+\mathcal{Y}_{ij}^k)$$
$$=a_i^k a_j^k\left(\frac{1}{aN}\sum_{\mu=1}^{p}\xi_i^{k\mu}\xi_j^{k\mu}+\delta_{ij}^k\right), \tag{35}$$

so that

$$\delta_{ij}^k=\frac{c_{ij}^k}{\mathcal{J}^k}(\mathcal{X}_{ij}^k+\mathcal{Y}_{ij}^k)-\frac{1}{aN}\sum_{\mu=1}^{p}\xi_i^{k\mu}\xi_j^{k\mu}, \tag{36}$$

where $\mathcal{J}^k$ is a normalization term that accounts for the dilution created by the synaptic refinement through c_{ij}^k. Note that we do not take into account any additional memorylike pattern terms, because they are assumed to be negligible with optimal gating via refinement, and thus $\mathcal{P}=p$. During recall in a particular context k with synaptic refinement, each gating variable c_{ij}^k introduces an effective correlation between the patterns within context [described by $\mathcal{X}_{ij}^k$ in Eq. (31)] and the weight matrix [described by $\mathcal{X}_{ij}^k+\mathcal{Y}_{ij}^k$ in Eq. (31)]. This correlation (and normalization) term is given by

$$\mathcal{J}^k=\frac{\left\langle c_{ij}^k\mathcal{X}_{ij}^k(\mathcal{X}_{ij}^k+\mathcal{Y}_{ij}^k)\right\rangle_{ij}}{\left\langle(\mathcal{X}_{ij}^k)^2\right\rangle_{ij}}, \tag{37}$$

where $\langle\cdot\rangle_{ij}$ represents the average over all connections (i.e., pairs i, j). We thus arrive at the variance of δ_{ij}^k given by

$$\frac{\left(\Delta^k\right)^2}{aN}=\frac{\left\langle\left(\mathcal{X}_{ij}^k+\mathcal{Y}_{ij}^k\right)^2 c_{ij}^k\right\rangle_{ij}}{(\mathcal{J}^k)^2}-\frac{\alpha_{\mathrm{cxt}}}{aN}=\Delta_0^2\frac{\alpha_{\mathrm{cxt}}}{aN}, \tag{38}$$

where (Appendix D)

$$1+\Delta_0^2=\frac{[1+a^2(s-1)]\pi^2 c+\pi a\sqrt{s-1}}{(\pi c+a\sqrt{s-1})^2}. \tag{39}$$

Combining Eqs. (8) and (39), we arrive at the overlap at capacity given by

$$m^*=m_\Delta^*\left(\frac{[1+a^2(s-1)]\pi^2 c+\pi a\sqrt{s-1}}{(\pi c+a\sqrt{s-1})^2}\right), \tag{40}$$

and combining Eqs. (9) and (39), we can estimate the single-context capacity as

$$\alpha_{\mathrm{cxt}}^*=\alpha_\Delta^*\left(\frac{[1+a^2(s-1)]\pi^2 c+\pi a\sqrt{s-1}}{(\pi c+a\sqrt{s-1})^2}\right), \tag{41}$$

where c is given by Eq. (33). Furthermore and importantly, when we assume the limit of an infinite number of contexts s, the overlap and single-context capacity become

$$\lim_{a^2 s\to\infty} m^*=m_\Delta^*\left(\frac{\pi^2}{2}\right)\quad\text{and}\quad\lim_{a^2 s\to\infty}\alpha_{\mathrm{cxt}}^*=\alpha_\Delta^*\left(\frac{\pi^2}{2}\right). \tag{42}$$

Both overlap and single-context memory capacity quickly approach the limit values [Eq. (42)] for refinement alone ($a=1$) when increasing the number of contexts s [Figs. 8(a) and 8(b)]. This results in total capacity and information content per synapse that increase roughly linearly with the number of contexts s [Figs. 8(c) and 8(d)], growing to large values ($\alpha^*\approx 5$ for $s=100$ contexts)—up to 30 times that of the standard Hopfield network [Fig. 8(c)]. Although both overlap and single-context memory capacity increase as the proportion of gated

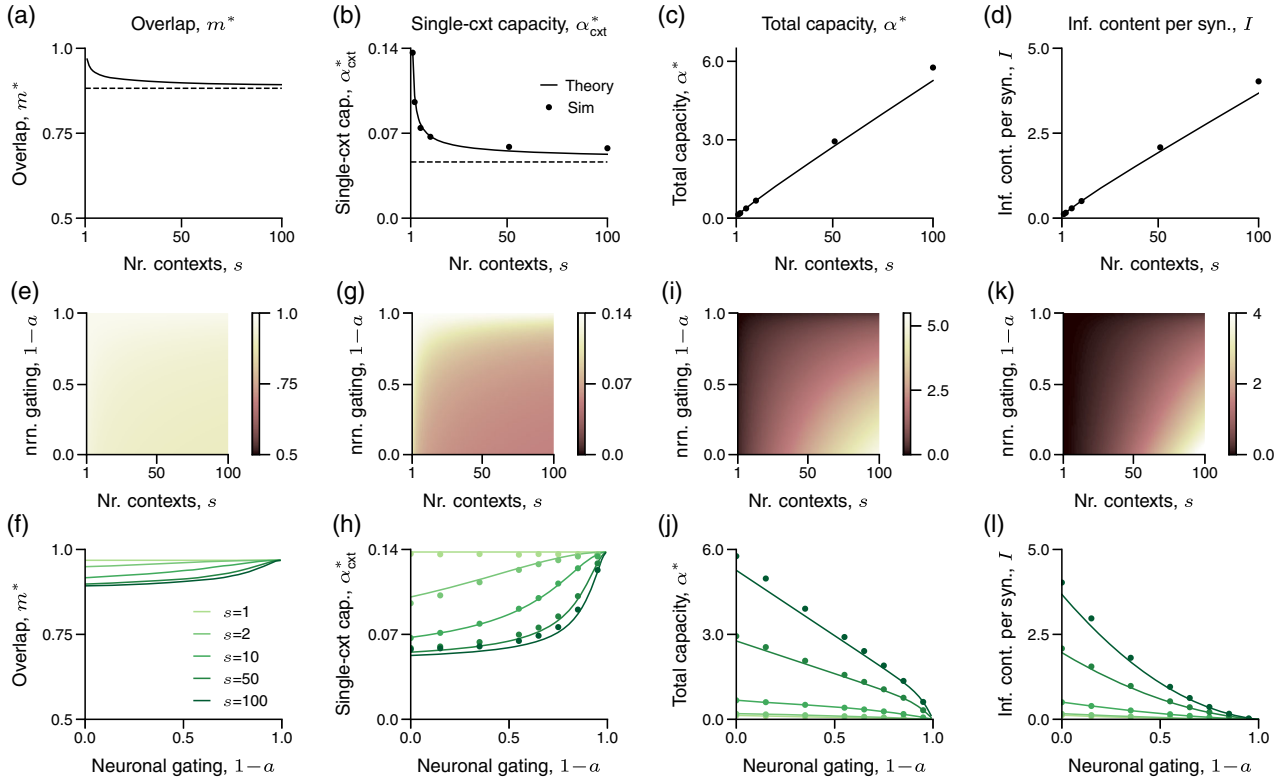


FIG. 8. Overlap, memory capacity, and information content per synapses following synaptic refinement. (a)–(d) Estimation of overlap m^* [(a); Eq. (40)], single-context capacity α^*_{cxt} [(b); Eq. (41)], total capacity α^* [(c); Eq. (4)], and information content per synapse I^* [(d); Eq. (12)], as a function of the number of contexts s for synaptic refinement alone ($a = 1$). (e) Estimation (color) of overlap m^* [Eq. (40)], as a function of the number of contexts s (x axis) and neuronal gating ratio, $1 - a$ (y axis). (f) Estimation of overlap m^* [Eq. (40)], as a function of neuronal gating, $1 - a$, and five different values for the number of contexts: $s = \{1, 2, 10, 50, 100\}$. (g),(h) Same as (e) and (f) for single-context capacity α^*_{cxt} [Eq. (41)]. (i),(j) Same as (e) and (f) for total capacity α^* [Eq. (4)]. (k),(l) Same as (e) and (f) for information content per synapse I^* [Eq. (12)]. Note that $s = 1$ corresponds to the standard Hopfield network. Dots in (b)–(d), (h), (j), and (l) correspond to the estimations from simulations with $N_{\text{cxt}} = 4000$ (Appendix G). As before, we note that (a) and (f) do not contain numerical simulation dots, as the theoretical overlap was used as the cutoff value by which to estimate capacity and information content.

neurons grows [Figs. 8(e)–8(h)], total memory capacity and information context are maximized when no neuron is gated [Figs. 8(i)–8(l)]. This results in the optimal neuronal gating being zero ($a_{\text{opt}} = 1$), and so the addition of neuronal gating to synaptic refinement does not further increase memory capacity nor information content. Numerical simulations confirmed the theory both for synaptic refinement alone [Figs. 8(b)–8(d)] and in combination with random neuronal gating [Figs. 8(h), 8(j), and 8(l)].

The average memory stability of both accessible and inaccessible memories can be calculated for synaptic refinement by adapting a signal-to-noise analysis to the theoretical insights gained from Eqs. (31)–(42). Doing so (Appendix F), we arrive at an average stability for accessible memories of

$$\bar{\kappa}_{\text{acc}} = \frac{\pi c + a\sqrt{s-1}}{\sqrt{\left[\pi c + \pi a^2 c(s-1) + a\sqrt{s-1}\right]\pi\alpha^*_{\text{cxt}}}}, \quad (43)$$

and an average stability of inaccessible memories of

$$\bar{\kappa}_{\text{inacc}} = \frac{a^2 c\sqrt{\pi}}{\sqrt{\left(\pi c + \pi a^2 c(s-1) + a\sqrt{s-1}\right)\alpha^*_{\text{cxt}}}}, \quad (44)$$

where c is given by Eq. (33). When taking the limit of infinite contexts, we arrive at the stability of accessible and inaccessible memories given by

$$\lim_{a^2 s\to\infty} \bar{\kappa}_{\text{acc}} = \sqrt{\frac{2}{\pi^2 \alpha^*_\Delta\left(\frac{\pi^2}{2}\right)}} \quad \text{and} \quad \lim_{a^2 s\to\infty} \bar{\kappa}_{\text{inacc}} = 0. \quad (45)$$

Similarly to single-context memory capacity, the stability of accessible memories increases with the number of gated synapses, quickly approaching the limit given by Eq. (45) as the number of contexts s increases [Fig. 9(a)]. The stability of inaccessible memories decreases rapidly with an increase in the number of contexts s and proportion of gated neurons, $1 - a$ [Fig. 9(b)]. Even for little to no

neuronal gating, $1-a \ll 1$, synaptic refinement alone is successful at reducing the inaccessible memory stability compared to accessible ones, as long as s is sufficiently large [Figs. 9(b)–9(d)]. This demonstrates that for networks with a low overall load, synaptic refinement will be unable to provide dynamic accessibility unless neuronal gating is also implemented, as the greater the number of gated neurons, the higher (lower) the stability of accessible (inaccessible) memories (Fig. 6). Counterintuitively, this suggests that for synaptic refinement, a noisier, less correlated, weight matrix actually facilitates better control of dynamic memory accessibility, especially when neuronal gating is not present ($a = 1$).

The theoretical values of both accessible and inaccessible average stability were confirmed numerically for networks of size $N = 4000$ neurons [Figs. 9(c) and 9(d)]. Overall, we thus see that synaptic refinement, with or without random neuronal gating, can support dynamic memory accessibility provided that the network load is large enough.

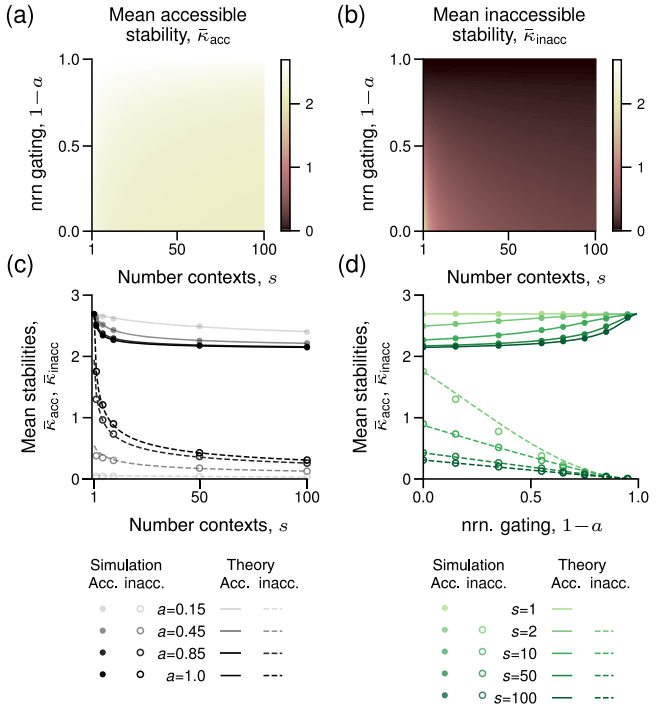


FIG. 9. Stability of accessible memories and inaccessible memories following synaptic refinement. (a),(b) Estimation of the average memory stability (color) of accessible [(a); Eq. (43)] and inaccessible [(b); Eq. (44)] memories as a function of the number of contexts s (x axis) and neuronal gating ratio, $1-a$ (y axis). (c) Average memory stability of accessible (full lines and closed circles) and inaccessible (dashed lines and open circles) memories as a function of the number of contexts s and four different values for neuronal gating ratio: $1-a = \{0, 0.15, 0.55, 0.85\}$. (d) Average memory stability of accessible (full lines and closed circles) and inaccessible (dashed lines and open circles) memories as a function of neuronal gating ratio, $1-a$, and five different values for the number of contexts: $s = \{1, 2, 10, 50, 100\}$. Note that $s = 1$ corresponds to the standard Hopfield network. Dots in (c) and (d) correspond to the estimations from simulations with $N = 4000$ (Appendix G).

B. Stabilization of shared memories

We now consider a variant of synaptic refinement in which we relax some of the initial assumptions of our model, and momentarily take a different perspective on context-dependent memory. Importantly, we break the requirement that each context has a set of unique, non-overlapping memory patterns. Since synaptic refinement is applied only to the network's recall dynamics, memories do not require this strict grouping. We therefore begin with a general pool of P memories, which are stored without explicitly specifying contextual states,

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{P} \tilde{\xi}_i^{\mu}\tilde{\xi}_j^{\mu}, \tag{46}$$

where $\tilde{\xi}_i^{\mu}$ $(i = 1, \ldots, N$ and $\mu = 1, \ldots, P)$ are context-independent memory patterns defined by $\Pr(\tilde{\xi}_i^{\mu} = \pm 1) = 1/2$. Next, we assign an arbitrary group of these memories, of size p, to be part of a context σ, and determine the appropriate contextual configurations, i.e., the set of gating variables c_{ij}^{σ} [Eq. (29)] that enables successful memory recall. This is possible only because all neurons remain ungated in each context, even if individual connections are gated off, which means that $a = 1$.

We explored the behavior of the network when we vary the total number of patterns stored in the network P, and thus the total load α. By taking the theoretical results for overlap m^*, synaptic gating $1-c$, single-context capacity α_{cxt}^*, and memory stability, $\bar{\kappa}_{acc}$ and $\bar{\kappa}_{inacc}$, as a function of the number of contexts s (cf. Figs. 8 and 9), we can plot the aforementioned quantities as a function of $\alpha = s\alpha_{cxt}^*$ (Fig. 10). We see that arbitrary sets of memory patterns can indeed be reliably retrieved with high overlap m^* [Fig. 10(a)], provided that the number of patterns stored in the active context is sufficiently small compared to the total load [Fig. 10(c)]. Intuitively, the single-context capacity α_{cxt}^* decreases with the increase in total load α [Fig. 10(c)]. As before, synaptic refinement increases the number of gated synapses during recall of memories in the active context as the total memory load is increased [Fig. 10(b)]. Importantly, the stability of the accessible memories remain high while the stability of inaccessible memories decreases as the total number of memories is increased [Fig. 10(d)]. This demonstrates that synaptic refinement can continue to be used following learning to dynamically adjust which memories are attributed to a particular context, even if some of these memories are already allocated to other contexts (see Sec. V).

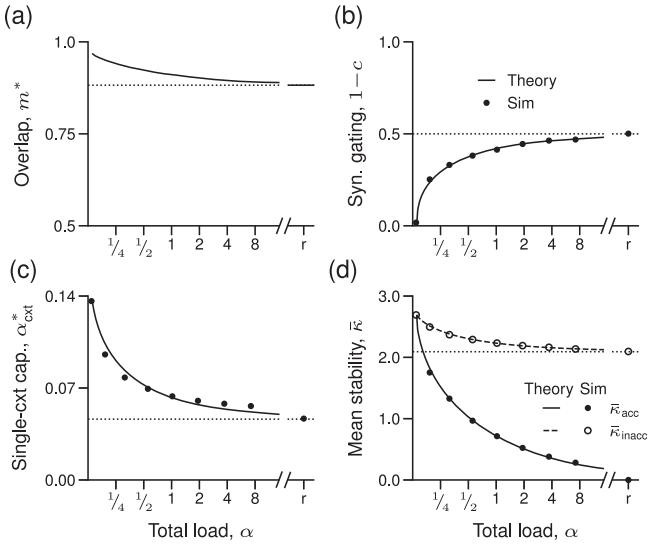


FIG. 10. Shared memory allocation through synaptic refinement. (a) Estimation of the overlap m^* [Eq. (40) with $\alpha = s\alpha^*_{\text{cxt}}$], as a function of the total load, $\alpha = P/N$. Dashed line shows the asymptotic value as α approaches infinity, which is equivalent to a a random Gaussian weight matrix [Eq. (51)]. (b)–(d) Same as (a) for synaptic gating ratio [(b); Eq. (33) with $\alpha = s\alpha^*_{\text{cxt}}$], single-context capacity [(c); Eq. (41) with $\alpha = s\alpha^*_{\text{cxt}}$], as well as stability of accessible and inaccessible memories [(d); Eqs. (43) and (44) with $\alpha = s\alpha^*_{\text{cxt}}$]. Dots correspond to the estimations from simulations with $N = 4000$ (Appendix G). Results for a weight matrix with symmetric elements drawn from a Gaussian distribution are shown as circles and indicated in the x axis by "r".

C. Stabilization of memory patterns in a random weight matrix

Given that synaptic gating modifies the energy landscape through a set of gating variables, even a randomly defined weight matrix may be utilized to recall memory patterns as long as the correct synaptic refinement is imposed. As a proof of concept, we thus generate a random weight matrix given by

$$J_{ij} = \eta_{ij}, \tag{47}$$

where η_{ij} is a symmetric random matrix with elements drawn from a Gaussian distribution with unitary variance,

$$\eta_{ij} \sim \mathcal{N}(0,1), \quad \eta_{ij} = \eta_{ji}, \quad \forall\, i,j \text{ and } \eta_{ii} = 0, \quad \forall\, i. \tag{48}$$

We then construct a set of p memory patterns for a hypothetical active context k as $\xi_i^{k\mu}$ $(i = 1,\ldots,N$ and $\mu = 1,\ldots,p)$, and define the gating variables c_{ij}^k for the active context k following Eq. (29).

The effective weight matrix for the active context k [cf. Eq. (5)] can be written as

$$\tilde{J}_{ij}^k = \frac{c_{ij}^k}{\mathcal{J}^k} J_{ij} = \left(\frac{1}{aN} \sum_{\mu=1}^{p} \xi_i^{k\mu} \xi_j^{k\mu} + \delta_{ij}^k \right), \tag{49}$$

where $\mathcal{J}^k$ is a normalization term used to account for the dilution created by the synaptic refinement. We assume no spurious memorylike pattern terms ($\mathcal{P} = p$). The noise term δ_{ij}^k is thus

$$\delta_{ij}^k = \frac{c_{ij}^k J_{ij}}{\mathcal{J}^k} - \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^{k\mu} \xi_j^{k\mu}. \tag{50}$$

From the variance of δ_{ij}^k (Appendix E) we arrive at overlap and context-specific memory capacity,

$$m^* = m^*_\Delta\left(\frac{\pi^2}{2}\right) \equiv m^*_{\text{rand}} \text{ and } \alpha^*_{\text{cxt}} = \alpha^*_\Delta\left(\frac{\pi^2}{2}\right) \equiv \alpha^*_{\text{rand}}, \tag{51}$$

which are equivalent to the limit of infinite contexts in the previous section [Eq. (42)]. The stability of accessible and inaccessible memories are

$$\bar{\kappa}_{\text{acc}} = \sqrt{\frac{2}{\pi^2 \alpha_{\text{rand}}}} \quad \text{and} \quad \bar{\kappa}_{\text{inacc}} = 0, \tag{52}$$

which are also equivalent to the limit of infinite context used in the previous section [Eq. (45)]. These results were confirmed numerically [Figs. 10(c) and 10(d)].

Remarkably, a random weight matrix can be successfully used to represent stable context-dependent memories, as long as the single-context load is less than α^*_{rand}. Intuitively, this is possible because a random weight matrix will have, on average, half of its elements with the correct sign according to the correlation structure of the desired subset of patterns [Fig. 10(b)]. This means that synaptic refinement can be used to impose any activity pattern as a stable attractor independent of the weight matrix; i.e., groups of memories are stored in the synaptic gating structure. Note that this requires the weight matrix to have a minimal amount of structure (e.g., randomly distributed with zero mean). This suggests that synaptic weights can be corrupted substantially without affecting performance, as long as the synaptic gating structure remains intact. Furthermore, it enables reliable recall with bounded (e.g., binary) synapses, an important consideration for biological memory [50]. In other words, by precisely controlling the synaptic gating variables, it is possible to recall a set of memories from a weight matrix that alone does not contain any information about those memories, but only when combined with the right mask.

V. DISCUSSION

The growing consensus perspective in neuroscience and psychology views biological memory not as a static storage device, but rather as a dynamic and flexible system [20,25]. Such a system adapts to current behavioral states, i.e., contexts, and is thus able to modify the understanding of

past experiences when needed [51]. Here, we revisited classic theoretical work on autoassociative memory in light of this perspective. Specifically, we proposed a novel context-modular memory network model in which the memory patterns and network connectivity are modified through neuronal and synaptic gating by a background contextual state. We found that such a scheme not only results in increased memory capacity, but is also effective at modulating the stability of stored patterns, thereby selecting a subset of memories for recall and reducing interference from others. Improved memory capacity is thus not the main feature of the context-modular memory network, but a side effect of increased control and flexibility which is traded off with the proportion of accessible memories at any given time.

We explored two contextual gating schemes: neuronal and synaptic. To implement these schemes, an inhibitory mask is imposed onto the recurrent connectivity to modulate the energy function of the network dynamics. Neuronal gating acts to remove particular neurons from the active context, thereby shrinking the effective network size and removing interference from masked neurons. The potential biological mechanisms for neuronal gating are numerous, including perisomatic inhibition [52] and modulation of intrinsic excitability [53]. Synaptic gating features the masking of individual synaptic elements, thereby making the network connectivity more sparse (diluted) in a particular context, reducing interference between accessible and inaccessible memory patterns. Though the precise and complex nature of synaptic gating makes individual synaptic control less plausible, there is some evidence of inhibitory connectivity at the level of individual dendritic spines, which could be interpreted as synapse-level control [54]. However, there is considerably more evidence for control of synaptic inputs at the level of dendritic branches [55], supported by the presence of specific dendrite-targeting inhibitory interneurons in both cortex and hippocampus [56,57], and findings such as dendritic clustering [58], localized dendritic inhibition [59], and localized neuromodulation [60]. These mechanisms could allow for the electrical isolation of particular areas of the dendrites [61], and clustering and control of specific input groups [62], thereby implementing a dendritic-branch-specific gating. One area for future work would be to explore this by relaxing the precision of synaptic gating, for which the analytical techniques described here may already be useful.

We proposed two methods for setting contextual states in our model: random allocation and refinement. In random gating, contextual states are imposed at the time of learning, which preallocates neurons and synapses to store memory patterns in a given context, and does not affect gated components. For random gating to be effective, contextual states are also reinstated during recall—such contextual reinstatement is hypothesized to occur in the brain [34–36]. We observed that neuronal gating was more effective than synaptic gating for random allocation, resulting in high memory capacity (Fig. 3) and low stability of inaccessible memories (Fig. 6). Such a finding is intriguing in light of recent evidence for allocation of particular neurons (through excitability changes) to memories in the amygdala and hippocampal system [24,53,63]. Specifically, the activity levels observed in such studies (∼20%–30%) correspond to optimal neuronal gating levels in our model for tens to hundreds of contexts (Fig. 5). Importantly, silencing of the allocated neurons—i.e., neurons with elevated excitability during memory formation—has been shown to decrease memory expression [63], in line with random neuronal gating during both learning and recall. Our model may be used to propose additional, context-dependent manipulations that may artificially bias recall toward or away from particular memories.

The second contextual gating method that we proposed was refinement, in which gating states are optimized to strengthen accessible memories after learning (i.e., contextual states are imposed only at retrieval). This scenario was proposed specifically with synaptic control in mind, as synapses were gated off in a particular context if they had a net harmful effect on recall. We found that synaptic refinement resulted in even larger gains in memory capacity (Fig. 8) and better control of memory accessibility (Fig. 9), with connectivity levels approaching 50% (Fig. 7). Such a scheme fits well with the idea that memories are constantly being updated [64], with each recall event possibly leading to changes [51]. Furthermore, recent work has also suggested that memory expression can be masked and controlled through inhibition [37,65]. Additionally, the finding that synaptic refinement can stabilize shared groups of memory patterns (even using a random weight matrix) suggests that the precision of synaptic learning may not be crucial to long-term maintenance of memories (see Refs. [50,66,67] for related studies). Lastly, we note that since refinement was defined at the level of synapses rather than neurons, we did not consider the refinement of neuronal gating here. However, while *random* neuronal gating proved to be surprisingly effective, it was still the case that many synapses remained unused when capacity was maximized [Fig. 5(c), inset]. This suggests that a more algorithmic allocation process (e.g., one that equalizes the number of contexts that a neuron participates in) may be possible, which would be interesting to explore in future work.

Memory capacity, stability, and information content results were obtained using conditional measures—i.e., we assumed an externally controlled context, with each measure conditioned upon this fixed external state. This choice was made due to reasons of analytical tractability, and the desire to model a flexible memory system more related to biological memory [20,25]. Furthermore, the questions addressed here did not depend upon an explicitly specified context representation, as the focus was rather on the consequences of contextual control for an associative

memory network. However, our conditional measures make comparisons with previous capacity estimates difficult, as performance metrics usually require stabilizing all memories at all times [3]. We therefore stress that our aim was not to compare the capacity with previous networks, but rather to propose a framework within which to study capacity while emphasizing flexibility and accessibility. In future work, an explicit implementation of a context-encoding network may potentially lead to interesting biological comparisons, including dynamic competition between memories [26], as well as descriptions of pathological memory states like in amnesia [68]. Overall, there may be a trade-off between capacity and flexibility, and it may be in the interest of the brain to favor dynamic accessibility over total amount of memories stored. Theoretically, it will be necessary to develop richer measures than capacity to determine the use and function of memory that includes flexibility, retrieval dynamics, and evolving memory patterns [69–71].

One simple way of considering the complexity of such contextual representations in the capacity estimations is to consider a Hopfield-like (context-encoding) network storing contextual states as memory patterns. For neuronal gating, our model would require the size of such a network to scale with the number of contexts s, which should be relatively small compared to the memory network size, and thus result in qualitatively similar results. For synaptic gating, however, our model would require fundamentally different dynamics in which synaptic weights are controlled by activities of other neurons [72]. Assuming this is possible, and constraining each context-encoding neuron to have on the order of N synapses, the network should scale with sN, leading to a context network that is potentially orders of magnitude larger than memory network, and reducing the memory capacity below the Hopfield limit with only a few contexts. Such intuitions argue against full and precise synaptic control. However, relaxations of such precise control, especially for a smaller number of contexts, may be possible, and our model serves as a proof of concept of such a situation.

Context-modular memory networks utilize several principles in common with previous associative memory models, including sparsity, modularity, and hierarchical organization [3,10]. Modularity can be used to mimic the architecture of the cortex—dense local with sparse long-ranged connectivity [12–14]—and several other models have explored a hierarchical organization of memories [8,73], or nontreelike organization of semantic memories [74]. Furthermore, other models have incorporated selective connectivity to enhance capacity [75]. Our model, in contrast, is unique in considering context as the main determinant of memory groupings (but see Ref. [76]), as well as modeling context as an external signal imposed on the memory network (in contrast to previous hierarchical models). However, more generally, the modularity of the model plays the same important role as in these previous models, namely to reduce crosstalk and interference between memories. By exploring the amount of overlapping connectivity [shared contexts per synapse, Figs. 5(c) and 5(d)], we found that the effectiveness of random gating is in part due to its ability to isolate the weights used for different contexts. Despite this, we still found that for most parameter settings, the majority of the weights were used in more than one context, and that this is crucial to the capacity gains seen in the model.

Extensions to associative memory networks that improve memory capacity, such as low activity patterns [6], alternative learning rules [19], or complex synapses [77], may be incorporated into the context-modular memory network architecture for additional benefits. We note that this also may include connections to recent advances in dense associative memory with exponential capacity [78–80], and heteroassociative memory architectures [81]. Furthermore, we note that while many of these classic models use very simple binary units for the neurons, there have been many successful extensions of this work to more realistic single-neuron dynamics [82]. While we limited ourselves here to the simple binary case, the neuroscientific implications of the work may be better explored in more realistic models. We also assume symmetric connectivity throughout, but a more realistic scenario would be asymmetric connectivity, for which other analyses apply (see Refs. [83,84] for details on asymmetric matrices).

Finally, other models have considered contextual control of learning and memory [28,32,62,85,86], as well as similar gating mechanisms applied in other situations (e.g., cortical information routing [38,39,87], deep learning [40], inhibitory learning [88], and motor learning [89]), highlighting the importance of contextual flexibility for cognition. Our work may serve as a theoretical framework with which to begin to study contextual gating in such additional scenarios.

In conclusion, the integration of associative memory models with retrieval processes and decision making is a promising area of future research [30]. Organizing memories around context leads to control over memory access, which the brain may harness. We propose the contextual gating framework introduced here as a promising model for flexible memory access which may lead to new insights into biological and artificial memory.

ACKNOWLEDGMENTS

We thank Helen Barron, Vezha Boboeva, Adam Packer, João Sacramento, Andrew Saxe, Misha Tsodyks, and Friedemann Zenke for helpful comments at various stages of this work, and Rubem Erichsen, Jr. for carefully reading the manuscript and valuable comments. This work was supported by a Sir Henry Dale Fellowship by the Wellcome Trust and the Royal Society [No. WT100000 (W. F. P., E. J. A., and T. P. V.)], a Wellcome Trust Senior Research Fellowship [No. 214316/Z/18/Z (E. J. A. and T. P. V.)], and

a Research Project Grant by the Leverhulme Trust [No. RPG-2016-446 (E. J. A.)].

DATA AVAILABILITY

All code used to run simulations and generate the figures herein will be made available upon publication of this work [90].

APPENDIX A: EQUIVALENCE OF $\{\pm 1\}$ AND $\{0, 1\}$ FORMULATIONS

We define the context-modular memory network considering three states, $S_i^{(k)} = \pm 1$ for neuron i allocated to the active context k and $S_i^{(k)} = 0$ for neuron i gated off in the active context k. This formulation can be modified to a $\{0, 1\}$ formulation, making it effectively a network of binary neurons. To see this, it is sufficient to show that the dynamics in either case are equivalent. In the $\{0, 1\}$ formulation [3], we consider the state of neuron i as $V_i^{(k)} = \frac{1}{2}(S_i^{(k)} + 1)$ for neuron i allocated to the active context k and $V_i^{(k)} = 0$ for neuron i gated off in the inactive context, and memory patterns given by $\eta_i^{\sigma^\mu} = \frac{1}{2}(\xi_i^{\sigma^\mu} + 1)$, with $\sigma = 1, \ldots, s$, $\mu = 1, \ldots, p$, and $i = 1, \ldots, N$. From these definitions, and considering a neuron i in the active context k, we arrive at

$$\begin{aligned} S_i^{(k)} &= \mathrm{sgn}\left(a_i^k \sum_{j=1}^{N} a_j^k c_{ij}^k J_{ij} S_j^{(k)} \right) = \mathrm{sgn}\left(\frac{a_i^k}{acN} \sum_{j=1}^{N} a_j^k c_{ij}^k \sum_{\sigma=1}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu} \xi_j^{\sigma^\mu} S_j^{(k)} \right) \\ &= \mathrm{sgn}\left(\frac{a_i^k}{acN} \sum_{j=1}^{N} a_j^k c_{ij}^k \sum_{\sigma=1}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} (2\eta_i^{\sigma^\mu} - 1)(2\eta_j^{\sigma^\mu} - 1)(2V_j^{(k)} - 1) \right) \\ &= \mathrm{sgn}\left(a_i^k \sum_{j=1}^{N} a_j^k c_{ij}^k w_{ij} \left(V_j^{(k)} - \frac{1}{2} \right) \right). \end{aligned} \tag{A1}$$

The input to neuron i can equivalently be given in terms of V variables,

$$V_i^{(k)} = a_i^k \mathcal{H}\left(\sum_{j=1}^{N} a_j^k c_{ij}^k w_{ij} V_j^{(k)} - \theta_i^k \right), \tag{A2}$$

where $\mathcal{H}(\cdot)$ is the Heaviside function, such that $2\mathcal{H}(x) - 1 = \mathrm{sgn}(x)$, w_{ij} is the weight matrix, and θ_i^k is a context-dependent threshold. In this formulation, the weight matrix w_{ij} is given by

$$w_{ij} = \frac{8}{acN} \sum_{\sigma=1}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \left(\eta_i^{\sigma^\mu} - \frac{1}{2} \right) \left(\eta_j^{\sigma^\mu} - \frac{1}{2} \right), \tag{A3}$$

and the (context-specific) threshold θ_i^k is

$$\theta_i^k = \frac{1}{2} \sum_{j=1}^{N} a_j^k c_{ij}^k w_{ij}. \tag{A4}$$

Note that we moved a_i^k to the outside of the argument of the Heaviside function in Eq. (A2) to avoid $\mathcal{H}(0) = 1/2$ which results from the identity $\mathcal{H}(x) = \frac{1}{2}[\mathrm{sgn}(x) + 1]$.

Thus, the deterministic dynamics of the two formulations are equivalent. We note that this holds also for the stochastic version of the dynamics, in which $\langle V_i \rangle = \sigma(h_i) = e^{2h_i}/(e^{2h_i} + 1)$ and $\langle S_i \rangle = \tanh(h_i)$, where $\langle \cdot \rangle$ is thermal average, due to the identity $\tanh(x) = 2\sigma(x) - 1$.

APPENDIX B: MEMORY CAPACITY

Here, we estimate the capacity of the context-modular memory network using a replica mean-field approach similar to previous work [3,18,41,44]. We consider a stochastic version of the Hopfield network (incorporating contextual states), with dynamics

$$\begin{cases} \Pr\left(S_i^{(k)} = \pm 1 \right) = \frac{1}{1 + \exp(\mp 2\beta h_i^k)} & \text{if } a_i^k = 1 \\ \Pr\left(S_i^{(k)} = 0 \right) = 1 & \text{otherwise,} \end{cases} \tag{B1}$$

where h_i^k is the input to neuron i if context k is active and $\beta = 1/T$ is the inverse of the temperature T.

To calculate the context-specific memory capacity α_{cxt}^* we consider that a given context k is active, for which we can write an effective weight matrix [cf. Eq. (18)],

$$\tilde{J}_{ij}^k = a_i^k a_j^k \left(\frac{1}{aN} \sum_{\mu=1}^{\mathcal{P}} \tilde{\xi}_i^{k^\mu} \tilde{\xi}_j^{k^\mu} + \delta_{ij}^k \right), \tag{B2}$$

where $\mathcal{P}$ is the total number of memorylike pattern terms per synapse when context k is active and δ_{ij}^k is a noise term

that incorporates the information related to memory patterns from inactive contexts and dilution.

We drop the context index k, considering only the N_{cxt} neurons allocated to the active context, without loss of generality, to simplify our notation and calculations. Rewriting $\tilde{\xi}_i^{k^\mu} = \xi_i^\mu$, we arrive at a general effective weight matrix give by

$$\tilde{J}_{ij} = \frac{1}{N_{\mathrm{cxt}}} \sum_{\mu=1}^{\mathcal{P}} \xi_i^\mu \xi_j^\mu + \delta_{ij}. \tag{B3}$$

The Hamiltonian considering this effective weight matrix is [3]

$$\begin{aligned} H = -\frac{1}{2} \sum_{i=1}^{N_{\mathrm{cxt}}} \sum_{j=1}^{N_{\mathrm{cxt}}} \tilde{J}_{ij} S_i S_j + \frac{\mathcal{P}}{2} &= -\frac{1}{2N_{\mathrm{cxt}}} \sum_{i=1}^{N_{\mathrm{cxt}}} \sum_{j=1}^{N_{\mathrm{cxt}}} \left(\sum_{\mu=1}^{\mathcal{P}} \xi_i^\mu \xi_j^\mu + N_{\mathrm{cxt}} \delta_{ij} \right) S_i S_j + \frac{\mathcal{P}}{2} \\ &= -\frac{1}{2N_{\mathrm{cxt}}} \left[\sum_{\mu=1}^{\mathcal{P}} \left(\sum_{i=1}^{N_{\mathrm{cxt}}} \xi_i^\mu S_i \right)^2 + N_{\mathrm{cxt}} \sum_{i=1}^{N_{\mathrm{cxt}}} \sum_{j=1}^{N_{\mathrm{cxt}}} \delta_{ij} S_i S_j \right] + \frac{\mathcal{P}}{2}, \end{aligned} \tag{B4}$$

whose partition function is

$$Z = \mathrm{Tr}_S[\exp(-\beta H\{S_i\})] = \exp\left(-\frac{\beta\mathcal{P}}{2}\right) \mathrm{Tr}_S \left\{ \exp\left[\frac{\beta}{2N_{\mathrm{cxt}}} \left(\sum_{\mu=1}^{\mathcal{P}} \left(\sum_{i=1}^{N_{\mathrm{cxt}}} \xi_i^\mu S_i \right)^2 + \beta N_{\mathrm{cxt}} \sum_{i=1}^{N_{\mathrm{cxt}}} \sum_{j=1}^{N_{\mathrm{cxt}}} \delta_{ij} S_i S_j \right) \right] \right\}. \tag{B5}$$

We then use the integral identity,

$$\int_{-\infty}^{\infty} \mathrm{e}^{-ax^2 \pm bx} \mathrm{d}x = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right), \tag{B6}$$

to write

$$\exp\left[\frac{\beta}{2N_{\mathrm{cxt}}} \left(\sum_{i=1}^{N_{\mathrm{cxt}}} \xi_i^\mu S_i \right)^2 \right] = \sqrt{\frac{2\pi}{\beta N_{\mathrm{cxt}}}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2} \beta N_{\mathrm{cxt}} (m^\mu)^2 + \beta m^\mu \sum_{i=1}^{N_{\mathrm{cxt}}} \xi_i^\mu S_i \right] \mathrm{d}m^\mu, \tag{B7}$$

thus arriving at

$$\begin{aligned} Z &= \left(\frac{2\pi}{\beta N_{\mathrm{cxt}}}\right)^{\mathcal{P}/2} \exp\left(-\frac{\beta\mathcal{P}}{2}\right) \mathrm{Tr}_S \left\{ \prod_{\mu=1}^{\mathcal{P}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2} \beta N_{\mathrm{cxt}} (m^\mu)^2 + \beta m^\mu \sum_{i=1}^{N_{\mathrm{cxt}}} \xi_i^\mu S_i + \beta N_{\mathrm{cxt}} \sum_{i=1}^{N_{\mathrm{cxt}}} \sum_{j=1}^{N_{\mathrm{cxt}}} \delta_{ij} S_i S_j \right] \mathrm{d}m^\mu \right\} \\ &= \left(\frac{2\pi}{\beta N_{\mathrm{cxt}}}\right)^{\mathcal{P}/2} \exp\left(-\frac{\beta\mathcal{P}}{2}\right) \mathrm{Tr}_S \left\{ \exp\left(\beta N_{\mathrm{cxt}} \sum_{i=1}^{N_{\mathrm{cxt}}} \sum_{j=1}^{N_{\mathrm{cxt}}} \delta_{ij} S_i S_j \right) \prod_{\mu=1}^{\mathcal{P}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2} \beta N_{\mathrm{cxt}} (m^\mu)^2 + \beta m^\mu \sum_{i=1}^{N_{\mathrm{cxt}}} \xi_i^\mu S_i \right] \mathrm{d}m^\mu \right\}. \end{aligned} \tag{B8}$$

To use the replica method, we need to first replicate the partition function,

$$Z^n = \exp\left(-\frac{\beta\mathcal{P}n}{2}\right) \mathrm{Tr}_S \left\{ \prod_{\rho=1}^{n} \left[\prod_{\mu=1}^{\mathcal{P}} \int_{-\infty}^{\infty} \sqrt{\frac{2\pi}{\beta N_{\mathrm{cxt}}}} \exp\left[-\frac{1}{2} \beta N (m_\rho^\mu)^2 \beta m_\rho^\mu \sum_{i=1}^{N_{\mathrm{cxt}}} \xi_i^\mu S_i^\rho \right] \mathrm{d}m_\rho^\mu \right] \prod_{\gamma=1}^{n} \exp\left(\beta N_{\mathrm{cxt}} \sum_{i=1}^{N_{\mathrm{cxt}}} \sum_{j=1}^{N_{\mathrm{cxt}}} \delta_{ij} S_i^\rho S_j^\gamma \right) \right\}. \tag{B9}$$

From the replicated partition function, we can calculate the averaged free energy per connected site using the replica trick,

$$f = -\lim_{cN_{\mathrm{cxt}}\to\infty}\frac{1}{\beta c N_{\mathrm{cxt}}}\langle\langle \ln(Z)\rangle_\delta\rangle_\xi = -\lim_{cN_{\mathrm{cxt}}\to\infty}\frac{1}{\beta c N_{\mathrm{cxt}}}\lim_{n\to 0}\frac{1}{n}\left[\langle\langle Z^n\rangle_\delta\rangle_\xi - 1\right]$$
$$= \frac{1}{2}\sum_{\mu=1}^{\mathcal{P}}(m_\mu)^2 + \frac{\alpha}{2\beta}\left[\ln(1-\chi) + \frac{\chi(1-\beta)}{1-\chi}\right] + \frac{\alpha r\chi}{2} - \frac{\Delta^2\chi^2}{4\beta} - \frac{1}{\beta}\int_{-\infty}^{\infty}\ln(2)\cosh\left[\beta\sqrt{\alpha r + \Delta^2 q} + \sum_{\mu=1}^{\mathcal{P}} m_\mu\right]\frac{\exp\left(-\frac{z^2}{2}\right)}{\sqrt{2\pi}}\mathrm{d}z, \tag{B10}$$

where $\langle\cdot\rangle_\delta$ indicates average over the variable δ_{ij}, $\langle\cdot\rangle_\xi$ indicates average over the memory patterns,

$$\chi = \frac{\beta}{N_{\mathrm{cxt}}}\sum_{i=1}^{N_{\mathrm{cxt}}}\left(\langle S_i^2\rangle - \langle S_i\rangle^2\right) = \beta(1-q), \tag{B11}$$

and

$$r = \frac{q}{(1-\chi)^2}. \tag{B12}$$

To calculate the average over δ_{ij} we assumed that this variable is random with zero mean and variance $\Delta^2/N_{\mathrm{cxt}}$:

$$\delta_{ij} \sim \mathcal{N}\left(0, \frac{\Delta^2}{N_{\mathrm{cxt}}}\right). \tag{B13}$$

Taking the derivative of f [Eq. (B10)] with respect to m_μ, q, χ, and r, equating them to zero, and considering a given recalled memory $\mathbf{m} = (m, 0, \ldots, 0)$ results in the following saddle point equations,

$$m = \int_{-\infty}^{\infty}\tanh\left[\beta\left(z\sqrt{\alpha r + \Delta^2 q} + m\right)\right]\frac{\exp\left(-\frac{z^2}{2}\right)}{\sqrt{2\pi}}\mathrm{d}z, \tag{B14}$$

$$q = \int_{-\infty}^{\infty}\tanh^2\left[\beta\left(z\sqrt{\alpha r + \Delta^2 q} + m\right)\right]\frac{\exp\left(-\frac{z^2}{2}\right)}{\sqrt{2\pi}}\mathrm{d}z, \tag{B15}$$

$$\chi = \frac{1}{\sqrt{\alpha r + \Delta^2}}\int_{-\infty}^{\infty}\tanh\left[\beta\left(z\sqrt{\alpha r + \Delta^2 q} + m\right)\right] \times \frac{z\exp\left(-\frac{z^2}{2}\right)}{\sqrt{2\pi}}\mathrm{d}z, \tag{B16}$$

and

$$r = \frac{q}{(1-\chi)^2}. \tag{B17}$$

Solving this set of equations for the limit $T \to 0$ (and thus $q = 1$), we arrive at

$$m = \mathrm{erf}\left[\frac{m}{\sqrt{2(\alpha r + \Delta^2)}}\right], \tag{B18}$$

$$r = \frac{1}{(1-\chi)^2}, \tag{B19}$$

and

$$\chi = \sqrt{\frac{2}{\pi(\alpha r + \Delta^2)}}\exp\left[-\frac{m^2}{2(\alpha r + \Delta^2)}\right]. \tag{B20}$$

In the context-modular memory networks, the noise term is proportional to α, so that $\Delta^2 = \Delta_0^2\alpha$. This allows us to rewrite Eqs. (B18)–(B20) to resemble the ones from the original Hopfield network as

$$m = \mathrm{erf}\left[\frac{m}{\sqrt{2\alpha\tilde{r}}}\right], \tag{B21}$$

$$\tilde{r} = \frac{1}{(1-\chi)^2} + \Delta_0^2, \tag{B22}$$

and

$$\chi = \sqrt{\frac{2}{\pi\alpha\tilde{r}}}\exp\left[-\frac{m^2}{2\alpha\tilde{r}}\right], \tag{B23}$$

where we have defined $\tilde{r} = r + \Delta_0^2$. Equations (B21)–(B23) can be solved numerically as a function of the noise level Δ_0. Given this dependence on the noise level, we refer to the two variables of interest, capacity and overlap, as α_Δ^* and m_Δ^*, respectively (Fig. 2). Furthermore, due to some simplifications in the algebra which will become apparent later, it is convenient to parametrize these variables not as a function of Δ_0, but rather as a function of $1 + \Delta_0^2$. As Δ_0^2 is multiplied by α to determine the absolute noise level, $1 + \Delta_0^2$ can be interpreted as the sum of the network load and the additional random noise.

Based on Δ_0^2 for the different gating methods, we can then estimate memory capacity considering the $\mathcal{P}$ stable patterns as

$$\alpha_{\mathcal{P}}^* = \alpha_\Delta^*(1+\Delta_0^2), \tag{B24}$$

and the context-specific memory capacity as

$$\alpha_{\mathrm{cxt}}^* = \frac{p}{\mathcal{P}}\alpha_\Delta^*(1+\Delta_0^2), \tag{B25}$$

where $\alpha_\Delta^*(1+\Delta_0^2)$ is the value of the capacity α in the solution of Eqs. (B21)–(B23) with a noise level Δ_0^2.

APPENDIX C: RANDOM NEURONAL AND SYNAPTIC GATING—ESTIMATION OF $\mathcal{P}$

To calculate the context-specific memory capacity for random neuronal and synaptic gating, we need to estimate the total number of memorylike pattern terms $\mathcal{P}$ in Eq. (18). To estimate $\mathcal{P}$, we first identify the number of elements in the sum over memory patterns of inactive contexts and then follow previous approach [41], taking the average over the gating variables c_{ij}^k, a_i^σ, and c_{ij}^σ in Eq. (18b), so that the effective weight matrix is equal to the averaged terms plus noise,

$$\frac{c_{ij}^k}{acN}\left(\sum_{\mu=1}^{p}\xi_i^{k^\mu}\xi_j^{k^\mu}+\sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s}a_i^\sigma a_j^\sigma c_{ij}^\sigma\sum_{\mu=1}^{p}\xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\right)=\frac{1}{aN}\left[\underbrace{\left(\frac{c_{ij}^k}{c}\sum_{\mu=1}^{p}\xi_i^{k^\mu}\xi_j^{k^\mu}\right)}_{p\text{ memory patterns}}+ac\underbrace{\left(\sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s}\frac{a_i^\sigma a_j^\sigma c_{ij}^k c_{ij}^\sigma}{ac^2}\sum_{\mu=1}^{p}\xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\right)}_{a^2c^2(s-1)p\text{ memory patterns}}\right] \tag{C1}$$

$$=\frac{1}{aN}\left(\sum_{\mu=1}^{p}\xi_i^{k^\mu}\xi_j^{k^\mu}+ac\sum_{\mu=1}^{a^2c^2(s-1)p}\hat{\xi}_i^{k^\mu}\hat{\xi}_j^{k^\mu}\right)+\delta_{ij}^k \tag{C2}$$

$$=\frac{1}{aN}\sum_{\mu=1}^{\mathcal{P}}\tilde{\xi}_i^{k^\mu}\tilde{\xi}_j^{k^\mu}+\delta_{ij}^k, \tag{C3}$$

where $\hat{\xi}_i^{k^\mu}\hat{\xi}_j^{k^\mu}$ are the nonzero terms in the sum over the $(s-1)p$ memory patterns from the inactive contexts, $a_i^\sigma a_j^\sigma c_{ij}^k c_{ij}^\sigma \xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}$. Note that we added a normalizing term ac^2 in Eq. (C1) to write the sum as the equivalent of a low-activity, diluted network with activity level a and dilution c^2. We multiplied the sum by ac to keep the left- and right-hand sides of Eq. (C1) equal. The $\mathcal{P}$ memory patterns are defined as follows: $\tilde{\xi}_i^{k^\mu}=\xi_i^{k^\mu}$ for $i=1,\ldots,N$ and $\mu=1,\ldots,p$; and $\mathrm{Prob}(\tilde{\xi}_i^{k^\mu}=\pm 1)=1/2$ for $i=1,\ldots,N$ and $\mu=p+1,\ldots,\mathcal{P}$, which is assumed without loss of generality. We then write the noise δ_{ij}^k based on the p patterns of the active context k plus the spurious synapse-specific memorylike pattern terms from the $s-1$ inactive contexts [combining both the left-hand side of Eq. (C1) and Eq. (C2)]:

$$\delta_{ij}^k=\frac{c_{ij}^k}{acN}\left(\sum_{\mu=1}^{p}\xi_i^{k^\mu}\xi_j^{k^\mu}+\sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s}a_i^\sigma a_j^\sigma c_{ij}^\sigma\sum_{\mu=1}^{p}\xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\right)-\frac{1}{aN}\left(\sum_{\mu=1}^{p}\xi_i^{k^\mu}\xi_j^{k^\mu}+ac\sum_{\mu=1}^{a^2c^2(s-1)p}\hat{\xi}_i^{k^\mu}\hat{\xi}_j^{k^\mu}\right). \tag{C4}$$

From Eq. (C4), the variance of the noise term δ_{ij}^k can be estimated as

$$\begin{aligned}\frac{\Delta^2}{aN}&=\frac{p}{a^2cN^2}+\frac{a^2(s-1)p}{a^2N^2}-\frac{p}{a^2N^2}-\frac{a^4c^4(s-1)p}{a^2N^2}\\&=\frac{1}{aN}\left[\frac{p}{aN}\left(\frac{1}{c}+a^2(s-1)\right)-\frac{p}{aN}(1+a^4c^4(s-1))\right].\end{aligned} \tag{C5}$$

We then rewrite the noise δ_{ij}^k based on the $\mathcal{P}$ memorylike pattern terms [combining both the left-hand side of Eq. (C1) and Eq. (C3)],

$$\delta_{ij}^k=\frac{c_{ij}^k}{acN}\left(\sum_{\mu=1}^{p}\xi_i^{k^\mu}\xi_j^{k^\mu}+\sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s}a_i^\sigma a_j^\sigma c_{ij}^\sigma\sum_{\mu=1}^{p}\xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\right)-\frac{1}{aN}\sum_{\mu=1}^{\mathcal{P}}\tilde{\xi}_i^{k^\mu}\tilde{\xi}_j^{k^\mu}, \tag{C6}$$

with variance given by

$$\frac{\Delta^2}{aN} = \frac{p}{a^2cN^2} + \frac{a^2(s-1)p}{a^2N^2} - \frac{\mathcal{P}}{a^2N^2} = \frac{1}{aN}\left[\frac{p}{aN}\left(\frac{1}{c} + a^2(s-1)\right) - \frac{\mathcal{P}}{aN}\right]. \quad \text{(C7)}$$

By comparing Eqs. (C5) and (C7), we arrive at

$$\mathcal{P} = p[1 + a^4c^4(s-1)]. \quad \text{(C8)}$$

Substituting p by $\mathcal{P}/[1 + a^4c^4(s-1)]$ in Eq. (C7) results in

$$\frac{\Delta^2}{aN} = \frac{1}{aN}\left[\frac{\mathcal{P}}{aN}\frac{\frac{1}{c} + a^2(s-1)}{1 + a^4c^4(s-1)} - \frac{\mathcal{P}}{aN}\right] = \frac{\alpha_{\mathcal{P}}}{aN}\left[\frac{\frac{1}{c} + a^2(s-1)}{1 + a^4c^4(s-1)} - 1\right], \quad \text{(C9)}$$

and thus

$$\Delta_0^2 + 1 = \frac{\frac{1}{c} + a^2(s-1)}{1 + a^4c^4(s-1)} = \frac{1 + a^2c(s-1)}{c[1 + a^4c^4(s-1)]}. \quad \text{(C10)}$$

APPENDIX D: SYNAPTIC REFINEMENT—ESTIMATION OF THE VARIANCE OF THE NOISE TERM

To calculate the context-specific memory capacity for the synaptic refinement method, we need to estimate the variance of the noise term δ_{ij}^k. We rewrite Eq. (38) here for convenience,

$$\frac{(\Delta^k)^2}{aN} = \frac{\left\langle \left(\mathcal{X}_{ij}^k + \mathcal{Y}_{ij}^k\right)^2 c_{ij}^k \right\rangle_{ij}}{(\mathcal{J}^k)^2} - \frac{\alpha_{\text{cxt}}}{aN}, \quad \text{(D1)}$$

where [Eq. (37)]

$$\mathcal{J}^k = \frac{\left\langle c_{ij}^k \mathcal{X}_{ij}^k \left(\mathcal{X}_{ij}^k + \mathcal{Y}_{ij}^k\right)\right\rangle_{ij} - \left\langle c_{ij}^k \mathcal{X}_{ij}^k \right\rangle_{ij}^2 \left\langle c_{ij}^k \left(\mathcal{X}_{ij}^k + \mathcal{Y}_{ij}^k\right)\right\rangle_{ij}^2}{\left\langle \left(\mathcal{X}_{ij}^k\right)^2 \right\rangle_{ij} - \left\langle \mathcal{X}_{ij}^k \right\rangle_{ij}^2} = \frac{\left\langle c_{ij}^k \mathcal{X}_{ij}^k \left(\mathcal{X}_{ij}^k + \mathcal{Y}_{ij}^k\right)\right\rangle_{ij}}{\left\langle \left(\mathcal{X}_{ij}^k\right)^2 \right\rangle_{ij}}. \quad \text{(D2)}$$

We calculate each individual term from Eqs. (D1) and (D2) below based on $\mathcal{X}_{ij}^k \sim \mathcal{N}(0, \sigma_x^2)$ and $\mathcal{Y}_{ij}^k \sim \mathcal{N}(0, \sigma_y^2)$ [σ_x^2 and σ_y^2 are defined in Eq. (32)]. The term averaged over i, j in Eq. (D1) can be calculated as

$$\begin{aligned}\left\langle (\mathcal{X}_{ij}^k + \mathcal{Y}_{ij}^k)^2 c_{ij}^k \right\rangle_{ij} &= \frac{1}{2\pi\sigma_x\sigma_y}\left[\int_0^\infty \exp\left(-\frac{x^2}{2\sigma_x^2}\right)\int_{-x}^\infty (x+y)^2 \exp\left(-\frac{y^2}{2\sigma_y^2}\right)\mathrm{d}y\mathrm{d}x \right. \\ &\quad \left. + \int_{-\infty}^0 \exp\left(-\frac{x^2}{2\sigma_x^2}\right)\int_{-\infty}^x (x+y)^2 \exp\left(-\frac{y^2}{2\sigma_y^2}\right)\mathrm{d}y\mathrm{d}x\right] \\ &= \left(1 + a^2(s-1) + \frac{a\sqrt{s-1}}{\pi c}\right)\sigma_x^2 c. \end{aligned} \quad \text{(D3)}$$

The normalization term $\mathcal{J}^k$ [Eq. (D2)] is

$$\begin{aligned}\mathcal{J}^k &= \frac{\left\langle c_{ij}^k \mathcal{X}_{ij}^k \left(\mathcal{X}_{ij}^k + \mathcal{Y}_{ij}^k\right)\right\rangle_{ij}}{\left\langle \left(\mathcal{X}_{ij}^k\right)^2 \right\rangle_{ij}} \\ &= \frac{1}{2\pi\sigma_x\sigma_y}\left[\int_0^\infty \exp\left(-\frac{x^2}{2\sigma_x^2}\right)\int_{-x}^\infty x(x+y)\exp\left(-\frac{y^2}{2\sigma_y^2}\right)\mathrm{d}y\mathrm{d}x \right. \\ &\quad \left. + \int_{-\infty}^0 \exp\left(-\frac{x^2}{2\sigma_x^2}\right)\int_{-\infty}^x x(x+y)\exp\left(-\frac{y^2}{2\sigma_y^2}\right)\mathrm{d}y\mathrm{d}x\right]\left[\frac{1}{\sqrt{2\pi}\sigma_x}\int_{-\infty}^\infty x^2 \exp\left(-\frac{x^2}{2\sigma_x^2}\right)\mathrm{d}x\right]^{-1} \\ &= \left(1 + \frac{a\sqrt{s-1}}{\pi c}\right)c. \end{aligned} \quad \text{(D4)}$$

The gating parameter c is given by Eq. (33). Note that the limits of the integrals in Eqs. (D3) and (D4) are not the same as in Eq. (33) because we need to integrate over the nongated synapses when taking into consideration the gating variable c_{ij}^k.

APPENDIX E: RANDOM WEIGHT MATRIX—ESTIMATION OF THE VARIANCE OF THE NOISE TERM

When the weight matrix is random, approximately half of the connections have the opposite of the desired sign, meaning that they will be gated off. The elements that are not gated off have the correct sign, but with a random value taken from a Gaussian distribution. We can thus rewrite the effective weight matrix as

$$\tilde{J}_{ij}^k = \frac{c_{ij}^k}{\mathcal{J}^k} J_{ij} = \frac{\tilde{\eta}_{ij}}{\mathcal{J}^k} \operatorname{sgn}\left(\sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu}\right), \tag{E1}$$

where $\tilde{\eta}_{ij}$ are the positive entries of η_{ij} [a random matrix with symmetric elements drawn from a Gaussian distribution as defined in Eq. (48)],

$$\tilde{\eta}_{ij} = \frac{\eta_{ij} + |\eta_{ij}|}{2}. \tag{E2}$$

Substituting Eq. (E1) in Eq. (50) results in

$$\delta_{ij}^k = \frac{\tilde{\eta}_{ij}}{\mathcal{J}^k} \operatorname{sgn}\left(\sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu}\right) - \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu}, \tag{E3}$$

with variance [41,43]

$$\frac{(\Delta^k)^2}{N} = \frac{1}{(\mathcal{J}^k)^2} \left\langle \left[\operatorname{sgn}\left(\sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu}\right)\right]^2 \right\rangle_\xi \langle (\tilde{\eta}_{ij})^2 \rangle_{ij} - \frac{\alpha_{\text{cxt}}}{N}, \tag{E4}$$

where $\langle \cdot \rangle_\xi$ and $\langle \cdot \rangle_{ij}$ represent average over all memory patterns and all connections, respectively. The correlation term $\mathcal{J}^k$ is [41,43]

$$\mathcal{J}^k = \frac{N}{\sqrt{p}} \left\langle \xi_i^{k^\mu} \xi_j^{k^\mu} \operatorname{sgn}\left(\sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu}\right) \right\rangle_\xi \langle \tilde{\eta}_{ij} \rangle_{ij}. \tag{E5}$$

We calculate below each term from Eqs. (E4) and (E5). First, the terms involving the memory patterns can be estimated as [41,43]

$$\left\langle \left[\operatorname{sgn}\left(\sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu}\right)\right]^2 \right\rangle_\xi = \int_{-\infty}^{\infty} \frac{[\operatorname{sgn}(x)]^2}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 \tag{E6}$$

and

$$\left\langle \xi_i^{k^\mu} \xi_j^{k^\mu} \operatorname{sgn}\left(\sum_{\mu=1}^{p} \xi_i^{k^\mu} \xi_j^{k^\mu}\right) \right\rangle_\xi = \int_{-\infty}^{\infty} \frac{\operatorname{sgn}(x)}{\sqrt{2\pi}} x e^{-x^2/2} dx = \sqrt{\frac{2}{\pi}}. \tag{E7}$$

Lastly, the terms involving the gated random weight matrix $\tilde{\eta}_{ij}$ are

$$\langle (\tilde{\eta}_{ij})^2 \rangle_{ij} = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} x^2 e^{-x^2/2} dx = \frac{1}{2} \tag{E8}$$

and

$$\langle \tilde{\eta}_{ij} \rangle_{ij} = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} x e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}}. \tag{E9}$$

Note that we define the integral limits in Eqs. (E8) and (E9) to account for the fact that $c = 1/2$—i.e., half of the connections are zero—and that all elements of $\tilde{\eta}_{ij}$ are positive.

Combining Eqs. (E4)–(E9), we arrive at

$$\frac{\Delta^2}{N} = \frac{\pi^2 \alpha_{\text{cxt}}}{2N} - \frac{\alpha_{\text{cxt}}}{N} = \left(\frac{\pi^2}{2} - 1\right) \frac{\alpha_{\text{cxt}}}{N} = \Delta_0^2 \frac{\alpha_{\text{cxt}}}{N}. \tag{E10}$$

APPENDIX F: STABILITY

For convenience, we rewrite the definition of accessibility [Eq. (14)],

$$\kappa_i^{k,\phi,\phi^\nu} = \frac{a_i^k a_i^\phi \xi_i^{\phi^\nu} \sum_{j=1}^{N} a_j^k c_{ij}^k J_{ij} \xi_j^{\phi^\nu}}{|J_i^k|}, \tag{F1}$$

with

$$|J_i^k| = \sqrt{\sum_{j=1}^{N} a_j^k c_{ij}^k J_{ij}^2}. \tag{F2}$$

Note that we changed the notation from σ to ϕ to use σ as the index in the sum over all contexts in the calculations below.

1. Random neuronal and synaptic gating

We start by opening Eq. (F1):

$$\begin{aligned}\kappa_i^{k,\phi,\phi^\nu} &= \frac{1}{acN|J_i^k|}\sum_{j=1}^{N}\sum_{\sigma=1}^{s} a_i^\phi a_i^k a_i^\sigma a_j^k a_j^\sigma c_{ij}^k c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\xi_i^{\phi^\nu}\xi_j^{\phi^\nu} \\ &= \frac{1}{acN|J_i^k|}\left(\sum_{j=1}^{N} a_i^k a_i^\phi a_j^k a_j^\phi c_{ij}^k c_{ij}^\phi \right. \\ &\quad + \sum_{j=1}^{N} a_i^k a_i^\phi a_j^k a_j^\phi c_{ij}^k c_{ij}^\phi \sum_{\substack{\mu=1,\\ \mu\neq\nu}}^{p} \xi_i^{\phi^\mu}\xi_j^{\phi^\mu}\xi_i^{\phi^\nu}\xi_j^{\phi^\nu} \\ &\quad \left. + \sum_{j=1}^{N}\sum_{\substack{\sigma=1,\\ \sigma\neq\phi}}^{s} a_i^\phi a_i^k a_i^\sigma a_j^k a_j^\sigma c_{ij}^k c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\xi_i^{\phi^\nu}\xi_j^{\phi^\nu}\right).\end{aligned} \tag{F3}$$

Below we use Eq. (F3) to calculate stability of accessible and inaccessible memory patterns for random gating. The term $|J_i^k|$ is proportional to the standard deviation of row i of the weight matrix when context k is active (considering $k=\phi$):

$$\begin{aligned}|J_i^k| &= \sqrt{\sum_{j=1}^{N} a_j^k c_{ij}^k \left(\frac{1}{acN}\sum_{\sigma=1}^{s}\sum_{\mu=1}^{p}\xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\right)^2} \\ &= \sqrt{\frac{[1+a^2c(s-1)]\alpha_{\text{cxt}}}{c}}.\end{aligned} \tag{F4}$$

a. Accessible patterns, $k=\phi$

First we consider the case of accessible memories, where $k=\phi$ in Eq. (F3). For convenience we set $a_i^k = a_i^\phi = 1$ in Eq. (F3). For this case, we arrive at

$$\begin{aligned}\kappa_i^{k,\phi,\phi^\nu} &= \frac{1}{acN|J_i^k|}\left(\sum_{j=1}^{N} a_j^\phi c_{ij}^\phi + \sum_{j=1}^{N} a_j^\phi c_{ij}^\phi \sum_{\mu=1,\mu\neq\nu}^{p} \xi_i^{\phi^\mu}\xi_j^{\phi^\mu}\xi_i^{\phi^\nu}\xi_j^{\phi^\nu} \right. \\ &\quad \left. + \sum_{j=1}^{N}\sum_{\sigma=1,\sigma\neq\phi}^{s} a_i^\sigma a_j^\phi a_j^\sigma c_{ij}^\phi c_{ij}^\sigma \sum_{\mu=1}^{p}\xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\xi_i^{\phi^\nu}\xi_j^{\phi^\nu}\right) \\ &= \frac{1+\zeta}{|J_i^k|},\end{aligned} \tag{F5}$$

where

$$\begin{aligned}\zeta &\sim \mathcal{N}\left(0, \frac{p[1+a^2c(s-1)]}{acN}\right) \\ &\sim \mathcal{N}\left(0, \frac{[1+a^2c(s-1)]\alpha_{\text{cxt}}}{c}\right).\end{aligned} \tag{F6}$$

We thus arrive at the stability of accessible patterns,

$$\begin{aligned}\kappa_i^{\phi,\phi^\nu} &= \frac{1+\zeta}{\sqrt{[1+a^2c(s-1)]\alpha_{\text{cxt}}/c}} \\ &= \sqrt{\frac{c}{[1+a^2c(s-1)]\alpha_{\text{cxt}}}} + \eta,\end{aligned} \tag{F7}$$

where $\eta \sim \mathcal{N}(0,1)$.

Finally, the average stability of accessible memory patterns becomes

$$\bar{\kappa}_{\text{acc}} = \sqrt{\frac{c}{[1+a^2c(s-1)]\alpha_{\text{cxt}}}}. \tag{F8}$$

b. Inaccessible patterns, $k\neq\phi$

For the case of inaccessible memory patterns we set $k\neq\phi$ in Eq. (F3). Following the same calculations as before for the accessible memory patterns, we arrive at

$$\kappa_i^{\phi,\phi^\nu} = \frac{ac+\zeta}{|J_i^k|}, \tag{F9}$$

where

$$\begin{aligned}\zeta &\sim \mathcal{N}\left(0, \frac{p}{N} + \frac{a^2c(s-1)p}{acN}\right) \\ &\sim \mathcal{N}\left(0, \left(a + \frac{a^2c(s-1)}{c}\right)\alpha_{\text{cxt}}\right).\end{aligned} \tag{F10}$$

The stability of inaccessible memory patterns is thus

$$\begin{aligned}\kappa_i^{k,\phi^\nu} &= \frac{ac+\zeta}{\sqrt{[1+a^2c(s-1)]\alpha_{\text{cxt}}/c}} \\ &= ac\sqrt{\frac{c}{[1+a^2c(s-1)]\alpha_{\text{cxt}}}} + \eta,\end{aligned} \tag{F11}$$

where

$$\eta \sim \mathcal{N}\left(0, \frac{ac+a^2c(s-1)}{1+a^2c(s-1)}\right), \tag{F12}$$

which is approximately $\mathcal{N}(0,1)$ for large s. We thus see that the stability of inaccessible patterns scales with the gating parameters a and c. Lastly, we recall that for neuronal gating, only a proportion a of the stability parameters will be nonzero, so that the overall mean of the inaccessible patterns is scaled by an additional factor of a.

We thus arrive at the average stability of inaccessible memory patterns as

$$\bar{\kappa}_{\text{inacc}} = a^2c\sqrt{\frac{c}{[1+a^2c(s-1)]\alpha_{\text{cxt}}}} = a^2c\bar{\kappa}_{\text{acc}}. \tag{F13}$$

2. Synaptic refinement

For synaptic refinement, we begin by returning to the general memory stability [Eq. (F1)]:

$$\kappa_i^{k,\phi,\phi^\nu} = \frac{1}{aN|J_i^k|}\sum_{j=1}^{N}\sum_{\sigma=1}^{s} a_i^\phi a_i^k a_i^\sigma a_j^k a_j^\sigma c_{ij}^k \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\xi_i^{\phi^\nu}\xi_j^{\phi^\nu}$$
$$= \frac{1}{aN|J_i^k|}\left(\sum_{j=1}^{N} a_i^k a_i^\phi a_j^k a_j^\phi c_{ij}^k + \sum_{j=1}^{N} a_i^k a_i^\phi a_j^k a_j^\phi c_{ij}^k \sum_{\substack{\mu=1,\\ \mu\neq\nu}}^{p} \xi_i^{\phi^\mu}\xi_j^{\phi^\mu}\xi_i^{\phi^\nu}\xi_j^{\phi^\nu} + \sum_{j=1}^{N}\sum_{\substack{\sigma=1,\\ \sigma\neq\phi}}^{s} a_i^\phi a_i^k a_i^\sigma a_j^k a_j^\sigma c_{ij}^k \sum_{\mu=1}^{p} \xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}\xi_i^{\phi^\nu}\xi_j^{\phi^\nu}\right). \quad \text{(F14)}$$

Unlike for the random gating case, we are no longer able to separate the effects of the different sets of patterns, so we will take an approach that follows the memory capacity results (Appendix D). First, and most straightforwardly, from the memory capacity calculations we have already derived the variance of the weights [Eq. (D3) for the Hebbian weight matrix and Eq. (E8) for the random weight matrix]. The square root of this term is the standard deviation of the weight matrix, $|J_i^k|$ [Eq. (F2)]. The other part of the calculation depends on whether or not we are measuring stability of accessible or inaccessible patterns.

a. Accessible patterns, $k = \phi$, for the Hebbian weight matrix

Accessible memory patterns will be correlated with the gating variables of the synaptic weights, leading to a nonzero mean for the stability. We can write this as

$$\kappa_i^{k,\phi,\phi^\nu} = \frac{\langle c_{ij}^k J_{ij}\rangle_{ij} + \zeta}{|J_i^k|}, \quad \text{(F15)}$$

where $\langle c_{ij}^k J_{ij}\rangle_{ij}$ was derived in Eq. (D4) and ζ is zero-mean Gaussian noise described by Eq. (D3) which can be written as

$$\zeta \sim \mathcal{N}\left(0, \frac{\alpha_{\text{cxt}}}{aN}\left(1 + a^2(s-1) + \frac{a\sqrt{s-1}}{\pi c}\right)\right). \quad \text{(F16)}$$

This results in an average stability of accessible memory patterns:

$$\bar{\kappa}_{\text{acc}} = \sqrt{\frac{c}{\alpha_{\text{cxt}}}}\frac{1 + \frac{a\sqrt{s-1}}{\pi c}}{\sqrt{1 + a^2(s-1) + \frac{a\sqrt{s-1}}{\pi c}}} = \frac{\pi c + a\sqrt{s-1}}{\sqrt{(\pi c + \pi a^2 c(s-1) + a\sqrt{s-1})\pi\alpha_{\text{cxt}}}}. \quad \text{(F17)}$$

b. Inaccessible patterns, $k \neq \phi$, for the Hebbian weight matrix

Inaccessible memory patterns are uncorrelated with the particular synaptic gating variables of the active context, and they can be treated as random variables. We thus have a setting which precisely follows the random gating setting, with two distinctions—first, the synaptic gating ratio c is set by the refinement rule instead of being pre-specified, and second, the standard deviation of the weight matrix will be different, as specified above. We thus have an average stability of inaccessible memory patterns of

$$\bar{\kappa}_{\text{inacc}} = a^2\sqrt{\frac{c}{\alpha_{\text{cxt}}}}\frac{1}{\sqrt{1 + a^2(s-1) + \frac{a\sqrt{s-1}}{\pi c}}} = \frac{a^2 c\sqrt{\pi}}{\sqrt{(\pi c + \pi a^2 c(s-1) + a\sqrt{s-1})\alpha_{\text{cxt}}}}. \quad \text{(F18)}$$

c. Accessible patterns, $k = \phi$, for the random weight matrix

Following the calculations above for the Hebbian weight matrix [Eq. (F15)], we use $\langle c_{ij}^k J_{ij}\rangle_{ij}$ derived in Eq. (E5) and arrive at

$$\bar{\kappa}_{\text{acc}} = \frac{\frac{1}{\pi\sqrt{\alpha_{\text{rand}}^*}}}{\sqrt{\frac{1}{2}}} = \sqrt{\frac{2}{\pi^2\alpha_{\text{rand}}^*}}. \quad \text{(F19)}$$

d. Inaccessible patterns, $k \neq \phi$, for the random weight matrix

Inaccessible memory patterns have, by definition, no correlation with the weight matrix drawn from a Gaussian distribution, which results in their stability being zero.

APPENDIX G: DETAILS OF NUMERICAL SIMULATIONS

We briefly describe the details of all numerical simulations here, used to generate memory capacity, information content, and stability in Figs. 3(d), 3(f), 3(h), 4(d), 4(f), 4(h), 6(c), 6(f), 8(a)–8(d), 8(h), 8(j), 8(l), 9(c), 9(d), 10(c), and 10(d). Code was written in PYTHON with the help of PYTORCH [91] for GPU-accelerated simulations. Memory capacity was estimated numerically by constructing finite-sized networks ($N_{\text{cxt}} = 4000$) initialized with a set of random patterns and connectivity as defined in the main text. To test stability of the patterns, we initialized the network in each memory state and simulated the dynamics until they either reached a steady state or they reached the maximum number of allowed time steps (50). Dynamics were run *synchronously* (i.e., all units were updated

simultaneously) according to Eq. (1), and recall performance was measured using the overlap,

$$m^{\sigma^\mu} = \frac{1}{N_{\text{cxt}}} \sum_{i=1}^{N} \xi_i^{\sigma^\mu} S_i, \tag{G1}$$

where $m^{\sigma^\mu} \in [-1, 1]$: $m^{\sigma^\mu} = 1$ when the state is exactly the pattern ξ^{σ^μ}, $m^{\sigma^\mu} = 0$ when it is uncorrelated with the pattern ξ^{σ^μ}, and $m^{\sigma^\mu} = -1$ when the state is the exact opposite (anticorrelated) of pattern ξ^{σ^μ}. We repeated this procedure for all patterns in a particular context to obtain an average overlap $\overline{m^\sigma} = p^{-1} \sum_{\mu=1}^{p} m^{\sigma^\mu}$. Capacity was set as the maximum number of patterns stored for which the average overlap was greater than or equal to the critical overlap curves from the mean-field theory in Figs. 3(a) and 3(b) and Figs. 4(a) and 4(b) for random neuronal gating and random synaptic gating, respectively, and Figs. 8(a), 8(e), and 8(f) for synaptic refinement. Synchronous dynamics were chosen for efficiency reasons despite potential convergence issues. Memory accessibility was numerically estimated by initializing networks (again of size $N_{\text{cxt}} = 4000$) at theoretical capacity, and then calculating accessible and inaccessible stabilities based on Eqs. (13)–(16).

[1] J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).

[2] D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, Cambridge, England, 1992).

[3] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, Reading, MA, 1991).

[4] R. Chaudhuri and I. Fiete, *Computational principles of memory*, Nat. Neurosci. **19**, 394 (2016).

[5] C. Miehl, S. Onasch, D. Festa, and J. Gjorgjieva, *Formation and computational implications of assemblies in neural circuits*, J. Physiol. **601**, 3071 (2022).

[6] M. V. Tsodyks and M. V. Feigelman, *The enhanced storage capacity in neural networks with low activity level*, Europhys. Lett. **6**, 101 (1988).

[7] H. Gutfreund, *Neural networks with hierarchically correlated patterns*, Phys. Rev. A **37**, 570 (1988).

[8] M. V. Tsodyks, *Hierarchical associative memory in neural networks with low activity level*, Mod. Phys. Lett. B **04**, 259 (1990).

[9] J. Steinberg and H. Sompolinsky, *Associative memory of structured knowledge*, Sci. Rep. **12**, 21808 (2022).

[10] A. Treves and E. T. Rolls, *What determines the capacity of autoassociative memories in the brain?*, Netw., Comput. Neural Syst. **2**, 371 (1991).

[11] B. Derrida, E. Gardner, and A. Zippelius, *An exactly solvable asymmetric neural network model*, Europhys. Lett. **4**, 167 (1987).

[12] C. F. Mari and A. Treves, *Modeling neocortical areas with a modular neural network*, BioSystems **48**, 47 (1998).

[13] A. Renart, N. Parga, and E. T. Rolls, *Associative memory properties of multiple cortical modules*, Netw., Comput. Neural Syst. **10**, 237 (1999).

[14] A. M. Dubreuil and N. Brunel, *Storing structured sparse memories in a multi-modular cortical network model*, J. Comput. Neurosci. **40**, 157 (2016).

[15] M. Naim, V. Boboeva, C. J. Kang, and A. Treves, *Reducing a cortical network to a Potts model yields storage capacity estimates*, J. Stat. Mech. (2018) 043304.

[16] V. Braitenberg and A. Schüz, *Anatomy of the Cortex: Statistics and Geometry* (Springer-Verlag, Berlin, 1991).

[17] A. Stepanyants, L. M. Martinez, A. S. Ferecskó, and Z. F. Kisvárday, *The fractions of short-and long-range connections in the visual cortex*, Proc. Natl. Acad. Sci. U.S.A. **106**, 3555 (2009).

[18] D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Storing infinite numbers of patterns in a spin-glass model of neural networks*, Phys. Rev. Lett. **55**, 1530 (1985).

[19] E. Gardner, *The space of interactions in neural network models*, J. Phys. A **21**, 257 (1988).

[20] E. Tulving and Z. Pearlstone, *Availability versus accessibility of information in memory for words*, J. Verbal Learning Verbal Behav. **5**, 381 (1966).

[21] S. M. Smith, *Remembering in and out of context*, Journal of experimental psychology: Human learning and memory **5**, 460 (1979).

[22] M. E. Bouton, *Context, time, and memory retrieval in the interference paradigms of Pavlovian learning*, Psychol. Bull. **114**, 80 (1993).

[23] S. Maren, K. L. Phan, and I. Liberzon, *The contextual brain: Implications for fear conditioning, extinction and psychopathology*, Nat. Rev. Neurosci. **14**, 417 (2013).

[24] S. A. Josselyn and P. W. Frankland, *Memory allocation: Mechanisms and function*, Annu. Rev. Neurosci. **41**, 389 (2018).

[25] P. W. Frankland, S. A. Josselyn, and S. Köhler, *The neurobiological foundation of memory retrieval*, Nat. Neurosci. **22**, 1576 (2019).

[26] M. C. Anderson and J. C. Hulbert, *Active forgetting: Adaptation of memory by prefrontal control*, Annu. Rev. Psychol. **72**, 1 (2021).

[27] Y. Zaki and D. J. Cai, *Memory engram stability and flexibility*, Neuropsychopharmacology **50**, 285 (2024).

[28] M. W. Howard and M. J. Kahana, *A distributed representation of temporal context*, J. Math. Psychol. **46**, 269 (2002).

[29] T. J. Ryan and P. W. Frankland, *Forgetting as a form of adaptive engram cell plasticity*, Nat. Rev. Neurosci. **23**, 173 (2022).

[30] M. N. Shadlen and D. Shohamy, *Decision making and sequential sampling from memory*, Neuron **90**, 927 (2016).

[31] M. J. Kahana, *Computational models of memory search*, Annu. Rev. Psychol. **71**, 107 (2020).

[32] S. J. Gershman, M.-H. Monfils, K. A. Norman, and Y. Niv, *The computational nature of memory modification*, eLife **6**, e23763 (2017).

[33] A. P. Yonelinas, C. Ranganath, A. D. Ekstrom, and B. J. Wiltgen, *A contextual binding theory of episodic memory:*

Systems consolidation reconsidered, Nat. Rev. Neurosci. **20**, 364 (2019).

[34] C. A. O. Coelho, A. J. Mocle, A. D. Jacob, A. I. Ramsaran, A. J. Rashid, S. Köhler, S. A. Josselyn, and P. W. Frankland, *Dentate gyrus ensembles gate context-dependent neural states and memory retrieval*, Sci. Adv. **10**, eadn9815 (2024).

[35] K. Z. Tanaka, A. Pevzner, A. B. Hamidi, Y. Nakazawa, J. Graham, and B. J. Wiltgen, *Cortical representations are reinstated by the hippocampus during memory retrieval*, Neuron **84**, 347 (2014).

[36] K. K. Cowansage, T. Shuman, B. C. Dillingham, A. Chang, P. Golshani, and M. Mayford, *Direct reactivation of a coherent neocortical memory of context*, Neuron **84**, 432 (2014).

[37] H. C. Barron, T. P. Vogels, U. E. Emir, T. R. Makin, J. O'shea, S. Clare, S. Jbabdi, R. J. Dolan, and T. E. J. Behrens, *Unmasking latent inhibitory connections in human cortex to reveal dormant cortical memories*, Neuron **90**, 191 (2016).

[38] T. P. Vogels and L. F. Abbott, *Gating multiple signals through detailed balance of excitation and inhibition in spiking networks*, Nat. Neurosci. **12**, 483 (2009).

[39] G. R. Yang, J. D. Murray, and X.-J. Wang, *A dendritic disinhibitory circuit mechanism for pathway-specific gating*, Nat. Commun. **7**, 12815 (2016).

[40] N. Y. Masse, G. D. Grant, and D. J. Freedman, *Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization*, Proc. Natl. Acad. Sci. U.S.A. **115**, E10467 (2018).

[41] H. Sompolinsky, *Neural networks with nonlinear synapses and a static noise*, Phys. Rev. A **34**, 2571 (1986).

[42] We note that it is common to express diluted network load as a function of average connectivity (i.e., p/K) instead of network size [3]. We do not follow this formulation here—though the active context imposes temporary dilution through synaptic gating, all synapses of the fully connected network are used when considering multiple contexts.

[43] H. Sompolinsky, *The theory of neural networks: The Hebb rule and beyond*, in *Heidelberg Colloquium on Glassy Dynamics*, edited by J. L. van Hemmen and I. Morgenstern (Springer, Berlin, 1987), pp. 485–527.

[44] W. K. Theumann and R. Erichsen, Jr., *Retrieval behavior and thermodynamic properties of symmetrically diluted Q-Ising neural networks*, Phys. Rev. E **64**, 061902 (2001).

[45] D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Statistical mechanics of neural networks near saturation*, Ann. Phys. (N.Y.) **173**, 30 (1987).

[46] D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Information storage in neural networks with low levels of activity*, Phys. Rev. A **35**, 2293 (1987).

[47] L. F. Abbott, *Learning in neural network memories*, Netw., Comput. Neural Syst. **1**, 105 (1990).

[48] D. Volk, *On the phase transition of Hopfield networks—Another Monte Carlo study*, Int. J. Mod. Phys. C **09**, 693 (1998).

[49] $\mathcal{B}(n; s, a^2c)$ is defined as the distribution for a binomial process which estimates the probability of n successful trials out of s independent Bernoulli trials with probability of success given by a^2c.

[50] S. Fusi and L. F. Abbott, *Limits on the memory storage capacity of bounded synapses*, Nat. Neurosci. **10**, 485 (2007).

[51] Y. Dudai, *The restless engram: Consolidations never end*, Annu. Rev. Neurosci. **35**, 227 (2012).

[52] T. F. Freund and I. Katona, *Perisomatic inhibition*, Neuron **56**, 33 (2007).

[53] A. P. Yiu, V. Mercaldo, C. Yan, B. Richards, A. J. Rashid, H.-L. L. Hsiang, J. Pressey, V. Mahadevan, M. M. Tran, S. A. Kushner, M. A. Woodin, P. W. Frankland, and S. A. Josselyn, *Neurons are recruited to a memory trace based on relative neuronal excitability immediately before training*, Neuron **83**, 722 (2014).

[54] D. M. Iascone, Y. Li, U. Sümbül, M. Doron, H. Chen, V. Andreu, F. Goudy, H. Blockus, L. F. Abbott, I. Segev, H. Peng, and F. Polleux, *Whole-neuron synaptic mapping reveals spatially precise excitatory/inhibitory balance limiting dendritic and somatic spiking*, Neuron **106**, 566 (2020).

[55] T. Branco and M. Häusser, *The single dendritic branch as a fundamental functional unit in the nervous system*, Curr. Opin. Neurobiol. **20**, 494 (2010).

[56] X. Jiang, S. Shen, C. R. Cadwell, P. Berens, F. Sinz, A. S. Ecker, S. Patel, and A. S. Tolias, *Principles of connectivity among morphologically defined cell types in adult neocortex*, Science **350**, aac9462 (2015).

[57] T. Klausberger and P. Somogyi, *Neuronal diversity and temporal dynamics: The unity of hippocampal circuit operations*, Science **321**, 53 (2008).

[58] G. Kastellakis, D. J. Cai, S. C. Mednick, A. J. Silva, and P. Poirazi, *Synaptic clustering within dendrites: An emerging theory of memory formation*, Prog. Neurobiol. **126**, 19 (2015).

[59] C. Q. Chiu, G. Lur, T. M. Morse, N. T. Carnevale, G. C. R. Ellis-Davies, and M. J. Higley, *Compartmentalization of GABAergic inhibition by dendritic spines*, Science **340**, 759 (2013).

[60] J. Urban-Ciecko, J.-S. Jouhanneau, S. E. Myal, J. F. A. Poulet, and A. L. Barth, *Precisely timed nicotinic activation drives SST inhibition in neocortical circuits*, Neuron **97**, 611 (2018).

[61] W. A. M. Wybo, B. Torben-Nielsen, T. Nevian, and M.-O. Gewaltig, *Electrical compartmentalization in neurons*, Cell Rep. **26**, 1759 (2019).

[62] E. J. Agnes and T. P. Vogels, *Co-dependent excitatory and inhibitory plasticity accounts for quick, stable and long-lasting memories in biological networks*, Nat. Neurosci. **27**, 964 (2024).

[63] G. Santoni, S. Astori, M. Leleu, L. Glauser, S. A. Zamora, M. Schioppa, I. Tarulli, C. Sandi, and J. Gräff, *Chromatin plasticity predetermines neuronal eligibility for memory trace formation*, Science **385**, eadg9982 (2024).

[64] Z. Ye, L. Shi, A. Li, C. Chen, and G. Xue, *Retrieval practice facilitates memory updating by enhancing and differentiating medial prefrontal cortex representations*, eLife **9**, e57023 (2020).

[65] H. C. Barron, T. P. Vogels, T. E. J. Behrens, and M. Ramaswami, *Inhibitory engrams in perception and memory*, Proc. Natl. Acad. Sci. U.S.A. **114**, 6666 (2017).

[66] G. Mongillo, S. Rumpel, and Y. Loewenstein, *Intrinsic volatility of synaptic connections—A challenge to the*

synaptic trace theory of memory, Curr. Opin. Neurobiol. **46**, 7 (2017).

[67] L. Susman, N. Brenner, and O. Barak, *Stable memory with unstable synapses*, Nat. Commun. **10**, 4441 (2019).

[68] O. Hardt, S.-H. Wang, and K. Nader, *Storage or retrieval deficit: The yin and yang of amnesia*, Learn. Mem. **16**, 224 (2009).

[69] S. Romani, I. Pinkoviezky, A. Rubin, and M. Tsodyks, *Scaling laws of associative memory retrieval*, Neural Comput. **25**, 2523 (2013).

[70] M. Naim, M. Katkov, S. Romani, and M. Tsodyks, *Fundamental law of memory recall*, Phys. Rev. Lett. **124**, 018101 (2020).

[71] O. H. Schnaack, L. Peliti, and A. Nourmohammad, *Learning and organization of memory for evolving patterns*, Phys. Rev. X **12**, 021063 (2022).

[72] D. Breuer, M. Timme, and R.-M. Memmesheimer, *Statistical physics of neural systems with nonadditive dendritic coupling*, Phys. Rev. X **4**, 011053 (2014).

[73] L. Kang and T. Toyoizumi, *Hopfield-like network with complementary encodings of memories*, Phys. Rev. E **108**, 054410 (2023).

[74] V. Boboeva, R. Brasselet, and A. Treves, *The capacity for correlated semantic memories in the cortex*, Entropy **20**, 824 (2018).

[75] F. Emina and E. Kropff, *Selective connectivity enhances storage capacity in attractor models of memory function*, Front. Syst. Neurosci. **16**, 983147 (2022).

[76] S. Doboli, A. A. Minai, and P. J. Best, *Latent attractors: A model for context-dependent place representations in the hippocampus*, Neural Comput. **12**, 1009 (2000).

[77] M. K. Benna and S. Fusi, *Computational principles of synaptic memory consolidation*, Nat. Neurosci. **19**, 1697 (2016).

[78] D. Krotov and J. J. Hopfield, *Dense associative memory for pattern recognition*, Adv. Neural Inf. Process. Syst. **29**, 1172 (2016).

[79] D. Krotov, *A new frontier for Hopfield networks*, Nat. Rev. Phys. **5**, 366 (2023).

[80] C. Lucibello and M. Mézard, *Exponential capacity of dense associative memories*, Phys. Rev. Lett. **132**, 077301 (2024).

[81] S. Sharma, S. Chandra, and I. Fiete, *Content addressable memory without catastrophic forgetting by heteroassociation with a fixed scaffold*, in *Proceedings of the 39th International Conference on Machine Learning. Proceedings of Machine Learning Research* (2022), Vol. 162, pp. 19658–19682, https://proceedings.mlr.press/v162/sharma22b.html.

[82] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski, *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition* (Cambridge University Press, Cambridge, England, 2014).

[83] J. A. Hertz, G. Grinstein, and S. A. Solla, *Memory networks with asymmetric bonds*, AIP Conf. Proc. **151**, 212 (1986).

[84] J. A. Hertz, G. Grinstein, and S. A. Solla, *Irreversible spin glasses and neural networks*, in *Heidelberg Colloquium on Glassy Dynamics*, edited by J. L. van Hemmen and I. Morgenstern, Lecture Notes in Physics Vol. 275 (Springer, Berlin, Heidelberg1987), pp. 538–546.

[85] S. M. Polyn, K. A. Norman, and M. J. Kahana, *A context maintenance and retrieval model of organizational processes in free recall*, Psychol. Rev. **116**, 129 (2009).

[86] J. B. Heald, D. M. Wolpert, and M. Lengyel, *The computational and neural bases of context-dependent learning*, Annu. Rev. Neurosci. **46**, 233 (2023).

[87] G. Christodoulou, T. P. Vogels, and E. J. Agnes, *Regimes and mechanisms of transient amplification in abstract and biological neural networks*, PLoS Comput. Biol. **18**, e1010365 (2022).

[88] E. J. Agnes, A. I. Luppi, and T. P. Vogels, *Complementary inhibitory weight profiles emerge from plasticity and allow flexible switching of receptive fields*, J. Neurosci. **40**, 9634 (2020).

[89] J. P. Stroud, M. A. Porter, G. Hennequin, and T. P. Vogels, *Motor primitives in space and time via targeted gain modulation in cortical networks*, Nat. Neurosci. **21**, 1774 (2018).

[90] https://github.com/wpodlaski/contextual-memory-nets

[91] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, *PYTORCH: An imperative style, high-performance deep learning library*, Adv. Neural Inf. Process. Syst. **32**, 8026 (2019).