bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

High capacity and dynamic accessibility in associative memory networks with context-dependent neuronal and synaptic gating

William F. Podlaski[*]
Centre for Neural Circuits and Behaviour, University of Oxford, OX1 3SR Oxford, United Kingdom and Champalimaud Neuroscience Programme, Champalimaud Foundation, 1400-038 Lisbon, Portugal

Everton J. Agnes[*,†]
Centre for Neural Circuits and Behaviour, University of Oxford, OX1 3SR Oxford, United Kingdom and Biozentrum, University of Basel, 4056 Basel, Switzerland

Tim P. Vogels[†]
Centre for Neural Circuits and Behaviour, University of Oxford, OX1 3SR Oxford, United Kingdom and Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria

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. Here, we bring this 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 (respectively synapses) have their activity (respectively 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 signal-to-noise and mean field analyses, we show that context-modular memory networks can exhibit substantially increased memory capacity with random neuronal gating, but not with random synaptic 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. Notably, synaptic refinement allows for patterns to be accessible in multiple contexts, stabilizing memory patterns even for weight matrices that do not contain any information about the memory patterns such as Gaussian random matrices. Lastly, we show that context modulates the relative stability of accessible versus inaccessible memories, thereby confirming that contextual control acts as a mechanism to temporarily hide or reveal particular memories. Overall, our model integrates recent ideas about context-dependent memory organization with classic associative memory models, highlights an intriguing trade-off between memory capacity and accessibility, and carries important implications for the understanding of biological memory storage and recall in the brain.

Discrete attractor networks offer a simple and powerful model for auto-associative 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[2,4,5], 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[2]. 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].

A common feature of much theoretical work on associative memory is the focus on memory capacity – the maximum number of stable patterns that the network is able to store[2,6,18,19]. While capacity is certainly a crucial quantity when considering the function of associative memory models, recent work in psychology and neuroscience has highlighted the flexibility of memory as an equally significant property. Memory *accessibility* – the relative ease by which a particular pattern can be retrieved – is not constant, and has been experimentally shown to vary over time depending upon the background context[20–25]. Such work has put forth the idea that the ability to flexibly store[26] and recall[24] memories depending upon the behavioral context may be a crucial aspect of memory's functional role. Classic models of associative memory typically feature static and uniform memory stability, and thus not only fail in capturing this crucial aspect of biological memory, but also miss out on possible computational advantages of such an approach.

Here, we introduce 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

[*] Correspondence should be addressed to
WFP: william.podlaski@gmail.com
EJA: everton.agnes@gmail.com
[†] Co-senior author
21st December 2022

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

synapses, termed *neuronal gating* and *synaptic gating*, respectively. These gating schemes, which are defined in Section 1, modify the effective network connectivity, and thereby also the accessibility of memories, i.e., their stability. Based on signal-to-noise analysis and mean-field theory, we show in Section 2 that random neuronal gating improves memory capacity considerably when compared to the classic Hopfield model, while random synaptic gating gating has only a minimal increase in memory capacity. We then introduce the concept of *refinement* in Section 3, in which synaptic gating (but not neuronal gating) can be optimized to improve memory stability. We show that memory capacity is drastically increased with synaptic refinement. Finally, we demonstrate the model's ability to stabilize memories in the active context and destabilize memories in the inactive context in Section 4, highlighting the model's powerful control over memory accessibility. 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.

1. Model

1.1. Model description. We introduce a context-dependent auto-associative memory network: the context-modular memory network. The main functionality of this model 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 can be 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 has a corresponding set of allocated neurons, N_{cxt}, while the other $N - N_{\text{cxt}}$ neurons are gated off. We use the parameter $a = N_{\text{cxt}}/N$ 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 neuron, K, with the other $N-K$ connections being gated off. We use the parameter $c = K/N_{\text{cxt}}$ 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 from the perspective of this active context, which we denote by k. Units are denoted $S_i^{(k)} \in \{\pm 1\}$ for neuron i when context k is active, with dynamics

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

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. 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 Section 3.3). Memory patterns are defined as $\xi_i^{\sigma^\mu} \in \{\pm 1\}$ for neuron i and memory μ of context σ: $\Pr(\xi_i^{\sigma^\mu} = \pm 1) = 1/2$. In the following sections, we explore connectivity matrices based on variants of Hebbian learning rules[1,6].

We note that 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 intuitive $\{0,1\}$ formulation of the network (see Appendix A).

1.2. 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 is thus intuitive 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[27], indicates the proportion of gated off synapses during recall of memories in the active context.

1.3. 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

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

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

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

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[28]. Due to analytical tractability, memory capacity is probed conditioned on an externally-imposed context without considering how context is implemented (see Discussion).

2. Random neuronal and synaptic gating

First we consider the case in which neurons and synapses are allocated uniformly at random to each context. That is, for each context k, variables a_j^k and c_{ij}^k (Eq. 1) are set to 1 or 0 with

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

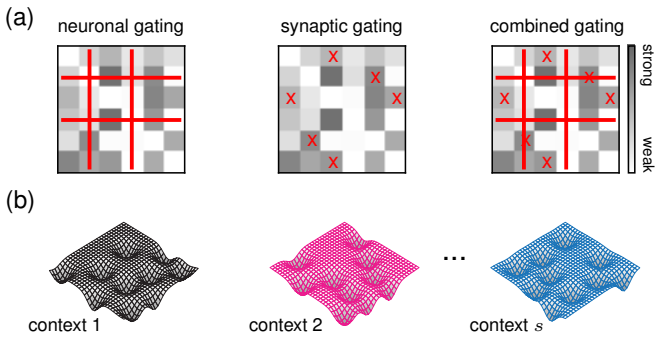

FIG. 1. The context-modular memory network. (a) Contextual configurations alter the effective connectivity matrix of the associative memory network: neuronal gating removes particular columns and rows (left), synaptic gating removes individual elements (center), and together, they implement both effects (right). (b) Each contextual state produces a distinct energy landscape with a unique set of attractor memory patterns, with others remaining hidden.

probabilities $\Pr(a_i^k = 1) = a$ and $\Pr(a_i^k = 0) = 1 - a$ for neuronal gating and $\Pr(c_{ij}^k = 1) = c$ and $\Pr(c_{ij}^k = 0) = 1 - c$ for synaptic gating. For simplicity and mathematical tractability, the synaptic gating matrix is symmetric; $c_{ij}^k = c_{ji}^k$. These quantities are used to define the weight matrix through a modified Hebbian rule as

$$J_{ij} = \frac{1}{acN} \sum_{\sigma=1}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma\mu} \xi_j^{\sigma\mu}. \tag{5}$$

Thus, synaptic weights only contain information about memory patterns for contexts in which that synapse, and the corresponding pair of neurons, are allocated. Below, we calculate the capacity of these networks conditioned on an active background context with respect to the parameters a, c, and s. We use two established heuristic methods: signal-to-noise analysis and mean-field theory.

2.1. Signal-to-noise analysis. We consider the general case of a network with s contexts, and with neuronal and synaptic gating (both random), determined by the parameters a and c defined above. We calculate memory capacity following a signal-to-noise analysis[2] that estimates the probability that a given neuron is in the incorrect state when recalling a memory pattern. To calculate this probability, we start by defining $h_i^{k\nu}$ as the total input to neuron i when the network configuration is set to pattern ν of context k. From this definition, we then calculate the probability that $\mathrm{sgn}(h_i^{k\nu}) = -\xi_i^{k\nu}$ considering that neuron i is allocated to context k, i.e., $a_i^k = 1$. Combining Eqs. 1 and 5, we obtain

$$h_i^{k\nu} = \sum_{j=1}^{N} a_j^k c_{ij}^k J_{ij} \xi_j^{k\nu} \tag{6}$$

$$= \xi_i^{k\nu} + \frac{1}{acN} \sum_{j=1}^{N} a_j^k c_{ij}^k \sum_{\substack{\mu=1,\\ \mu\neq\nu}}^{p} \xi_i^{k\mu} \xi_j^{k\mu} \xi_j^{k\nu} + \frac{1}{acN} \sum_{j=1}^{N} a_j^k c_{ij}^k \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{\sigma\mu} \xi_j^{\sigma\mu} \xi_j^{k\nu}. \tag{7}$$

The first term on the right-hand side of Eq. 7 is the memory pattern whose stability is being evaluated. The second and third terms on the right-hand side of Eq. 7 represent two sources of crosstalk noise that may disrupt the stability of stored memories. The second term is analogous to the noise term in a standard Hopfield network[2], accounting for the noise introduced by other patterns stored in the same context. The third term accounts for the noise introduced by patterns stored in different contexts. We multiply Eq. 7 by $-\xi_i^{k\nu}$, obtaining

$$-h_i^{k\nu} \xi_i^{k\nu} = -1 - \frac{1}{acN} \sum_{j=1}^{N} a_j^k c_{ij}^k \sum_{\substack{\mu=1,\\ \mu\neq\nu}}^{p} \xi_i^{k\nu} \xi_i^{k\mu} \xi_j^{k\mu} \xi_j^{k\nu} - \frac{1}{acN} \sum_{j=1}^{N} a_j^k c_{ij}^k \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{k\nu} \xi_i^{\sigma\mu} \xi_j^{\sigma\mu} \xi_j^{k\nu} \tag{8}$$

$$= -1 + \zeta_i^{k\nu}, \tag{9}$$

where $\zeta_i^{k\nu}$ represents the combined effect of the two crosstalk terms multiplied by $-\xi_i^{k\nu}$,

$$\zeta_i^{k\nu} = -\frac{1}{acN} \sum_{j=1}^{N} a_j^k c_{ij}^k \sum_{\substack{\mu=1,\\ \mu\neq\nu}}^{p} \xi_i^{k\nu} \xi_i^{k\mu} \xi_j^{k\mu} \xi_j^{k\nu} - \frac{1}{acN} \sum_{j=1}^{N} a_j^k c_{ij}^k \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} a_i^\sigma a_j^\sigma c_{ij}^\sigma \sum_{\mu=1}^{p} \xi_i^{k\nu} \xi_i^{\sigma\mu} \xi_j^{\sigma\mu} \xi_j^{k\nu}. \tag{10}$$

From Eq. 9 we note that if $\zeta_i^{k\nu} > 1$, the input to neuron i has the opposite of the desired sign. To estimate memory capacity, we can thus calculate the probability of neuron i being in the incorrect state for memory ν in context k, $\Pr(\zeta_i^{k\nu} > 1)$.

We observe that, due the effects of the gating variables, both terms on the right-hand side of Eq. 10 are sums of discrete random variables, $\{-1, 0, 1\}$. Considering a sequence of Bernoulli random variables X_j such that $\Pr(X_j = 0) = \Pr(X_j = 1) = 1/2$, we can approximate $\zeta_i^{k\nu}$ as

$$\zeta_i^{k\nu} \approx \frac{1}{acN} \sum_{j=1}^{\rho_1+\rho_2 z} [2X_j - 1], \tag{11}$$

with the limits of the sum defined as

$$\rho_1, \rho_2 \sim \mathcal{B}(Np, ac) \tag{12}$$

$$z \sim \mathcal{B}(s-1, a^2 c), \tag{13}$$

where $\mathcal{B}(x, y)$ is a binomial distribution with x trials and probability of success y. Here, we can intuitively interpret ρ_1 and ρ_2 as the number of non-zero terms in the sums over j and μ, and z as the number of non-zero terms in the sum over σ in Eq. 10. In the thermodynamic limit, we can assume that ρ_1 and ρ_2 (Eq. 12) are well approximated by Gaussian noise. However, since the number of contexts s is finite, z is not well approximated by a Gaussian and must be treated differently. We thus assume that for pattern ν of context k, the realization of z takes a value $z_i^{k\nu}$ for neuron i. With this assumption, $\zeta_i^{k\nu}$ can be well-approximated as a Gaussian random variable, and using Wald's equation (Appendix B) we obtain

$$\zeta_i^{k\nu} \sim \mathcal{N}\left(0, \frac{\alpha_{\mathrm{cxt}}(1 + z_i^{k\nu})}{c}\right). \tag{14}$$

Here we take a shortcut by noting a correspondence between the noise term $\zeta_i^{k\nu}$ and previous results for the standard and diluted Hopfield networks[2,18,27]. Capacity should be reached

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

when the variance approaches that of the standard Hopfield (for neuronal gating alone, $c = 1$) or diluted Hopfield (for synaptic gating, $c < 1$) networks. In those classic models, the crosstalk variance is approximately $\frac{1}{c}\alpha_{\mathrm{D}}^*(c)$ (Appendix C), where $\alpha_{\mathrm{D}}^*(c = 1) = \alpha_{\mathrm{H}}^* \approx 0.138$. We thus define an equality between the variance of Eq. 14 and the diluted Hopfield variance. By rearranging terms and averaging over all neurons, patterns and contexts, we obtain a single-context capacity of

$$\alpha_{\mathrm{cxt}}^* = \left\langle \frac{\alpha_{\mathrm{D}}^*(c)}{1+z} \right\rangle_{z\sim\mathcal{B}(s-1,a^2c)}, \tag{15}$$

where $\langle f(z)\rangle_{z\sim\mathcal{B}(x,y)}$ denotes an average over the binomial random variable z. The denominator of Eq. 15 can be interpreted as the effective number of contexts whose patterns contribute to the crosstalk noise — the number 1 refers to the currently-active context, which always contributes, and z refers to the contribution from other contexts, which depends on the parameters s, a, and c. In other words, each context can be viewed as a standalone diluted Hopfield network of size N_{cxt} storing p patterns, but with additional noise in the weights due to the influence of effectively zp other memories. The denominator $1+z$ is always greater than or equal to 1, and so the single-context capacity is upper bounded by the diluted Hopfield network limit ($\alpha_{\mathrm{cxt}}^* \leqslant \alpha_{\mathrm{D}}^*(c), 0 \leqslant c \leqslant 1$).

From Eq. 15, we obtain a total capacity

$$\alpha^* = \left\langle \frac{as\alpha_{\mathrm{D}}^*(c)}{1+z} \right\rangle_{z\sim\mathcal{B}(s-1,a^2c)}. \tag{16}$$

Eqs. 15 and 16 apply generally for cases of arbitrary numbers of contexts, s, proportion of gated neurons, $1-a$, and proportion of gated synapses, $1-c$. We make several observations about the signal-to-noise estimates of single-context capacity, α_{cxt}^*, and total capacity, α^* (Fig. 2). Single-context capacity is upper-bounded by the Hopfield limit (Fig. 2(a,b)). For fixed proportion of gated neurons or synapses, the single-context capacity decreases as the number of contexts, s, increases (Fig. 2(a,b)). We observe that for high proportions of gated neurons ($1-a \leqslant 0.75$ and $c = 1$), single-context capacity degrades gradually (Fig. 2(a)), leading to large total capacity (Fig. 2(c)). For synaptic gating alone ($a = 1$), single-context capacity is very small (Fig. 2(b)) which leads to only a modest improvement in total capacity (Fig. 2(d)).

Memory capacity is well described by Eq. 16, but it cannot provide a direct intuition of how different parameters affect capacity. To arrive at such an intuition, we consider the simplified case in which each neuron is gated in the same number of contexts. This is equivalent to replacing the variable z with its expected value in Eqs. 15 and 16. Doing so, we arrive at a single-context capacity given by

$$\alpha_{\mathrm{cxt}}^* = \frac{\alpha_{\mathrm{D}}^*(c)}{1+a^2c(s-1)}, \tag{17}$$

and a total capacity of

$$\alpha^* = \frac{as\alpha_{\mathrm{D}}^*(c)}{1+a^2c(s-1)}. \tag{18}$$

We can make several observations from these simplified equations. First, the optimal neuronal gating ratio, $1-a_{\mathrm{opt}}$, for a fixed number of contexts and no dilution ($c = 1$) is $1-a_{\mathrm{opt}} = 1-s^{-1/2}$ (Fig. 3(a), red line). Furthermore, taking the limit of large s

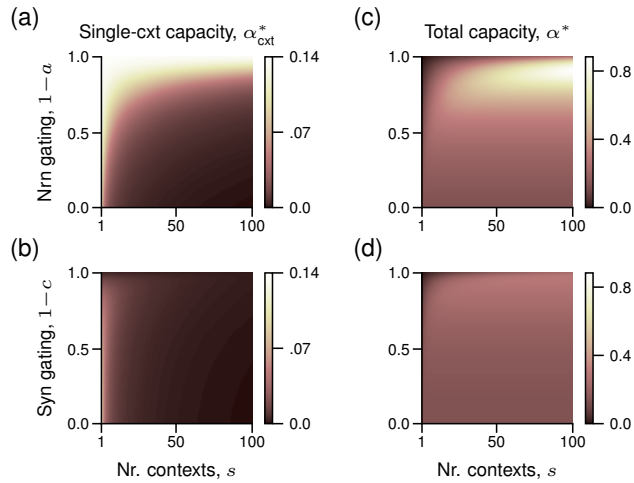

FIG. 2. Memory capacity of the context-modular memory network with random neuronal and synaptic gating. (a) Capacity estimation for single-context capacity, α_{cxt}^*, as a function of the number of contexts, s, and neuronal gating, $1-a$, for neuronal gating alone ($c = 1$) from Eq. 15. (b) Capacity estimation for single-context capacity, α_{cxt}^*, as a function of the number of contexts, s, and synaptic gating, $1-c$, for synaptic gating along ($a = 1$) from Eq. 15. (c,d) Total capacity for the networks considered in panels (a,b) based on Eq. 16.

(such that $1 + a^2(s-1) \approx a^2 s$), the total capacity approaches $\alpha^* \approx \alpha_{\mathrm{H}}/a$ (Fig. 3(b), red line), and thus $\alpha^* \approx \alpha_{\mathrm{H}}\sqrt{s}$ considering optimal neuronal gating. It follows that the high capacity emerges due to the patterns having low activity, growing sublinearly as a function of the number of contexts. The optimal synaptic gating ratio, $1-c_{\mathrm{opt}}$, cannot be easily calculated in our simplified analysis because $\alpha_{\mathrm{D}}(c)$ does not have a closed-form solution[27]. However, $1-c_{\mathrm{opt}} \propto 1-s^{-3/4}$ fits well the optimal synaptic gating ratio (Fig. 3(a), purple line). Taking the limit of large s for synaptic gating alone ($a = 1$), the total capacity approaches $\alpha^* \approx \alpha_{\mathrm{D}}(c)/c$ (Fig. 3(b), purple line). We also compared the results to a low-activity variant of the Hopfield network[6] (Fig. 3(b), dashed gray line; Appendix). While the low-activity capacity is larger, it is only achieved for very sparse activity levels compared to neuronal gating. Overall, we see that the context-modular memory network uses contextual gating to harnesses sparsity in activity and connectivity in a similar way to previous models.

2.2. Mean-field theory. To validate our signal-to-noise results, we confirmed them using a heuristic mean-field approach[2]. We consider a stochastic version of the Hopfield network, with dynamics

$$\begin{cases} \Pr\left(S_i^{(k)} = \pm 1\right) = \frac{1}{1+\exp\left(\mp 2\beta h_i^k\right)} & \text{if } a_i^k = 1 \\ \Pr\left(S_i^{(k)} = 0\right) = 1 & \text{otherwise,} \end{cases} \tag{19}$$

where h_i^k is the input to neuron i and $\beta = 1/T$ is the inverse of the temperature, T. We replace the fluctuating input h_i^k in Eq. 19 by its average value $\langle h_i^k\rangle = \sum_{j=1}^N a_j^k c_{ij}^k J_{ij} \langle S_j^{(k)}\rangle$ (where $\langle\cdot\rangle$ represents thermal average), to obtain the average activation

$$\left\langle S_i^{(k)}\right\rangle = a_i^k \tanh\left(\beta\left\langle h_i^k\right\rangle\right) = a_i^k \tanh\left(\beta \sum_{j=1}^N a_j^k c_{ij}^k J_{ij} \left\langle S_j^{(k)}\right\rangle\right), \tag{20}$$

if context k is active.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

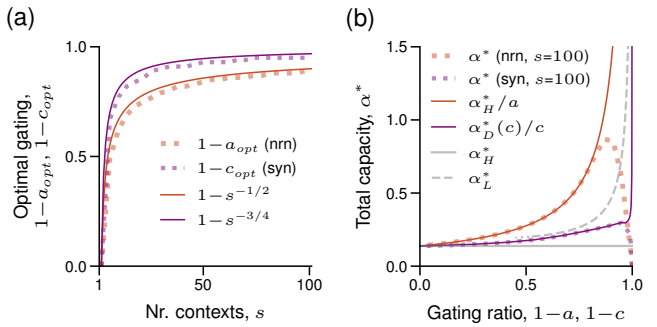

FIG. 3. Optimal gating parameter and simplified total capacity. (a) Optimal neuronal gating ratio, $1-a_{\mathrm{opt}}$, for neuronal gating alone ($c=1$, red), and optimal synaptic gating ratio, $1-c_{\mathrm{opt}}$, for synaptic gating alone ($a=1$, purple), as a function of the number of contexts, s. Dashed lines are the optimal values, numerically estimated from Eqs. 15 and 16. Full lines are the functions indicated in the legend. (b) Total capacity, α^*, as a function of neuronal and synaptic gating ratios, $1-a$ and $1-c$, respectively. Dashed thick lines are the estimated capacity for a fixed number of contexts ($s=100$) as a function of the neuronal gating ratio, $1-a$, for synaptic gating alone ($c=1$, red), and synaptic gating ratio, $1-c$, for synaptic gating alone ($a=1$, purple) from Eq. 15. Red and purple full lines are the simplified estimated capacity with functions indicated in the legend. Full gray line is the Hopfield limit, α^*_{H}, and dashed gray line is the estimated capacity of the low-activity Hopfield network, α^*_{L} (Appendix C).

Based on the mean-field theory (see Appendix D for detailed calculation), we arrive at the following three equations for the context-modular memory network,

$$m_1^k = \frac{1}{aN}\sum_{i=1}^{N} a_i^k \xi_i^{k^1} \left\langle S_i^{(k)} \right\rangle = \left\langle\!\!\left\langle \tanh\left(\beta(m_1^k + \sigma^k\eta)\right)\right\rangle\!\!\right\rangle_\eta \tag{21}$$

$$r^k = \frac{1}{\alpha_{\mathrm{cxt}}}\sum_{\mu=2}^{p}\left(m_\mu^k\right)^2 = \frac{q^k}{[1-\beta(1-q^k)]^2} \tag{22}$$

$$q^k = \left\langle\!\!\left\langle \tanh^2\left(\beta(m_1^k + \sigma^k\eta)\right)\right\rangle\!\!\right\rangle_\eta, \tag{23}$$

where m_1^k is the *overlap* between the average network state $\langle S_i^{(k)}\rangle$ and pattern 1 of the active context k; r^k is the *mean square overlap* of the system configuration with the non-retrieved patterns in context k; q^k is the Edwards-Anderson order parameter[29,30]; $\langle\!\langle\cdot\rangle\!\rangle_\eta$ represents a Gaussian average over the variable $\eta \sim \mathcal{N}(0,1)$; and σ^k represents the crosstalk noise on pattern 1 of context k due to all other patterns. Then, writing m_1^k, r^k, q^k and σ^k as m, r, q and σ, respectively, we estimate the noise as

$$\sigma^2 = \tilde{\alpha} r + \frac{\tilde{\alpha}(1-c)}{c} q, \tag{24}$$

where $\tilde{\alpha}$ is given by

$$\tilde{\alpha} = \alpha_{cxt}(1+z), \tag{25}$$

and $z \sim \mathcal{B}(s-1, a^2c)$ is a binomial random variable (analogous to Eq. 13). For expository purposes, we can consider the expected value for z to obtain

$$\tilde{\alpha} = \alpha_{cxt}\left(1 + (s-1)ca^2\right). \tag{26}$$

We observe that setting a, c, and s to 1 recovers the original Hopfield network capacity with $\sigma^2 = \alpha_{\mathrm{H}} r$. Furthermore, for $s=1$ and $a=1$, we can recover the diluted Hopfield capacity, with $\sigma^2 = \alpha_{\mathrm{D}} r + \alpha_{\mathrm{D}}(1-c)q/c$. We solve the mean-field equations in the zero temperature limit ($\beta \to \infty$)[2], obtaining precisely the same memory capacity as Eq. 16, thereby validating the signal-to-noise analysis. While the signal-to-noise analysis usually requires an arbitrary noise cutoff value[2], we overcame this by directly comparing to the mean-field results for the standard and diluted Hopfield networks, α^*_{H} and $\alpha^*_{\mathrm{D}}(c)$.

Importantly, the mean-field theory can provide an estimate of the overlap between the stored memory patterns and the attractor states of the dynamics at capacity — such a measure is needed for numerical estimations of capacity. The level of overlap is well captured by the standard Hopfield network when there is no dilution ($c=1$), and it's well described by the levels of dilution otherwise[27] (Appendix C). We can thus write the mean-field overlap $m(c,a,s)$ as a function of the diluted network result, as

$$m(c,a,s) = m_{\mathrm{D}}(c). \tag{27}$$

For neuronal gating alone ($c=1$), the overlap between memory patterns and the fixed-points of the network's dynamics is independent of a and s ($m^* \approx 0.97$; Fig. 4, red line and dots)[30]. For synaptic gating alone ($a=1$), the overlap is independent of s, but decreases with increasing dilution, $1-c$ (Fig. 4, purple line and dots)[27]. This indicates that there is increased error in recall for large amounts of dilution ($c \ll 1$).

To validate our analyses, we compared the theoretical results with numerical simulations of size $N = 10{,}000$ (Figs. 4, 5; simulation details can be found in Appendix G). We first verified the accuracy of the theoretical overlap, m^*, by simulating networks at theoretical capacity and measuring the average overlap. Simulations of neuronal gating alone showed overlap values very close to the theoretical prediction (Fig. 4, red points). However, while the simulations of synaptic gating alone resulted in a decreased overlap as the synaptic gating ratio, $1-c$, increased, it was substantially higher than expected, and highlights the difficulties in numerically simulating such diluted networks.

As for memory capacity, simulations and theory behave similarly for neuronal gating alone ($c=1$; Fig. 5(a,b)), with discrepancies likely due to finite-size effects, and the approximations in the analysis described above. For synaptic gating alone ($a=1$), there is an increased discrepancy between simulation and theory (Fig. 5(c,d)), again due to finite-size effects, but also due to the difficulty in achieving an accurate overlap cutoff, as mentioned above (Fig. 4).

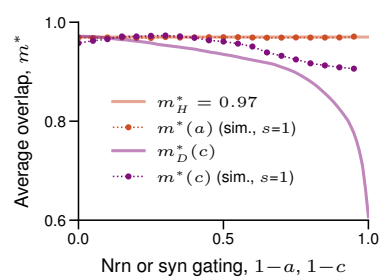

FIG. 4. Mean-field overlap. Average mean-field overlap between memory patterns and fixed-points of the network's dynamics for neuronal gating alone ($c=1$, red) and synaptic gating alone ($a=1$, purple) as a function of neuronal and synaptic gating ratios, $1-a$ and $1-c$, respectively. Full lines indicate theoretical results from mean-field theory. Circles connected by dotted lines indicate results from simulations for networks at theoretical capacity for a single context ($s=1$). See Appendix G for details.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

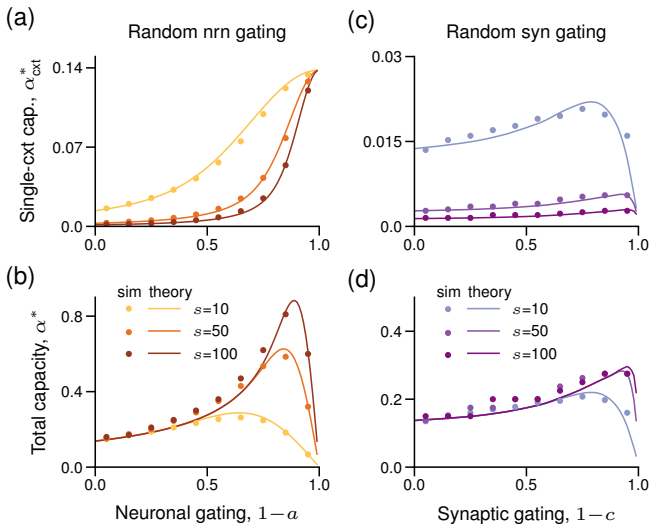

FIG. 5. Comparison between mean-field theory and simulations for random neuronal and synaptic gating. (a,b) Mean-field capacity estimation (solid lines) and numerical simulations (dots) for single-context capacity, $\alpha^*_{\rm cxt}$ (panel (a)) from Eq. 15, and total capacity, α^* (panel (b)) from Eq. 16, as a function of neuronal gating ratio, $1-a$, for neuronal gating alone ($c = 1$). Plots for three different number of context, $s = 10$, $s = 50$, and $s = 100$. (c,d) Same as panels (a,b) but as a function of synaptic gating ratio, $1-c$, for synaptic gating alone ($a = 1$). $N = 10{,}000$ for all simulations (Appendix G).

2.3. Information content. While high storage capacity is a relevant and important aspect of an associative memory network, the information stored per pattern is also an essential aspect, as both should be maximized. The information content can be calculated as the total entropy (Shannon information) across all retrieved patterns[19]. This calculation is straightforward for neuronal gating, because the average overlap between the minima of the dynamics (retrieved patterns) and the original patterns is close to unity, $m \approx 0.97$ (Fig. 4, red line). Due to the low values of overlap for synaptic gating (Fig. 4), we calculate the information content as the total entropy considering the original memory patterns minus the *missing information* due to inaccurate recall[31]. Based on the above definition, we obtain the following information content of the context-modular network (see Appendix E for detailed calculation),

$$I = spN_{\rm cxt}\left[S_{1/2} - S(m_{\rm D}(c))\right] \tag{28}$$

$$= \alpha a N^2 \left[\frac{1+m_{\rm D}(c)}{2}\log_2\left(1+m_{\rm D}(c)\right) + \frac{1-m_{\rm D}(c)}{2}\log_2\left(1-m_{\rm D}(c)\right)\right], \tag{29}$$

where $S_{1/2}$ is the entropy per neuron given that $\Pr(\xi_i^{\sigma^\mu} = \pm 1) = 1/2$, and $S(m)$ is the missing information given that the average overlap between original memory patterns and corresponding attractors is m[31]. In the following, we consider information content relative to that of the standard Hopfield network (Appendix E).

The information content of the context-modular memory network with random neuronal gating alone ($c = 1$) and $s = 100$ contexts, when at maximum capacity ($\alpha = \alpha^*$ in Eq. 29), follows closely that of the standard Hopfield network at maximum capacity ($\alpha a = \alpha^*_{\rm H}$ and $m_{\rm D}(c) = m^*_{\rm H}$ in Eq. 29) for small gating ratios $1 - a < 0.75$ (Fig. 6, red). For larger gating ratios, we observe that the information content drops off steeply to zero. Similarly to storage capacity (Fig. 3(b)), this discrepancy is due to the dependency on the number of contexts, s, and suggests that a similar information content may be obtained for an arbitrarily high gating ratio provided that s is large enough. The information content of a network with random synaptic gating alone ($a = 1$) and $s = 100$ contexts is similar in magnitude to that of the standard Hopfield network, and grows larger with increased gating ratio, $1 - c$ (Fig. 6, purple). Analogously to neuronal gating, the information content for synaptic gating drops to zero as the gating ratio approaches 1, due to a finite number of contexts s.

Overall, we see that while storage capacity grows large for context-modular memory networks with high gating ratios $1-a$ and $1 - c$ (Fig. 3(b), red and purple lines), information content remains similar to the standard Hopfield network (Fig. 6). This is similar to the classic low-activity variant[6], whose storage capacity diverges as activity levels approach zero (Fig. 3(b) dashed grey line), but whose information content remains finite ($I^*_{\rm L} = N^2/(2\ln 2) \approx 6 I^*_{\rm H}$[6,19]).

3. Synaptic refinement

We have shown that random synaptic gating does not provide considerable improvements to memory capacity. We thus devise a more selective way of imposing synaptic control which we refer to as synaptic *refinement*. Connections are refined after learning of memory patterns, so we consider that synaptic gating can only act on the recall dynamics (Eq. 1) and not on the weight matrix, J_{ij}. We thus define the weight matrix based on the standard Hebbian learning rule (including random neuronal gating),

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

To intuitively describe the refinement scheme, we first consider the case of no neuronal gating ($a = 1$). In this case, Eq. 30 sets the synaptic weight between each pair of neurons according to the correlation in their activity across all patterns, acting to stabilize the majority of patterns[2]. However, the set of memories in the active context may produce substantially different correlations, such that some synaptic weights decrease the stability of the majority of memories in this active set. Such synaptic weights have a net harmful effect on memory recall, and performance would improve if these weights were gated off. This reasoning leads to a well-defined criterion for synaptic gating: if there is a mismatch between the sign of the synaptic weight serving to stabilize all memories versus serving to stabilize memories belonging to a specific context, then this synapse is gated off for this context. We then observe that this criterion can be combined with the inclusion of random neuronal gating ($a < 1$) before learning without one affecting the other.

More precisely, we find an optimal set of synaptic gates c_{ij}^σ for stability of memory patterns in each context. First, we define a hypothetical "isolated" weight matrix for each individual context, e.g., context σ, with elements

$$J_{ij}^\sigma = \frac{a_i^\sigma a_j^\sigma}{aN}\sum_{\mu=1}^{p} \xi_i^{\sigma^\mu}\xi_j^{\sigma^\mu}. \tag{31}$$

Notice that Eq. 31 only incorporates patterns from context σ, such that $J_{ij} = \sum_{\sigma=1}^{s} J_{ij}^\sigma$. Then, the optimized gating state of each

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

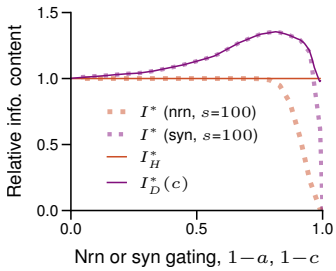

FIG. 6. Information content. Relative information content of the context-modular memory network with random neuronal and synaptic gating. Dashed lines shows the information content of random neuronal gating alone ($c = 1$ and $s = 100$, red) and random synaptic gating alone ($a = 1$ and $s = 100$, purple) as a function of neuronal gating parameter, $1 - a$, and synaptic gating parameter, $1 - c$, normalized by the information content of the standard Hopfield network (solid red line). Solid purple shows the information content of the diluted Hopfield network as a function of connection density, $1 - c$, normalized by the information content of the standard Hopfield network.

synapse is defined as

$$c^{\sigma}_{ij} = \begin{cases} 0 & \text{if } \left(J_{ij}J^{\sigma}_{ij}\right) < 0 \\ 1 & \text{otherwise.} \end{cases} \tag{32}$$

This method ensures that for context σ, the sign of each ungated weight reflects the sign of the correlation between neurons i and j over patterns in context σ, but not over all other patterns. Due to the symmetry in the weight matrices J_{ij} and J^{σ}_{ij}, synaptic gating is symmetric ($c^{\sigma}_{ij} = c^{\sigma}_{ji}$). Furthermore, and importantly, the synaptic gating ratio, $1 - c$, is no longer a predetermined parameter of the network, but is a function of the number of contexts, s, and the neuronal gating ratio, $1 - a$.

3.1. Synaptic gating ratio. To estimate the expected number of gated synapses as a function of network parameters s and a, we assume that the overall weight distributions from Eqs. 30 and 31 are Gaussian with zero mean and variances proportional to the amount of crosstalk that they contribute, which is determined by a binomial random variable $z \sim \mathcal{B}(s-1, a^2)$ analogous to the signal-to-noise analysis from Section 2.1. This leads to synaptic gating given by (see Appendix F for detailed calculation)

$$1 - c(s,a) = \frac{1}{\pi}\left\langle \arctan\left(\sqrt{z}\right)\right\rangle_{z\sim\mathcal{B}(s-1,a^2)}. \tag{33}$$

For synaptic refinement alone ($a = 1$), we observe that the synaptic gating ratio, $1 - c$, quickly approaches 0.5 (i.e., 50%) as the number of contexts increases (Fig. 7(a)). This trend also holds for synaptic refinement in combination with random neuronal gating (Fig. 7(b)) for parameter ranges with high overlap between contexts (low neuronal gating, $1-a$, and large number of contexts, s). 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 $N = 10{,}000$ and confirm that our approximation for the overall connectivity holds (Fig. 7(c)).

3.2. Heuristic estimation of storage capacity and information content. To estimate the memory capacity with synaptic refinement, we need to take into account optimized synaptic gating. Synaptic refinement introduces correlations between the two crosstalk terms in Eq. 10, i.e., the within-context crosstalk and across-context crosstalk are no longer independent. We observe that, when neuronal gating is high ($1 - a > 0.9$), few connections are gated, and the single-context capacity is close to that of the standard Hopfield network ($\alpha^*_{\text{cxt}} = \alpha^*_{\text{H}} \approx 0.138$). For lower amounts of neuronal gating (e.g., $1 - a < 0.5$), synaptic gating rises to 0.5. In this regime, a given synapse is either zero (with ~ 50% probability) or assumes a value with the same sign as the correlation between all patterns in the active context, i.e., the sign of that synapse for a weight matrix of the active context (Eq. 31). However, the value of a given synapse may not reflect the actual correlation between the patterns in the active context, due to the noise added by the patterns stored in all inactive contexts. Such a network may thus be well approximated by a Hopfield network with binary weights[2,32], but also scaled roughly linearly by sparsity[27]. The connections of a Hopfield network with binary weights follows a standard Hebbian learning rule that is passed through a sign-function, and weights are set to +1 or −1. In this case, the single-context capacity of a context-modular memory network can be approximated as

$$\alpha^*_{\text{cxt}} \approx \tfrac{1}{2}\alpha^*_{\text{B}} = 0.05, \tag{34}$$

where $\alpha^*_{\text{B}} = 0.1$ is the memory capacity of a network with binary synapses[32]. Therefore, to a rough approximation, we may be able to estimate the capacity in between two extremes – no synaptic gating, $c(s,a) \approx 1$, and half of connections gated-off, $c(s,a) \approx 0.5$ – by linearly interpolating between α_{H} and $\frac{1}{2}\alpha^*_{\text{B}}$. This interpolated estimation can be written as

$$\alpha^*_{\text{cxt}} = c(s,a)\left([2c(s,a)-1]\alpha^*_{\text{H}} + 2[1-c(s,a)]\alpha^*_{\text{B}}\right), \tag{35}$$

where $1 - c(s,a)$ is the synaptic gating ration defined by Eq. 33.

Based on this approximation, we observe that single-context capacity for synaptic refinement alone ($a = 1$) is upper-bounded by the Hopfield limit, α^*_{H}, decreasing as the number of contexts increases (Fig. 7(d)), and following similar trends as for random neuronal gating alone for different neuronal gating ratios, $a < 1$ (Fig. 7(e,f)). However, this decrease in single-context capacity appears to have a lower bound as well, near $\frac{1}{2}\alpha^*_{\text{B}} \approx 0.05$ (Fig. 7(d,f), dashed gray line), as predicted by the estimated theoretical capacity (Eq. 35). This finite lower bound leads to a total capacity that increases roughly linearly with the number of contexts, s, and decreases roughly linearly with neuronal gating ratio, $1 - a$ (Fig. 7(g,h,i)), growing to large values ($\alpha^* \approx 5$ for $s = 100$ contexts) – up to 30 times that of the standard Hopfield network (Fig. 7(g)). Numerical simulations confirmed the theory both for synaptic refinement alone (Fig. 7(a,d,g)) and in combination with random neuronal gating (Fig. 7(c,f,i)). For both the theoretical and numerical results, we observe that the optimal neuronal gating is zero, i.e., $a_{\text{opt}} = 1$, and so the addition of neuronal gating to synaptic refinement does not further increase memory capacity.

The information content of the context-modular memory network with synaptic refinement grows to much larger values than both standard and low-activity Hopfield networks. Writing the information content as a function of the information content of a standard Hopfield network, we arrive at $I^* = a\alpha^*/\alpha^*_{\text{H}}$. This means that the information content grows

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

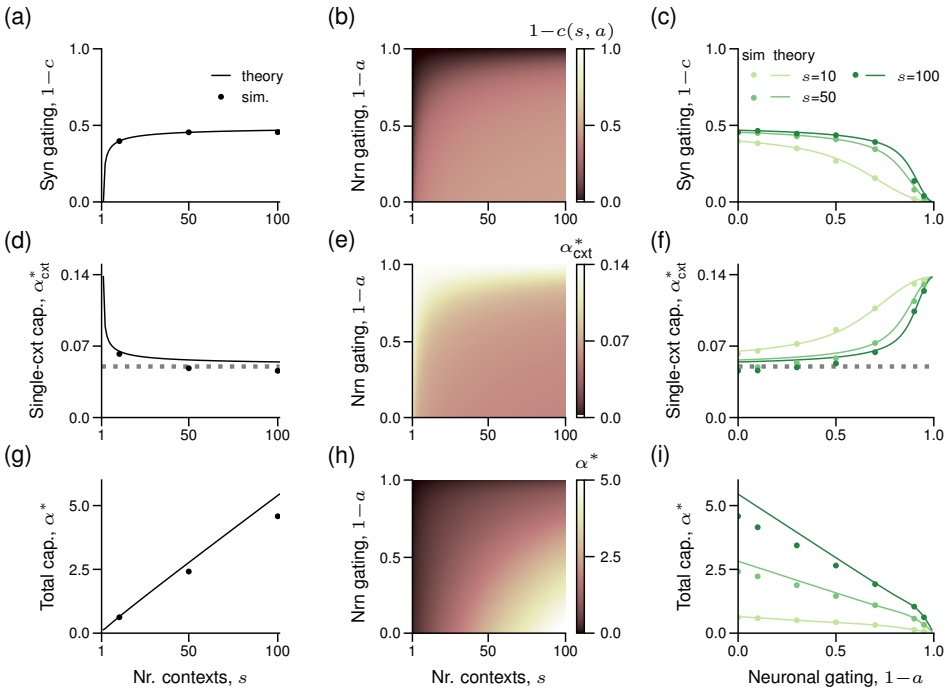

FIG. 7. Synaptic gating ratio and memory capacity following synaptic refinement. (a-c) Estimation of the resulting synaptic gating ratio, $1 - c$ (Eq. 33), following synaptic refinement alone (panel (a)) or synaptic refinement with random neuronal gating (panels (b,c)). Synaptic gating ratio as a function of: number of contexts, s (panel (a)); number of contexts, s, and neuronal gating, $1 - a$ (panel (b)); and neuronal gating, $1 - a$ (panel (c)). (d-f) Estimation of single-context capacity, α^*_{cxt} (Eq. 35), for synaptic refinement alone (panel (d)) or synaptic refinement with random neuronal gating (panels (e,f)). Single-context capacity as a function of: number of contexts, s (panel (d)); number of contexts, s, and neuronal gating, $1 - a$ (panel (e)); and neuronal gating, $1 - a$ (panel (f)). (g-i) Total capacity estimation, $\alpha^* = \alpha^*_{\text{cxt}} sa$ (Eq. 35), for synaptic refinement alone (panel (g)) or synaptic refinement with random neuronal gating (panels (h,i)). Total capacity as a function of: number of contexts, s (panel (g)); number of contexts, s, and neuronal gating, $1 - a$, (panel (h)); and neuronal gating, $1 - a$ (panel (i)). Comparison with numerical simulations (dots) for three different numbers of contexts in panels (a,c,d,f,g,i): $s = 10$, $s = 50$, and $s = 100$. All numerical simulations from panels (a,c,d,f,g,i) were done with $N = 10{,}000$ (Appendix G).

with the total capacity, α^*. For example, for $a = 1$ and $s = 100$, $I^* \approx 36 I^*_{\text{H}} \approx 6 I^*_{\text{L}}$. For the total capacity, α^*, and the total information context, I^*, to be high, the contextual control requires that half of synapses to be gated off for each context. However, this type of control requires many more additional neurons for autonomous implementation, which decreases the overall information content (see Discussion).

3.3. Arbitrary allocation of memories with synaptic refinement. Synaptic refinement is only applied to the network's recall dynamics, and thus memories do not have a strict grouping as they did for random gating. We therefore momentarily take a different perspective on context-dependent memory, in which a base set of memories is stored in the weight matrix, and then arbitrary memory subsets may be stabilized in different contexts at the time of recall. To illustrate this feature, we first store P memory patterns without explicitly specifying contextual states,

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{P} \tilde{\xi}^{\mu}_i \tilde{\xi}^{\mu}_j, \tag{36}$$

where $\tilde{\xi}^{\mu}_i$ ($i = 1 \ldots N$ and $\mu = 1 \ldots P$) are context-independent memory patterns defined by $\Pr(\tilde{\xi}^{\mu}_i = \pm 1) = 1/2$. 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^{σ}_{ij} that enables successful memory recall (Eq. 32). Importantly, this is only possible because all neurons remain ungated in each context, even if individual connections are gated off, which means that $a = 1$.

We observe that arbitrary sets of memory patterns can be reliably retrieved (Fig. 8(a)), provided that the number of patterns stored in the active context is sufficiently small ($\alpha_{\text{cxt}} \leqslant 0.03N$). Performance begins to degrade when the number of patterns per context approaches the single-context capacity, α^*_{cxt}, which depends on the overall load, α. Intuitively, the single-context capacity, α^*_{cxt}, decreases with the increase in total load, α (Fig. 8(b)). 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 (not shown). This indicates that synaptic refinement can be used at any moment in time after memories are learned to stabilize an arbitrary group of memories, even if some of these memories are also allocated to other contexts.

3.4. Stabilization of memory patterns in a random weight matrix with synaptic refinement. 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

$$J_{ij} \sim \mathcal{N}(0, 1). \tag{37}$$

We then construct a set of p memory patterns, $\tilde{\xi}^{\mu}_i$ ($i = 1 \ldots N$ and $\mu = 1 \ldots p$), and define the gating variables c^{σ}_{ij} for the active context σ following Eq. 32. Remarkably, a random weight matrix can be successfully used to represent stable context-dependent memories (Fig. 8, dashed lines), as long as the single-context load is small ($\alpha_{\text{cxt}} \leq 0.03$).

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. 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 biolo-

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

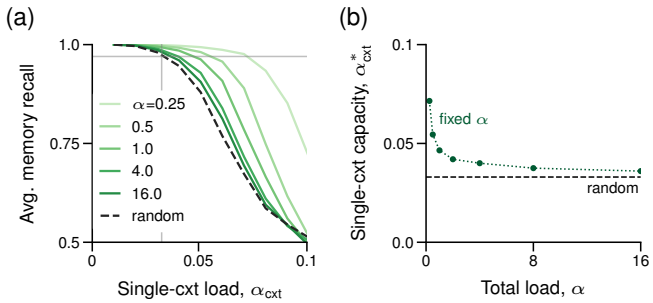

FIG. 8. Arbitrary memory allocation through synaptic refinement. (a) Network performance as measured numerically by the average overlap between initial pattern and steady-state across all patterns for the active context, plotted for various total loads, α, as a function of the single-context load, α_{cxt}. Simulation with a random weight matrix is shown by the grey dashed line. $N = 1000$ for all simulations. (b) Single-context capacity, α^*_{cxt}, as a function of total load, α, estimated numerically from the simulations shown in panel (a). Single-context capacity for the random weight matrix is plotted as the dashed black line, also estimated from panel (a).

gical memory[33]. In other words, by precisely controlling the synaptic gating variables, it is possible to recall a set of memories from a weight matrix that doesn't contain any information about those memories.

4. Accessibility (stability)

A crucial feature of the model is the assumption that only a subset of memories need to be stable in each context. While storage capacity provides an indirect measure of memory stability for accessible memory patterns, our model also assumes that inaccessible memories are hidden, and thus unstable, in contexts for which they do not belong. Furthermore, as the number of contexts, s, increases, the proportion of accessible memories decreases as $1/s$ (Fig. 9(a)), 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. To confirm that accessible memories remain stable while inaccessible memories remain unstable, we measured the stability of all memories.

To do so, we measure the so-called *stability parameter* of each memory pattern[34], which provides a rough measure of the size of each pattern's basin of attraction. We write the stability of a particular pattern ν of context σ (for an arbitrary neuron i) when a (potentially) different context k is active as

$$\kappa_i^{k,\sigma^\nu} = \frac{a_i^k a_i^\sigma \xi_i^{\sigma^\nu} h_i^{k,\sigma^\nu}}{|J_i^k|} = \frac{a_i^k a_i^\sigma \xi_i^{\sigma^\nu} \sum_{j=1}^N a_j^k c_{ij}^k J_{ij} \xi_j^{\sigma^\nu}}{|J_i^k|}, \tag{38}$$

where h_i^{k,σ^ν} denotes the input to unit i when context k is active but the network is representing pattern ν of context σ, and

$$|J_i^k| = \sqrt{\sum_{j=1}^N a_i^k a_j^k c_{ij}^k J_{ij}^2}. \tag{39}$$

Averaging over all neurons, patterns, and contexts, we obtain an average stability parameter of accessible memories when setting $k = \sigma$ in Eq. 38,

$$\kappa_{\text{acc}} = \frac{\sqrt{\alpha_{\text{H}}^*}}{spN} \sum_{i=1}^N \sum_{\sigma=1}^s \sum_{\mu=1}^p \kappa_i^{\sigma,\sigma^\mu}, \tag{40}$$

where we include the factor $\sqrt{\alpha_{\text{H}}^*}$ to make the measure relative to the stability of the standard Hopfield network (with mean stability $1/\sqrt{\alpha_{\text{H}}^*}$)[34]. We can then also measure the stability of inaccessible memory patterns by setting $k \neq \sigma$ in Eq. 38. As above, averaging over all neurons, patterns, and contexts we obtain an average stability parameter of inaccessible memories,

$$\kappa_{\text{inacc}} = \frac{2\sqrt{\alpha_{\text{H}}^*}}{s(s-1)pN} \sum_{i=1}^N \sum_{k=1}^{s-1} \sum_{\sigma=k+1}^s \sum_{\mu=1}^p \kappa_i^{k,\sigma^\mu}. \tag{41}$$

We measured the stability of both accessible and inaccessible memories at full capacity ($N = 10{,}000$, see Appendix G). For random neuronal and synaptic gating, as well as synaptic refinement, we observe that the stability of accessible memories is close to the stability levels of the standard Hopfield network and independent of the number of contexts, s, as well as neuronal and synaptic gating ratios, $1 - a$ and $1 - c$ (Fig. 9(b), open circles). As for inaccessible memory patterns, we find that their average stability is always equal to or lower than that of accessible memory patterns (Fig. 9(b), filled circles). The average stability of inaccessible memory patterns depends upon gating ratios, but remains relatively independent of the number of contexts, s (not shown). For neuronal gating, the stability of inaccessible memories increases roughly with a^2 (Fig. 9(b), orange line), whereas for synaptic gating, the stability of inaccessible memories increases with c (Fig. 9(b), purple line). For synaptic refinement with random neuronal gating, the stability of inaccessible memories decreases as the number of contexts, s, increases (not shown), because of the dependency of the synaptic gating ratio on the number of contexts. In the limit as the gating ratio increases to 50% (i.e., for $s > 100$), the stability of inaccessible memories remains very low for all values of a, and is well approximated by $a/8$ (Fig. 9(b), green line). Overall, we conclude that the modular organization of memories successfully suppresses inaccessible memories while maintaining high stability of accessible memories, as intended.

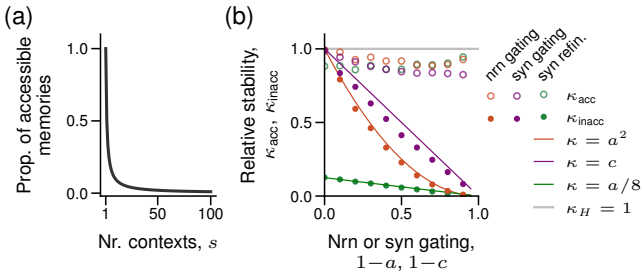

FIG. 9. Memory accessibility. (a) Proportion of accessible memories, i.e., ratio of memories stored per context divided by the total number of memories, $1/s$, as a function of the number of contexts, s. (b) Average memory stability for accessible memories, κ_{acc} (Eq. 40, open circles), and inaccessible memories, κ_{inacc} (Eq. 41, closed circles), with neuronal gating alone (red circles), synaptic gating alone (purple circles), or neuronal gating with synaptic refinement (green circles). Three comparison functions (full lines) are described in the legend. Stabilities are plotted relative to that of the standard Hopfield network ($\kappa_{\text{H}} = 1$).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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[20,24]. Such a system adapts to current behavioral states, i.e., contexts, and is thus able to modify the understanding of past experiences when needed[35]. Here, we revisited classic theoretical work on auto-associative memory in light of this perspective. Specifically, we studied a 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 peri-somatic inhibition[36] and modulation of intrinsic excitability[37]. 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. Given the precise and complex nature of synaptic gating, it is likely that full individual synaptic control is not plausible (but see ref. 38). However, there is considerable evidence for control of synaptic inputs at the level of dendritic branches[39], supported by findings such as dendritic clustering[40], localized dendritic inhibition[41], and localized neuromodulation[42]. One area for future work would be to relax the precision of synaptic gating, and explore a continuum between full individual control and coarse control of all inputs (as in neuronal gating).

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 pre-allocates 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 re-instated during recall — such contextual reinstatement is hypothesized to occur in the brain[43,44]. We observed that neuronal gating was much more effective than synaptic gating for random allocation, resulting in high memory capacity (Fig. 2) and low stability of inaccessible memories (Fig. 9). 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[26,37]. Specifically, the activity levels observed in such studies ($\sim 20-30\%$) correspond to optimal neuronal gating levels in our model for tens to hundreds of contexts (Fig. 3).

The second contextual gating method 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 (refinement of neuronal activity would result in altered memory patterns, so it was not considered). We found that synaptic refinement resulted in even larger gains in memory capacity (Fig. 7) 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, with each recall event possibly leading to changes[35]. Furthermore, recent work has also suggested that memory expression can be masked and controlled through inhibition[45,46]. Finally, the finding that synaptic refinement can stabilize arbitrary 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. 33,47,48 for related studies).

Memory capacity, stability, and information content results were obtained using conditional measures — i.e., we assumed that context is externally controlled, and memory stability is measured conditioned upon this fixed 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[20,24]. Furthermore, we contend that the questions addressed here do 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[2]. We therefore stress that our aim is 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[25], as well as descriptions of pathological memory states like in amnesia[49]. Overall, there may be a trade-off between capacity and flexibility, and it may be in the interest of the brain to favor flexibility over capacity. 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[50–52].

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[53]. 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 prin-

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ciples which are common in previous associative memory models, including sparsity, modularity, and hierarchical organization[2,10]. 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], or non-tree-like organization of semantic memories[54]. Our model bears a resemblance to these previous models, but is unique in considering context as the main determinant of memory groupings (but see ref. 55), as well as modeling context as an external signal imposed on the memory network (in contrast to previous hierarchical models). Extensions to associative memory networks that improve memory capacity, such as low activity patterns[6], alternative learning rules[19], or complex synapses[56] may be incorporated into the context-modular memory network architecture for additional benefits. Finally, other models have considered contextual control of learning and memory[57–60], as well as similar gating mechanisms applied in other situations (e.g., cortical information routing[61–63], deep learning[64], inhibitory learning[65], and motor learning[66]), highlighting the importance of contextual flexibility for cognition.

In conclusion, the integration of associative memory models with retrieval processes and decision making is a promising area of future research[67]. 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.

Acknowledgements. 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 (WT100000; WFP, EJA and TPV), a Wellcome Trust Senior Research Fellowship (214316/Z/18/Z; EJA and TPV), and a Research Project Grant by the Leverhulme Trust (RPG-2016-446; EJA). For the purpose of open access, the authors have applied a CC BY public copyright license to any authors accepted manuscript version arising from this submission.

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[2], 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 $\eta_i^{\sigma^\mu} = \frac{1}{2}(\xi_i^{\sigma^\mu} + 1)$. From these definitions, and considering a neuron i in the active context k, we arrive at

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

and thus the input to neuron i can equivalently be given in terms of V variables. We can thus write

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

where $H(\cdot)$ is the Heaviside function, such that $2H(x) - 1 = \mathrm{sgn}(x)$, 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

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

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

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

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 $V_i = \sigma(h_i) = \mathrm{e}^{2h_i}/(\mathrm{e}^{2h_i} + 1)$ and $S_i = \tanh(h_i)$, due to the identity $\tanh(x) = 2\sigma(x) - 1$.

Appendix B: Wald's equation

For the signal-to-noise analysis (Section 2.1), we make use of a result in statistics known as Wald's equation[68], which we summarize here. Consider a binomial random variable $Z \sim \mathcal{B}(N_Z, p_Z)$. Let Y be the sum of a sequence of independent identically distributed

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

random variables X_i of length Z, i.e., $Y = \sum_{i=1}^{Z} X_i$. Assuming that each X_i is independent of Z, then the mean and variance of Y can then be written as

$$\mathbb{E}[Y] = \mathbb{E}[Z]\mathbb{E}[X] \tag{B1}$$

and

$$\mathrm{Var}[Y] = \mathbb{E}[Z]\mathrm{Var}[X] + (\mathbb{E}[X])^2\mathrm{Var}[Z], \tag{B2}$$

where we have dropped the index on X because each random variable X_i comes from an identical distribution.

Appendix C: The standard, diluted, and low-activity Hopfield models

We compare the results obtained in this study with that of the standard Hopfield model[1]. For a network of N neurons storing p patterns, we use $\alpha_{\mathrm{H}} = p/N$ to denote the load (with $\alpha_{\mathrm{H}}^* = p_{\max}/N \approx 0.138$ at capacity) and m to denote the overlap of the state configuration with the retrieved pattern ($m^* \approx 0.97$ at capacity)[18,30]. For convenience of the reader, we write the three mean field equations for the standard Hopfield network here, as

$$m = \langle\!\langle \tanh\left(\beta(m + \sigma_{\mathrm{H}}\eta)\right)\rangle\!\rangle_\eta \tag{C1}$$

$$q = \langle\!\langle \tanh^2\left(\beta(m + \sigma_{\mathrm{H}}\eta)\right)\rangle\!\rangle_\eta \tag{C2}$$

$$r = \left\langle (m_\nu)^2 \right\rangle = \frac{q}{[1-\beta(1-q)]^2} \tag{C3}$$

where $\sigma_{\mathrm{H}}^2 = \alpha_{\mathrm{H}} r$ is the crosstalk noise and $\langle\!\langle\cdot\rangle\!\rangle_\eta$ represents a Gaussian average, i.e., $\eta \sim \mathcal{N}(0,1)$.

Next, we utilize previous results for a diluted Hopfield network[27] with dilution parameter c (analogous to the synaptic gating parameter in this work). The diluted connectivity was shown to be well-approximated by Gaussian noise[27],

$$J_{ij} = c_{ij}\frac{1}{cN}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu \approx \frac{1}{cN}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu + \delta_{ij}, \tag{C4}$$

with

$$\delta_{ij} \sim \mathcal{N}\left(0, \frac{\alpha(1-c)}{cN}\right). \tag{C5}$$

This weight matrix results in analogous mean field equations as the Hopfield network (Eqs. C1–C3), but with a noise term

$$\sigma_{\mathrm{D}}^2 = \alpha_{\mathrm{D}} r + \frac{\alpha_{\mathrm{D}}(1-c)}{c} q \tag{C6}$$

in place of σ_{H}. This results in a capacity that scales roughly linearly with c ($\alpha_{\mathrm{D}}^*(c) \approx c\alpha_{\mathrm{H}}^*$), and an overlap that also degrades as c becomes small (Fig. 4, purple line). Using the mean-field results, it can also be shown that a signal-to-noise analysis of a diluted Hopfield network results in an input (to neuron i for a pattern ν) of

$$h_i^\nu \approx \xi_i^\nu + \mathcal{N}(0, \alpha_{\mathrm{D}}(c)/c) \tag{C7}$$

which we utilize in Section 2.1.

Finally, we consider an associative memory variant with low-activity patterns[6] (Fig. 3(b) and Section 2.3). The capacity of the low-activity Hopfield network, α_{L}^*, scales with the activity level as

$$\alpha_{\mathrm{L}}^*(\tilde{a}) = \frac{\alpha_{\mathrm{H}}^*}{2\tilde{a}|\ln\tilde{a}|}, \tag{C8}$$

for $\tilde{a} \ll 1$[6] (the α_{H}^* factor is included to achieve a more fair comparison for more intermediate $\tilde{a}$ values).

Appendix D: Mean-field derivations

We used a heuristic mean-field approach[2,69,70] to compute the capacity in Section 2.2, which we detail here. Following ref. 27 (Eq. C4), we simplify the weight matrix used in the random neuronal and synaptic gating scheme (Eq. 5) using an effective matrix, $\tilde{J}_{ij}^k$, so that dilution is not explicitly included,

$$\tilde{J}_{ij}^k = \frac{1}{aN}\sum_{\mu=1}^{p} a_i^k\xi_i^{k\mu}a_j^k\xi_j^{k\mu} + \frac{1}{aN}\sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s}\sum_{\mu=1}^{cp} a_i^\sigma\xi_i^{\sigma\mu}a_j^\sigma\xi_j^{\sigma\mu} + \delta_{ij}, \tag{D1}$$

where δ_{ij} represents the effects of dilution within context k,

$$\delta_{ij} \sim \mathcal{N}\left(0, \frac{\tilde{\alpha}(1-c)}{acN}\right), \tag{D2}$$

and $\tilde{\alpha}$ is the effective noise in the weights due to all stored patterns, which we leave unspecified for now. Synaptic gating reduces

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the original sum over $(s-1)p$ terms to an average of $c(s-1)p$ – the second term of the right-hand side of Eq. D1 – which we take into account by changing the limit of the sum over patterns to cp.

We describe the dynamics of the system using three pairs of order parameters. First, we use the two load parameters as defined in Eqs. 3 and 4 (single-context and total load). Second, considering that the currently active context is k, we use m_ν^k and $n_{\sigma^\mu}^k$ to denote the *overlap* between the average network state $\langle S_i^{(k)} \rangle$ and a particular pattern ν of context k, $\xi_i^{k^\nu}$, and a pattern μ of context σ, $\xi_i^{\sigma^\mu}$, respectively,

$$m_\nu^k = \frac{1}{aN} \sum_{i=1}^{N} a_i^k \xi_i^{k^\nu} \left\langle S_i^{(k)} \right\rangle \tag{D3}$$

and

$$n_{\sigma^\mu}^k = \frac{1}{aN} \sum_{i=1}^{N} a_i^k a_i^\sigma \xi_i^{\sigma^\mu} \left\langle S_i^{(k)} \right\rangle . \tag{D4}$$

Third, we use r^k and r_n^k to denote the *mean square overlap* of the system configuration with the nonretrieved patterns in context k, and all other contexts, respectively,

$$r^k = \frac{1}{\alpha_{\text{cxt}}} \sum_{\substack{\mu=1,\\ \mu\neq\nu}}^{s} \left(m_\mu^k \right)^2 \tag{D5}$$

and

$$r_n^k = \frac{1}{\alpha} \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} \sum_{\mu=1}^{p} \left(n_{\sigma^\mu}^k \right)^2 . \tag{D6}$$

We consider that the network state is close to pattern 1 of context k. This means that if this pattern is a fixed-point of the network's dynamics when context k is active, $m_1^k \gg m_\mu^k, \mu \neq 1$.

We then follow analogous steps outlined in refs. 2,69,70 by plugging the mean field equations for $\langle S_i^{(k)} \rangle$ from Eq. 20 into Eq. D3, and expanding the weight matrix using the expression in Eq. D1,

$$m_\nu^k = \frac{1}{aN} \sum_{i=1}^{N} a_i^k \xi_i^{k^\nu} \tanh\left(\frac{\beta}{aN} \sum_{j=1}^{N} \sum_{\mu=1}^{p} a_i^k \xi_i^{k^\mu} a_j^k \xi_j^{k^\mu} \left\langle S_j^{(k)} \right\rangle + \frac{\beta}{aN} \sum_{j=1}^{N} \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} \sum_{\mu=1}^{cp} a_i^\sigma \xi_i^{\sigma^\mu} a_j^k a_j^\sigma \xi_j^{\sigma^\mu} \left\langle S_j^{(k)} \right\rangle + \beta \sum_{j=1}^{N} \delta_{ij} \left\langle S_j^{(k)} \right\rangle \right). \tag{D7}$$

We then rearrange terms and substitute the other m_μ^k and $n_{\sigma^\mu}^k$ terms into the equation to obtain

$$m_\nu^k = \frac{1}{aN} \sum_{i=1}^{N} a_i^k \xi_i^{k^\nu} \tanh\left(\beta \sum_{\mu=1}^{p} a_i^k \xi_i^{k^\mu} m_\mu^k + \beta \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} \sum_{\mu=1}^{cp} a_i^\sigma \xi_i^{\sigma^\mu} n_{\sigma^\mu}^k + \beta \sum_{j=1}^{N} \delta_{ij} \left\langle S_j^{(k)} \right\rangle \right), \tag{D8}$$

and then separating terms and multiplying by $\xi_i^{k^1} \xi_i^{k^1} = 1$, we get

$$m_\nu^k = \frac{1}{aN} \sum_{i=1}^{N} a_i^k \xi_i^{k^\nu} \xi_i^{k^1} \tanh\left(\beta \left(a_i^k m_1^k + \zeta_i^{k^1} \right) \right), \tag{D9}$$

where we use $\zeta_i^{k^1}$ to denote the crosstalk terms,

$$\zeta_i^{k^1} = \sum_{\mu=2}^{p} a_i^k \xi_i^{k^\mu} \xi_i^{k^1} m_\mu^k + \sum_{\substack{\sigma=1,\\ \sigma\neq k}}^{s} \sum_{\mu=1}^{cp} a_i^\sigma \xi_i^{\sigma^\mu} \xi_i^{k^1} n_{\sigma^\mu}^k + \sum_{j=1}^{N} \xi_i^{k^1} \delta_{ij} \left\langle S_j^{(k)} \right\rangle . \tag{D10}$$

We then repeat these steps twice more (plugging Eqs. D1 and 20 into Eq. D3) — once using m_1^k (the retrieved pattern) in place of an arbitrary m_ν^k, and another for $n_{\sigma^\mu}^k$ (using Eq. D4 in place of Eq. D3), to obtain

$$m_1^k = \frac{1}{N_{\text{cxt}}} \sum_i \tanh\left(\beta (m_1^k + \zeta_i^{k^1}) \right) \tag{D11}$$

and

$$n_{\sigma^\nu}^k = \frac{1}{aN} \sum_{i=1}^{N} a_i^k a_i^\sigma \xi_i^{k^\nu} \xi_i^{\sigma^\mu} \tanh\left(\beta (a_i^k m_1^k + \zeta_i^{k^1}) \right). \tag{D12}$$

These expressions are analogous to that of the standard Hopfield network and we refer the reader to refs. 2,69,70 for the details involved in manipulating these expressions to obtain the final mean field equations. One final detail unique to the context-modular memory network is estimating the variance of the crosstalk noise in Eq. D10. Due to the similarity with the diluted Hopfield

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

network (Appendix C), we can write the standard deviation of the term as

$$\sigma^2 = \tilde{\alpha} r + \frac{\tilde{\alpha}(1-c)}{c} q, \tag{D13}$$

with $\tilde{\alpha}$ being the effective load of the network due to the first two terms in Eq. D10. Note that the second term comes from the definition of δ_{ij} above (Eq. D2). Following similar arguments to the signal to noise analysis, we expect the effective load to depend on a particular realization of the number of overlapping contexts,

$$\tilde{\alpha} = \alpha_{cxt}(1+z), \tag{D14}$$

where z is a binomial random variable distributed as $\mathcal{B}(s-1, a^2c)$ (see Eq. 13) and with expected value $(s-1)a^2c$ (see Eq. 17).

Finally, from Eqs. D9, D11 and D12, we obtain the three mean-field equations

$$m = \langle\!\langle \tanh\left(\beta\left(m + \sigma z\right)\right) \rangle\!\rangle_\eta \tag{D15}$$

$$r = \left\langle \left(m_\nu\right)^2 \right\rangle = \frac{q}{[1-\beta(1-q)]^2} \tag{D16}$$

$$r_n = \left\langle \left(n_{\phi^\nu}\right)^2 \right\rangle = \frac{aq}{[1-\beta(1-q)]^2} = ar, \tag{D17}$$

where q is defined as

$$q = \langle\!\langle \tanh^2\left(\beta\left(m + \sigma z\right)\right) \rangle\!\rangle_\eta, \tag{D18}$$

and $\langle\!\langle \cdot \rangle\!\rangle_\eta$ represents a Gaussian average ($\eta \sim \mathcal{N}(0,1)$), and σ is defined in Eq. D13. Similar to the signal to noise analysis, the binomial random variable z can be dealt with by performing a binomial average at the end, or by approximating it with the expected value.

These equations can be solved numerically using standard methods. We again take a shortcut here and note the correspondence to the standard and diluted Hopfield networks, enabling us to avoid numerically solving the mean field equations. Following Section 2.1, we can set the variance term here equal to that of previous results, and obtain Eqs. 15 and 16. By simplifying the binomial random variable z to the expected value, we obtain Eqs. 17 and 18. Finally, we can write the mean field overlap with pattern 1 of context k as $m(c,a,s) = m_{\mathrm{D}}(c)$ as written in the main text.

Appendix E: Information content

Information content can be estimated based on the entropy (Shannon information) across all retrieved patterns[19]. Importantly, we consider that the total magnetization is always zero in our calculations — i.e., for both standard and diluted Hopfield networks, we consider that half of the neurons assume value +1, while half assume value −1. We can write the entropy per neuron considering the original patterns as

$$S_{½} = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = 1. \tag{E1}$$

The entropy per neuron considering the missing information and average overlap, m, is given by

$$\begin{aligned} S(m) &= -\frac{1+m}{2}\log_2\left(\frac{1+m}{2}\right) - \frac{1-m}{2}\log_2\left(\frac{1-m}{2}\right) \\ &= -\frac{1+m}{2}\log_2\left(1+m\right) - \frac{1-m}{2}\log_2\left(1-m\right) + 1. \end{aligned} \tag{E2}$$

Combining Eqs. E1 and E2 we arrive at the total information content for the diluted Hopfield network as

$$I_{\mathrm{D}}(c) = pN\left[S_{½} - S(m_{\mathrm{D}}(c))\right] \tag{E3}$$

$$= \alpha_{\mathrm{D}}(c) N^2 \left[\frac{1+m_{\mathrm{D}}(c)}{2}\log_2\left(1+m_{\mathrm{D}}(c)\right) + \frac{1-m_{\mathrm{D}}(c)}{2}\log_2\left(1-m_{\mathrm{D}}(c)\right)\right], \tag{E4}$$

with $I_{\mathrm{H}} = I_{\mathrm{D}}(c=1)$ being the information content of the standard Hopfield network.

Using the same argument, the maximum information content of the low-activity Hopfield network, when the activity level goes to zero, can be written as

$$I_{\mathrm{L}}(\tilde{a}) = pN\left[S_{\tilde{a}} - S(m_{\mathrm{L}}(\tilde{a}))\right] \tag{E5}$$

$$= \alpha_{\mathrm{L}}(\tilde{a}) N^2 \left[\frac{1+m_{\mathrm{L}}(\tilde{a})}{2}\log_2\left(1+m_{\mathrm{L}}(\tilde{a})\right) + \frac{1-m_{\mathrm{L}}(\tilde{a})}{2}\log_2\left(1-m_{\mathrm{L}}(\tilde{a})\right)\right], \tag{E6}$$

where $S_{\tilde{a}} = \tilde{a}\log_2\tilde{a} - (1-\tilde{a})\log_2(1-\tilde{a})$ is the entropy per neuron for an activity level $\tilde{a}$ and $m_{\mathrm{L}}(\tilde{a})$ is the average overlap between original patterns and minima of the network's dynamics for activity level $\tilde{a}$. For the limit of vanishing activity level, we get[19]

$$I_{\mathrm{L}}^* = \frac{N^2}{2\ln 2} \approx 0.72N^2 \approx 6I_{\mathrm{H}}^*. \tag{E7}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

For the context-modular memory network, we consider the conditional information content of patterns stored in networks of size N_{cxt}. This way we can rewrite Eq. E4 substituting $\alpha_{\text{D}}(c)N^2$ by $\alpha a N^2$, arriving in Eq. 29.

Appendix F: Connectivity density for synaptic refinement

The expected synaptic gating ratio, $1 - c(s,a)$, for synaptic refinement with random neuronal gating (with parameter a), can be estimated assuming that the weight distributions for a single context and across all contexts are Gaussian with zero mean and variances proportional to the amount of crosstalk that they contribute. Given this, the probability that a particular weight will be removed can be approximated as

$$1 - c(s,a) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x^2}{2\sigma_1^2}\right) \int_x^\infty \frac{1}{\sqrt{2\pi\sigma_{\text{A}}^2}} \exp\left(-\frac{y^2}{2\sigma_{\text{A}}^2}\right) dy dx \tag{F1}$$

$$= \frac{1}{\pi} \arctan(\sigma_{\text{A}}/\sigma_1), \tag{F2}$$

where σ_1 and σ_{A} are the crosstalk standard deviations from patterns within the same context and all other contexts, respectively, whose ratio is

$$\frac{\sigma_{\text{A}}}{\sigma_1} = \sqrt{\frac{z\alpha_{\text{cxt}}/N}{\alpha_{\text{cxt}}/N}} = \sqrt{z}, \tag{F3}$$

and z being the binomial random variable distributed as $\mathcal{B}(s-1, a^2 c)$. Doing a binomial average, we obtain the same equation in the main text,

$$1 - c(s,a) = \frac{1}{\pi} \left\langle \arctan\left(\sqrt{z}\right) \right\rangle_{z \sim \mathcal{B}(s-1, a^2)}. \tag{F4}$$

Appendix G: Details of numerical simulations

We briefly describe the details of all numerical simulations here, used to generate the overlap and capacity plots in Figs. 4, 5, 7, and 8. Code was written in python with the help of Pytorch[71]. Memory capacity was estimated numerically by building finite-sized networks (N = 10,000) 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 (60). 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 (anti-correlated) of pattern ξ^{σ^μ}. We repeated this procedure for all patterns in a particular context to obtain an average overlap $\overline{m^\sigma} = N_{\text{cxt}}^{-1} \sum_{\mu=1}^{p} m^{\sigma^\mu}$. Capacity was set as the maximum number of patterns stored for which the average overlap was greater than or equal to the overlap curves in Fig. 4 (as a function of gating ratios $1 - a$ and $1 - c$). Synchronous dynamics were chosen for efficiency reasons despite potential convergence issues. To generate the overlap plots in Fig. 4, we simulated networks with the theoretically-expected number of patterns at capacity (using dynamics as specified above), and measured the resulting average overlap. All code used to run simulations and generate the paper figures will be made available upon publication of this work.

References

[1] J J Hopfield, "Neural networks and physical systems with emergent collective computational abilities," Proceedings of the National Academy of Sciences **79**, 2554–2558 (1982).
[2] J Hertz, A Krogh, and R G Palmer, *Introduction to the theory of neural computation* (Addison-Wesley/Addison Wesley Longman, 1991).
[3] D J Amit, *Modeling brain function: The world of attractor neural networks* (Cambridge university press, 1992).
[4] R Chaudhuri and I Fiete, "Computational principles of memory," Nature Neuroscience **19**, 394–403 (2016).
[5] C Miehl, S Onasch, D Festa, and J Gjorgjieva, "Formation and computational implications of assemblies in neural circuits," The Journal of Physiology **In Press** (2022).
[6] M V Tsodyks and M V Feigelman, "The enhanced storage capacity in neural networks with low activity level," Europhysics Letters , 101–105 (1988).
[7] H Gutfreund, "Neural networks with hierarchically correlated patterns," Physical Review A **37**, 570–577 (1988).
[8] M V Tsodyks, "Hierarchical associative memory in neural networks with low activity level," Modern Physics Letters B **4**, 259–265 (1990).
[9] J Steinberg and H Sompolinsky, "Associative memory of structured knowledge," Scientific Reports **12**, 21808 (2022).
[10] A Treves and E T Rolls, "What determines the capacity of autoassociative memories in the brain?" Network: Computation in Neural Systems **2**, 371–397 (1991).
[11] B Derrida, E Gardner, and A Zippelius, "An exactly solvable asymmetric neural network model," Europhysics Letters **4**, 167 (1987).
[12] C F Mari and A Treves, "Modeling neocortical areas with a modular neural network," Biosystems **48**, 47–55 (1998).
[13] A Renart, N Parga, and E T Rolls, "Associative memory properties of multiple cortical modules," Network: Computation in Neural Systems **10**, 237–255 (1999).
[14] A M Dubreuil and N Brunel, "Storing structured sparse memories in a multi-modular cortical network model," Journal of Computational Neuroscience **40**, 157–175 (2016).
[15] M Naim, V Boboeva, C J Kang, and A Treves, "Reducing a cortical network to a potts model yields storage capacity estimates," Journal of Statistical Mechanics: Theory and Experiment **2018**, 043304 (2018).
[16] V Braitenberg and A Schüz, *Anatomy of the cortex: statistics and geometry* (Springer-Verlag Berlin Heidelberg, 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," Proceedings of the National Academy of Sciences **106**, 3555–3560 (2009).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.08.898528; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[18] D J Amit, H Gutfreund, and H Sompolinsky, "Storing infinite numbers of patterns in a spin-glass model of neural networks," Physical Review Letters **55**, 1530–1533 (1985).
[19] E Gardner, "The space of interactions in neural network models," Journal of Physics A: Mathematical and General **21**, 257–270 (1988).
[20] E Tulving and Z Pearlstone, "Availability versus accessibility of information in memory for words," Journal of Verbal Learning and Verbal Behavior **5**, 381–391 (1966).
[21] S M Smith, "Remembering in and out of context," Journal of Experimental Psychology: Human Learning and Memory **5**, 460–471 (1979).
[22] M E Bouton, "Context, time, and memory retrieval in the interference paradigms of pavlovian learning," Psychological Bulletin **114**, 80–99 (1993).
[23] S Maren, K L Phan, and I Liberzon, "The contextual brain: implications for fear conditioning, extinction and psychopathology," Nature Reviews Neuroscience **14**, 417–428 (2013).
[24] P W Frankland, S A Josselyn, and S Köhler, "The neurobiological foundation of memory retrieval," Nature Neuroscience **22**, 1576–1585 (2019).
[25] M C Anderson and J C Hulbert, "Active forgetting: Adaptation of memory by prefrontal control," Annual Review of Psychology **72**, 1–36 (2021).
[26] S A Josselyn and P W Frankland, "Memory allocation: mechanisms and function," Annual Review of Neuroscience **41**, 389–413 (2018).
[27] H Sompolinsky, "Neural networks with nonlinear synapses and a static noise," Physical Review A **34**, 2571 (1986).
[28] 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[2]. 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.
[29] S F Edwards and P W Anderson, "Theory of spin glasses," Journal of Physics F: Metal Physics **5**, 965 (1975).
[30] D J Amit, H Gutfreund, and H Sompolinsky, "Statistical mechanics of neural networks near saturation," Annals of Physics **173**, 30–67 (1987).
[31] D J Amit, H Gutfreund, and H Sompolinsky, "Information storage in neural networks with low levels of activity," Physical Review A **35**, 2293–2303 (1987).
[32] J L van Hemmen and R Kühn, "Nonlinear neural networks," Physical Review Letters **57**, 913–916 (1986).
[33] S Fusi and L F Abbott, "Limits on the memory storage capacity of bounded synapses," Nature Neuroscience **10**, 485–493 (2007).
[34] L F Abbott, "Learning in neural network memories," Network: Computation in Neural Systems **1**, 105–122 (1990).
[35] Y Dudai, "The restless engram: consolidations never end," Annual Review of Neuroscience **35**, 227–247 (2012).
[36] T F Freund and I Katona, "Perisomatic inhibition," Neuron **56**, 33–42 (2007).
[37] 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**, 722–735 (2014).
[38] 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**, 566–578 (2020).
[39] T Branco and M Häusser, "The single dendritic branch as a fundamental functional unit in the nervous system," Current Opinion in Neurobiology **20**, 494–502 (2010).
[40] G Kastellakis, D J Cai, S C Mednick, A J Silva, and P Poirazi, "Synaptic clustering within dendrites: an emerging theory of memory formation," Progress in Neurobiology **126**, 19–35 (2015).
[41] 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**, 759–762 (2013).
[42] 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**, 611–625 (2018).
[43] 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**, 347–354 (2014).
[44] 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**, 432–441 (2014).
[45] 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**, 191–203 (2016).
[46] H C Barron, T P Vogels, T E J Behrens, and M Ramaswami, "Inhibitory engrams in perception and memory," Proceedings of the National Academy of Sciences **114**, 6666–6674 (2017).
[47] G Mongillo, S Rumpel, and Y Loewenstein, "Intrinsic volatility of synaptic connections – a challenge to the synaptic trace theory of memory," Current Opinion in Neurobiology **46**, 7–13 (2017).
[48] L Susman, N Brenner, and O Barak, "Stable memory with unstable synapses," Nature Communications **10**, 4441 (2019).
[49] O Hardt, S-H Wang, and K Nader, "Storage or retrieval deficit: the yin and yang of amnesia," Learning & Memory **16**, 224–230 (2009).
[50] S Romani, I Pinkoviezky, A Rubin, and M Tsodyks, "Scaling laws of associative memory retrieval," Neural Computation **25**, 2523–2544 (2013).
[51] M Naim, M Katkov, S Romani, and Misha Tsodyks, "Fundamental law of memory recall," Physical Review Letters **124**, 018101 (2020).
[52] O H Schnaack, L Peliti, and A Nourmohammad, "Learning and organization of memory for evolving patterns," Physical Review X **12**, 021063 (2022).
[53] D Breuer, M Timme, and R-M Memmesheimer, "Statistical physics of neural systems with nonadditive dendritic coupling," Physical Review X **4**, 011053 (2014).
[54] V Boboeva, R Brasselet, and A Treves, "The capacity for correlated semantic memories in the cortex," Entropy **20**, 824 (2018).
[55] S Doboli, A A Minai, and P J Best, "Latent attractors: a model for context-dependent place representations in the hippocampus," Neural Computation **12**, 1009–1043 (2000).
[56] M K Benna and S Fusi, "Computational principles of synaptic memory consolidation," Nature Neuroscience **19**, 1697–170 (2016).
[57] M W Howard and M J Kahana, "A distributed representation of temporal context," Journal of Mathematical Psychology **46**, 269–299 (2002).
[58] S M Polyn, K A Norman, and M J Kahana, "A context maintenance and retrieval model of organizational processes in free recall," Psychological Review **116**, 129 (2009).
[59] S J Gershman, M-H Monfils, K A Norman, and Y Niv, "The computational nature of memory modification," eLife **6**, e23763 (2017).
[60] E J Agnes and T P Vogels, "Codependent excitatory and inhibitory plasticity accounts for quick, stable and long-lasting memories in biological networks," BioRxiv **2021.04.01**, 437962 (2022).
[61] T P Vogels and L F Abbott, "Gating multiple signals through detailed balance of excitation and inhibition in spiking networks," Nature Neuroscience **12**, 483 (2009).
[62] G R Yang, J D Murray, and X-J Wang, "A dendritic disinhibitory circuit mechanism for pathway-specific gating," Nature Communications **7**, 12815 (2016).
[63] G Christodoulou, T P Vogels, and E J Agnes, "Regimes and mechanisms of transient amplification in abstract and biological neural networks," PLoS Computational Biology **18**, e1010365 (2022).
[64] N Y Masse, G D Grant, and D J Freedman, "Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization," Proceedings of the National Academy of Sciences **115**, E10467–E10475 (2018).
[65] E J Agnes, A I Luppi, and T P Vogels, "Complementary inhibitory weight profiles emerge from plasticity and allow flexible switching of receptive fields," Journal of Neuroscience **40**, 9634–9649 (2020).
[66] 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," Nature Neuroscience **21**, 1774–1783 (2018).
[67] M N Shadlen and D Shohamy, "Decision making and sequential sampling from memory," Neuron **90**, 927–939 (2016).
[68] S M Ross, *Introduction to probability models* (Academic press, 2014).
[69] P Peretto, "On learning rules and memory storage abilities of asymmetrical neural networks," Journal de Physique France **49**, 711–726 (1988).
[70] T Geszti, *Physical models of neural networks* (World Scientific, 1990).
[71] A Paszke, S Gross, F Massa, A Lerer, J Bradbury, G Chanan, T Killeen, Z Lin, N Gimelshein, L Antiga, A Desmaison, A Kopf, E Yang, Z DeVito, M Raison, A Tejani, S Chilamkurthy, B Steiner, L Fang, J Bai, and S Chintala, "Pytorch: An imperative style, high-performance deep learning library," Advances in Neural Information Processing Systems **32** (2019).