



OPEN ACCESS

EDITED BY

Adrian G. Palacios,
Universidad de Valparaiso, Chile

REVIEWED BY

Alessandro Treves,
International School for Advanced
Studies (SISSA), Italy
Jose Bargas,
National Autonomous University of
Mexico, Mexico

*CORRESPONDENCE

Emilio Kropff
ekropff@leloir.org.ar

RECEIVED 30 July 2022

ACCEPTED 22 August 2022

PUBLISHED 15 September 2022

CITATION

Emina F and Kropff E (2022) Selective
connectivity enhances storage
capacity in attractor models of
memory function.
Front. Syst. Neurosci. 16:983147.
doi: 10.3389/fnsys.2022.983147

COPYRIGHT

© 2022 Emina and Kropff. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Selective connectivity enhances storage capacity in attractor models of memory function

Facundo Emina^{1,2} and Emilio Kropff^{2*}

¹Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Física, Buenos Aires, Argentina, ²Leloir Institute—IIBBA/CONICET, Buenos Aires, Argentina

Autoassociative neural networks provide a simple model of how memories can be stored through Hebbian synaptic plasticity as retrievable patterns of neural activity. Although progress has been made along the last decades in understanding the biological implementation of autoassociative networks, their modest theoretical storage capacity has remained a major constraint. While most previous approaches utilize randomly connected networks, here we explore the possibility of optimizing network performance by selective connectivity between neurons, that could be implemented in the brain through creation and pruning of synaptic connections. We show through numerical simulations that a reconfiguration of the connectivity matrix can improve the storage capacity of autoassociative networks up to one order of magnitude compared to randomly connected networks, either by reducing the noise or by making it reinforce the signal. Our results indicate that the signal-reinforcement scenario is not only the best performing but also the most adequate for brain-like highly diluted connectivity. In this scenario, the optimized network tends to select synapses characterized by a high consensus across stored patterns. We also introduced an online algorithm in which the network modifies its connectivity while learning new patterns. We observed that, similarly to what happens in the human brain, creation of connections dominated in an initial stage, followed by a stage characterized by pruning, leading to an equilibrium state that was independent of the initial connectivity of the network. Our results suggest that selective connectivity could be a key component to make attractor networks in the brain viable in terms of storage capacity.

KEYWORDS

autoassociative memory, Hopfield network, structural plasticity, connectivity optimization, storage capacity, computational models, attractor dynamics

1. Introduction

The dynamics and functionality of neural networks, both artificial and biological, are strongly influenced by the configuration of synaptic weights and the architecture of connections. The ability of a network to modify synaptic weights plays a central role in learning and memory. Networks in cortical areas and in the hippocampus are believed to store patterns of activity through synaptic plasticity, making possible their retrieval at

a later stage, or equivalently the replay of past network states (Citri and Malenka, 2008; Carrillo-Reid et al., 2016). Autoassociative neural networks have emerged as crucial models to describe this mode of computation by means of attractor dynamics. Memories can be thought of as attractor states that arise as a consequence of configuring synaptic weights following a Hebbian learning rule, where the synapse between two neurons is strengthened or weakened depending on how correlated is their activity (Hopfield, 1982).

A second, less explored way in which interactions between neurons in autoassociative memories can be modified is by adding or deleting connections. Evidence suggests that topological characteristics of biological neural networks are far from random, and might result from a trade-off between energy consumption minimization and performance maximization (Bullmore and Sporns, 2012). Over the course of evolution, what is now the human brain grew by several orders of magnitude in terms of the number of neurons (N), but the number of connections per neuron (c) has remained rather stable (Assaf et al., 2020). A limitation for increasing c is the scaling of the volume and mass associated to it, and the fact that white matter (long-range connections) represents already approximately half of the total mass of the human brain might indicate a tight compromise (Herculano-Houzel et al., 2010). The number of connections, however, is not constant throughout human life. The formation of neuronal connections in the central nervous system is a highly dynamic process, consisting of simultaneous events of elimination and formation of connections (Hua and Smith, 2004). Studies suggest that a common rule in many parts of the brain consists of an initial stage where the creation of connections dominates, peaking during childhood at around 2 years old, and a later reversion stage where pruning dominates until an equilibrium is reached during late adolescence or early adulthood (Huttenlocher, 1979; Lichtman and Colman, 2000; Navlakha et al., 2015). In addition, evidence such as profound changes in dendritic branching in some cortical areas after exposing animals to complex environments and cortical axon remodeling after lesions of the sensory periphery support the idea that structural changes in the wiring diagram might be a vital complement of the learning scheme based on synaptic plasticity (Chklovskii et al., 2004).

Several theoretical variations of the Hopfield network have been proposed to describe the functionality of the hippocampus and other brain areas by means of attractor dynamics, with varying degree of biological plausibility (Amit, 1989; Treves and Rolls, 1991; Kropff and Treves, 2007; Roudi and Latham, 2007). All these variants, however, share a key limitation of the original Hopfield model regarding the number of memories that a network can store and successfully retrieve. Fully connected Hopfield models, for example, can only store a number of patterns equal to a fraction $\alpha_c \sim 0.138$ of the connections per neuron. If more memories are stored, a phase transition occurs and the network loses its ability to retrieve any of the

stored patterns. Real brains would need to stay far from the transition point to avoid this risk, a constraint under which networks in the human brain, which has $N \sim 10^{11}$ neurons but on average only $c \sim 10^4$ connections per neuron, would be able to store $< p \sim 10^3$ memories, a rather modest number. Error-correcting iterative algorithms as an alternative technique to set the synaptic weight configuration have been shown to increase the storage capacity up to $\alpha_c = 2$ (Forrest, 1988; Gardner, 1988). However, the use of fully connected networks and non-Hebbian modification of synaptic weights limits the applicability of these ideas to model real brains. Random sparse connectivity offers only a relative gain in network performance, allowing the storage capacity to go up to $\alpha_c = 2/\pi$ in ultra-diluted networks, where $c \ll \ln(N)$ (Derrida et al., 1987), or a slightly lower limit in less diluted networks where $c \ll N$ (Arenzon and Lemke, 1994). Some studies suggest that selective pruning might be much more effective than random dilution in terms of storage capacity, making it diverge for ultra-diluted networks (Montemurro and Tamarit, 2001) or increasing the memory robustness (Janowsky, 1989).

Autoassociative networks gain an order of magnitude in storage capacity when including a more realistic sparse activity (involving $\sim 5 - 10\%$ of neurons per pattern) (Tsodyks and Feigel'man, 1988; Tsodyks, 1989), although this gain is compensated by similar losses related to the incorporation of other biologically plausible elements in either the modeled neurons or the statistics of the stored data (Kropff and Treves, 2007; Roudi and Latham, 2007). Other models that consider the cortex as a network of networks suggest that a hierarchical strategy does not yield considerable benefits, since the number of bits of information that can be stored per synaptic variable is very similar to that of simpler models (Kropff and Treves, 2005).

What other strategies could have been developed by our brains to increase the storage capacity beyond the limit of hundreds of memories predicted by the Hopfield model? In this work we explore the possibility of introducing modifications to the architecture of connections with the aim of improving the signal-to-noise ratio. This process could be analogous to the formation and pruning of connections that reshapes our brains throughout maturation. In the first section we show through numerical simulations that autoassociative networks are able to increase their storage capacity up to around seven-fold by minimizing the noise. In the second section, we show that if the cost function aims to reinforce the signal rather than minimizing the noise, a gain of up to almost one order of magnitude can be obtained. In the last section we implement an algorithm where connections are constantly added and pruned as the network learns new patterns, showing that it converges to the same connectivity with optimal storage capacity regardless of the starting conditions, and that if initial conditions are of low connectivity it reaches an early maximum followed by a long period of decay, as is the case generally in the human cortex.

2. Materials and methods

2.1. Hopfield model

2.1.1. Autoassociative network

For simplicity, we utilized a network similar to the one originally proposed by Hopfield, capable of storing information in the form of activity patterns that can be later retrieved by a partial cue. The network consists of N recurrently connected neurons, each receiving c pre-synaptic connections with no self-connections. The state of neuron i at time t is represented by s_i^t and can take two possible values: $s_i^t = 1$ (“active”) or $s_i^t = -1$ (“inactive”). The activity of the network evolves synchronously at discrete time steps. At each time step neurons receive a local activity field given by the weighted sum of the activity of other neurons

$$h_i^t = \frac{1}{c} \sum_{j=1}^N W_{ij} C_{ij} s_j^t, \quad (1)$$

where W_{ij} represents the synaptic weight between the pre-synaptic neuron j and the post-synaptic neuron i , C_{ij} is a binary matrix taking a value of 1 if this physical connection exists and 0 otherwise and c is the number of pre-synaptic connections targeting a given neuron, so that $\sum_{j=1}^N C_{ij} = c$ for all i ($C_{ii} = 0$ because there are no self-connections). Since we are interested in studying the effects of adding and removing connections, it is convenient to consider separately a synaptic matrix \overline{W} with values that only depend on the stored patterns and a matrix \overline{C} that only indicates whether or not the connection exists.

After calculating h_i^t , neurons update their state following a deterministic update rule,

$$s_i^{t+1} = \text{sgn}(h_i^t), \quad (2)$$

where the function $\text{sgn}(x)$ returns the sign of x .

Synaptic weights are computed following a linear Hebbian rule,

$$W_{ij} = \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu, \quad (3)$$

where ξ_i^μ represents the state of neuron i in pattern μ , from a total amount of p stored patterns. In the classic Hopfield model (and in this work), activity patterns come from the random distribution,

$$P(\xi_i^\mu) = \frac{1}{2} \delta(\xi_i^\mu - 1) + \frac{1}{2} \delta(\xi_i^\mu + 1). \quad (4)$$

2.1.2. Storage capacity

The relevant parameter that describes the storage capacity of a network is $\alpha = p/c$. A critical value α_c exists where, if more

patterns are loaded, the network suffers a phase transition to an amnesic state. This phase transition can be understood in terms of the local activity field each neuron receives. If the network has p patterns ideally stored (i.e., every pattern is exactly a fix point of the network dynamics) and at time t the network is initialized so that the state of each neuron corresponds to a given pattern $\bar{\xi}^v$ (i.e., $s_i^t = \xi_i^v$ for all neurons), then no changes in the neuronal states should occur as a consequence of the update rule in Equation (2), i.e., $s_i^{t+1} = s_i^t = \xi_i^v$. To understand this, the local activity field that neuron i receives $h_i^t \equiv h_i^v$ can be split into a signal term ξ_i^v (resulting from the contributions of stored information corresponding to pattern v) and a noise term R_i^v (which can be thought of as the contribution of other stored patterns)

$$s_i^{t+1} = \text{sgn}(h_i^v),$$

$$h_i^v = \xi_i^v + \frac{1}{c} \sum_{j=1}^N \sum_{\mu \neq v}^p \xi_i^\mu \xi_j^\mu \xi_j^v C_{ij} = \xi_i^v + R_i^v. \quad (5)$$

If the local field has the same sign as ξ_i^v , or equivalently, if the aligned local field is positive for each neuron ($h_i^v \xi_i^v \equiv 1 + R_i^v \xi_i^v > 0$), then the pattern $\bar{\xi}^v$ will be exactly retrieved. In other words, pattern retrieval depends on whether or not the noise term flips the sign of the aligned local activity field from positive to negative. If connectivity is diluted and random, the aligned local field can be approximated by a random variable following a normal distribution of unitary mean and standard deviation $\sigma \sim \sqrt{p/c}$. Since this standard deviation increases monotonically with the number of stored patterns, there is a critical point beyond which the noise fluctuations are large enough to destabilize all patterns and prevent them from being recovered (i.e., $h_i^v \xi_i^v < 0$ for a critical number of neurons).

2.1.3. Basin of attraction

A fundamental characteristic of the attractor states of a Hopfield network is their basin of attraction. It is a quantification of the network's tolerance to errors in the initial state. The basin of attraction depends on the connectivity of the network and the number of patterns stored. In randomly connected networks with low memory load ($p \ll p_c$), every pattern can be retrieved if the cue provided to the network represents at least 50% of the pattern (<50% is a cue for the retrieval of a stable spurious state represented by flipping all elements in the pattern). As the memory load approaches the critical value p_c , tolerance to error smoothly weakens.

We studied memory robustness by simulating networks with different connectivity and memory load. For each initial error, we counted the number of patterns the network could successfully retrieve (e.g., initializing the network in pattern $\bar{\xi}^v$ with an error of 0.2 implied that 20% of the neurons initially deviated from the pattern). We studied the percentage of

successfully retrieved patterns as a function of error and memory load.

2.2. Simulations

Simulations were run in custom made scripts written in MATLAB (RRID:SCR_001622). In all sections, the connectivity matrix \bar{C} of each simulated network was initially constructed pseudo-randomly. We selected for each of the N neurons its c pre-synaptic connections using MATLAB's function `randperm(N - 1, c)`, randomly obtaining the c indexes among their $N - 1$ possible inputs.

In a typical simulation to study storage capacity, we initialized the network in a given pattern v , i.e., $s_i^0 = \xi_i^v$, and updated the network following Equation (2) for 100 iterations or until the overlap between the pattern and the network state (i.e., $\frac{1}{N} \sum_{j=1}^N s_j^t \xi_j^v$) remained constant. After this, the pattern was classified as successfully retrieved if the overlap between the final state of the network and the pattern v was >0.7 (implying that the attractor associated to the pattern was stable but possibly slightly distorted by interference with other patterns). Note that, as is usual when working close to the storage capacity limit, we did not require patterns to be exact fixed points of the network dynamics in order to be considered as successfully stored. Unless otherwise specified, the storage capacity of the network was defined as the maximum number of patterns for which the network could successfully retrieve all of them, normalized by c .

2.3. Connectivity optimization

2.3.1. Noise reduction

As described above, the retrieval of memories can be compromised by random fluctuations in the noise term making the aligned local field negative. We asked whether a non-random connectivity matrix could substantially reduce the local noise contribution each neuron received, resulting in an increase in the storage capacity. In order to find an optimal connectivity matrix, we proposed each neuron to select its c pre-synaptic connections by minimizing an energy function

$$E_i^0 = \sum_{v=1}^p \left(\sum_{j=1}^N \sum_{\mu \neq v}^p \xi_i^\mu \xi_j^\mu \xi_i^v \xi_j^v C_{ij} \right)^2 \propto \sum_{v=1}^p (R_i^v)^2. \quad (6)$$

Note that in an ideal connectivity configuration that cancels E_i^0 , neuron i would receive zero noise during the retrieval of any of the stored patterns. In order to use the number of pre-synaptic connections as a control parameter, we used a fixed c , implying that this minimization problem is subject to the constraint $\sum_{j=1}^N C_{ij} = c$. Other constraints inherent to the

nature of the connectivity matrix is that it is binary ($C_{ij} \in \{0, 1\}$) with $C_{ii} = 0$. Thus, the minimization of Equation (6) belongs to the family of quadratic constrained binary optimization problems. To obtain a computationally efficient approximate solution to this problem we implemented an adaptation of the simulated annealing algorithm. We applied independently to each neuron's pre-synaptic connectivity an annealing schedule where temperature T was decreased by a factor of 0.99 at each step. For each neuron i , we randomly selected two elements in the i -th row of the connectivity matrix \bar{C} and permuted them, which ensured an invariant c . The change was always accepted if the cost function E_i^0 decreased ($\Delta E_i^0 < 0$), or otherwise with a probability equal to $e^{-\Delta E_i^0/T}$. Initial temperature was estimated following the method detailed by Yang (2014) and final temperature was set to 10^{-4} .

2.3.2. Signal reinforcement

We proposed a second cost function which can be thought of as a generalization of Equation (6). In this case, the aim is not to minimize the noise but instead to reinforce the signal, contributing positively to the aligned local field by making $R_i^v \xi_i^v > 0$,

$$E_i^\epsilon = \sum_{v=1}^p \left(\sum_{j=1}^N \sum_{\mu \neq v}^p \xi_i^\mu \xi_j^\mu \xi_i^v \xi_j^v C_{ij} - \epsilon \right)^2 \propto \sum_{v=1}^p (R_i^v \xi_i^v - \epsilon)^2, \quad (7)$$

where ϵ is a non-negative parameter. Note that setting $\epsilon = 0$ makes this problem equivalent to the noise reduction scenario.

Results shown in Section 3.2 correspond to optimizations with $\epsilon = p$ (we refer to the energy function in Equation 7 as E_i^p). This arbitrary choice is justified by the observation in a preliminary analysis that the optimization with this value of ϵ nearly maximized the number of patterns where noise contributed positively to the local field for a wide range of values of c , p , and N .

2.3.3. Online algorithm

As mentioned in Section 1, addition and pruning of connections are features of brain maturation. To gain an insight into the role they could play in a learning network, we also proposed an online optimization algorithm where the incorporation of memories and the modification of the connectivity through signal reinforcement occurred in parallel. Our aims were to understand if a similar improvement in storage capacity could be achieved through this on-line approach and to study the dynamics of connectivity.

Algorithm 1 consisted of a heuristic optimization where each neuron would independently attempt to minimize E_i^ϵ (Equation 7) by eliminating and generating input connections. The main difference with the previous approach is that in order

not to disrupt the natural dynamics of the network, annealing was not used (equivalent to setting $T = 0$ throughout the optimization). In other words, the algorithm was deterministic in the sense that changes in connectivity were only accepted if they diminished the cost function. The resulting greedy optimization scheme, however, often failed to escape local minima. We found that the relaxation of the performance criterion (90% of patterns retrieved instead of 100%) gave enough flexibility for the algorithm to escape these local minima and actually achieve an overall performance similar to the one obtained with simulated annealing. Therefore, we defined the variable p_{eff} as the effective number of patterns the network could successfully retrieve, from a total amount of p , where $p_{eff} > 0.9p$.

```

1:  $\{\bar{C} \in \mathbb{R}^{N \times N} / \sum_{j=1}^N C_{ij} = c_0; C_{ii} = 0; 1 \leq i \leq N\}$   $\triangleright$  Generate
   random connectivity matrix
2:  $p = p_c^{rand}$ 
3:  $W_{ij} = \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu$   $\triangleright$  Store  $p = p_c^{rand}$  random patterns  $\xi_i^\mu$ 
4: Compute  $p_{eff}$ 
5: while  $p_{eff} > 0.9 \cdot p$  do
6:    $p = p + 10$ 
7:    $W_{ij} = \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu$ 
8:   Compute  $p_{eff}$ 
9:   while  $p_{eff} \leq 0.9 \cdot p$  do
10:    for  $i = 1:N$  do  $\triangleright$  This for can be done in
        parallel
11:      Compute  $E_i^c[\bar{C}_i]$   $\triangleright$  Neuron  $i$  local energy
12:      for  $k = 1:10$  do
13:         $\bar{C}_i^f = \bar{C}_i$   $\triangleright$  Dummy connectivity
14:        Select randomly  $C_{ij}^f$  such that  $j \neq i$ 
15:         $C_{ij}^f = 1 - C_{ij}^c$   $\triangleright$  Assess corresponding
        change
16:        Compute  $E_i^c[\bar{C}_i^f]$ 
17:        if  $E_i^c[\bar{C}_i^f] - E_i^c[\bar{C}_i] < 0$  then
18:           $E_i^c[\bar{C}_i] = E_i^c[\bar{C}_i^f]$ 
19:           $\bar{C}_i = \bar{C}_i^f$ 
20:        end if
21:      end for
22:    end for
23:
24:    if stopping condition is met then
25:      break while
26:    end if
27:
28:  end while
29: end while

```

Algorithm 1. Online algorithm.

Given N neurons and c_0 connections per neuron, we constructed the network's initial random connectivity matrix

following the steps detailed before in Section 2.2. Then, an initial number p_0 of patterns was loaded, equal to its maximum memory capacity, following Equation (3). Given these initial conditions, we alternated the optimization of the network connectivity and the loading of 10 new patterns. Given neuron i , a pre-synaptic neuron j was randomly selected using $randi(N - 1)$ function. If a connection existed for this combination of pre- and post-synaptic neurons ($C_{ij} = 1$), the effect of eliminating the connection was assessed, and inversely the effect of adding the connection was assessed if $C_{ij} = 0$. In both cases, the modification of C_{ij} was kept only if it reduced the local energy. This trial-and-error sequence was repeated 10 times for each neuron, after which pattern stability was tested. If more than 90% of patterns were successfully retrieved, a new set of 10 patterns was loaded. Otherwise, the optimization of C_{ij} was repeated until the retrieval condition was met. The algorithm stopped whenever the network's connectivity matrix did not vary for 50 consecutive repetitions. Since the memory load was no longer fixed, we set the parameter ϵ to a constant value throughout the optimization. In Section 3, we fixed $\epsilon = N/2$.

3. Results

3.1. Noise reduction

3.1.1. Storage capacity

We first simulated networks of different size, with a number of neurons $N \in \{500; 2,000; 5,000\}$ and a number of pre-synaptic connections per neuron covering the full range from highly diluted to fully connected. In each simulation, the connectivity matrix was optimized by implementing the simulated annealing algorithm, which minimized the cost function E_i^0 for each neuron (Equation 6). As a result, we obtained an increment in the storage capacity for all values of N and c . We plotted the ratio between the maximum number of patterns stored by optimized vs. random networks in otherwise identical settings, p_c^{opt}/p_c^{rand} (Figure 1A). This ratio peaked for a connectivity around $c/N \sim 0.20$, independently of the network size, with a storage capacity for optimized networks up to seven-fold higher than the capacity for random networks. For more diluted networks the factor of improvement over random networks decreased. In the highly diluted extreme, $c/N \rightarrow 0$, the improvement ratio tended to 3, while in the opposite limit $c/N \rightarrow 1$ it tended to 1, which can be explained by the absence of degrees of freedom left for the optimization to proceed.

For simulations of increasing c and fixed N , the optimization tended to break the monotonic dependence between p_c and c found in random networks (where p_c is maximal for fully connected networks; Figure 1B). We found that the maximum number of stored patterns (not normalized by c) peaked at a

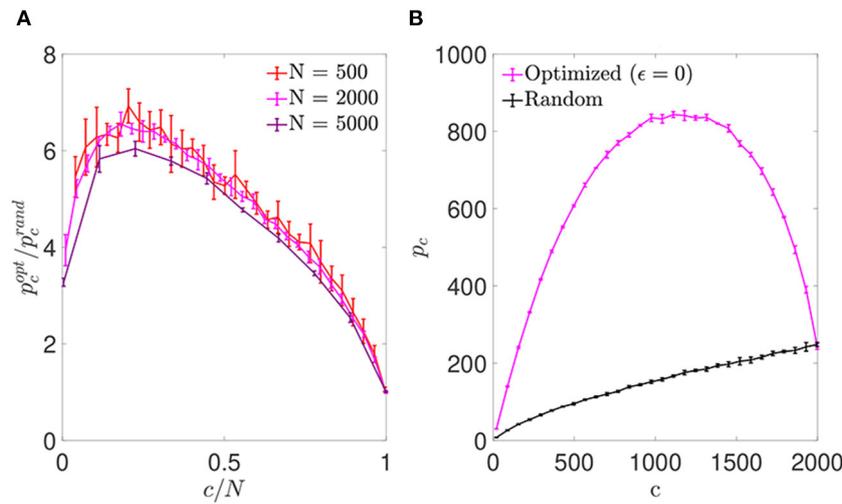


FIGURE 1 Networks optimized by noise reduction outperform randomly connected ones in terms of storage capacity. **(A)** Ratio between the storage capacity for optimized (p_c^{opt}) and random (p_c^{rand}) connectivity as a function of the number of connections per neuron for networks of different N (color coded; mean \pm s.d.). **(B)** Overall maximum number of patterns (not normalized by c) that can be stored and retrieved in an optimized (magenta) or random (black) network of $N = 2,000$ as a function of c (mean \pm s.d.). Note that the maxima for p_c^{opt}/p_c^{rand} **(A)** and p_c^{opt} **(B)** occur for different connectivity levels, implying that more patterns can be stored using a higher number of less efficient connections. Pool of data corresponding to five simulations for each connectivity.

connectivity near $c/N \sim 0.59$, different from the connectivity for which the improvement ratio maximized. This could be explained by a compromise allowing to maximize the overall amount of stored information by including a greater number of less efficient connections.

3.1.2. Structural connectivity analysis

We estimated from the resulting connectivity matrices the conditional probability distribution that, given the synaptic weight W_{ij} , neurons i and j were connected in the optimized scheme. In a network with random connectivity, this probability is independent of W_{ij} because by definition $P(C_{ij} = 1) = \frac{c}{N-1}$, but this was no longer the case in optimized networks. To improve visual comparison of results obtained for different c and p , we plotted this conditional probability $P(C_{ij} = 1|W_{ij})$ vs. W_{ij}/W^M (Figure 2), which represents the available weights normalized by the maximum absolute value reached by the synaptic weight matrix,

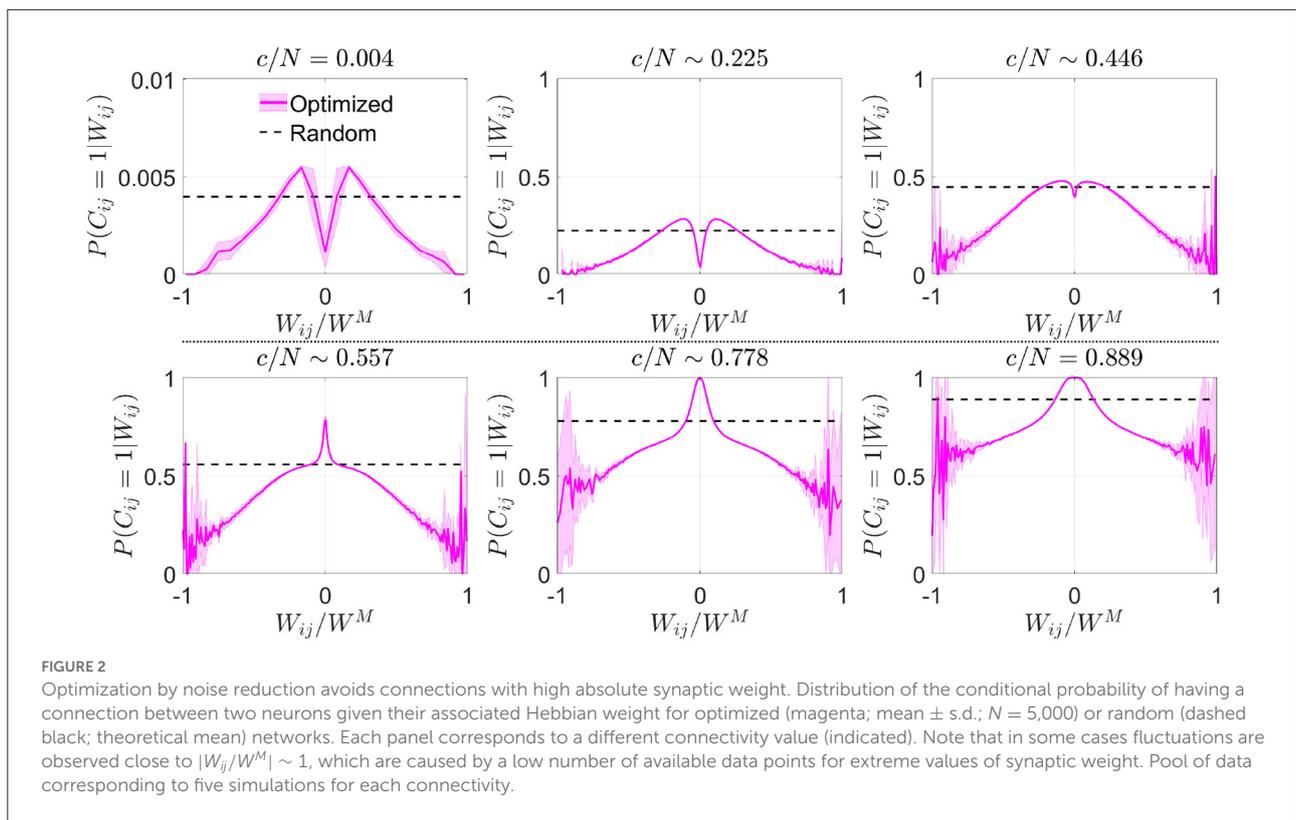
$$W^M = \max_{\substack{i \neq j \\ 1 \leq i, j \leq N}} \{|W_{ij}|\}. \tag{8}$$

Note that W^M is not necessarily equal to p , the maximum theoretical absolute value for W_{ij} . The probability that a given W_{ij} takes a value p is equivalent to the probability of success in p consecutive coin flips, which becomes increasingly unlikely as p grows.

In comparison with the uniform distribution, we observed that the optimized network in the sparse connectivity region tended to favor synapses within a range of low weight values, avoiding those with extreme high or low absolute weight (Figure 2). As connectivity increased, the network started to make use of close-to-zero weights, while still avoiding synapses with high absolute weight. After reaching the connectivity region near $c/N \sim 0.59$ (where pattern storage peaked), a sudden boost in the use of connections with close-to-zero weight was observed. We speculate that the reason behind this behavior is that optimization at this point minimizes the contribution of new connections to the local field. Eventually, synapses with high absolute weight were included, but always with a probability lower than in random networks.

3.1.3. Signal-to-noise analysis

We next characterized the aligned local field distribution ($h_i^v \xi_i^v$) in optimized networks. Given that the heuristic minimization of E_i^0 never took the cost function to exactly zero, neurons received, on average, a non-zero noise contribution, expected to be lower than in randomly connected networks. We analyzed this distribution for different memory loads in a network with fixed c and N (Figure 3A). In contrast to randomly connected networks, the mean aligned field tended to decrease with α , which implied that the optimization found it convenient to anti-correlate the noise and the signal as a means of minimizing E_i^0 . In spite of this anti-correlation, retrieval was possible for a larger range of α



values than what was observed with otherwise identical random networks.

We next studied the standard deviation of the local field to understand if a decrease in variability was compensating for the a priori negative effect of a decrease in mean aligned local field. We observed that, indeed, the increment in the storage capacity could be explained by a substantially narrower noise distribution than the one obtained with random connectivity (Figure 3B). This suggests that fluctuations of the noise term were small enough to let the network store patterns beyond the classical limit, even if they occurred around a negative mean.

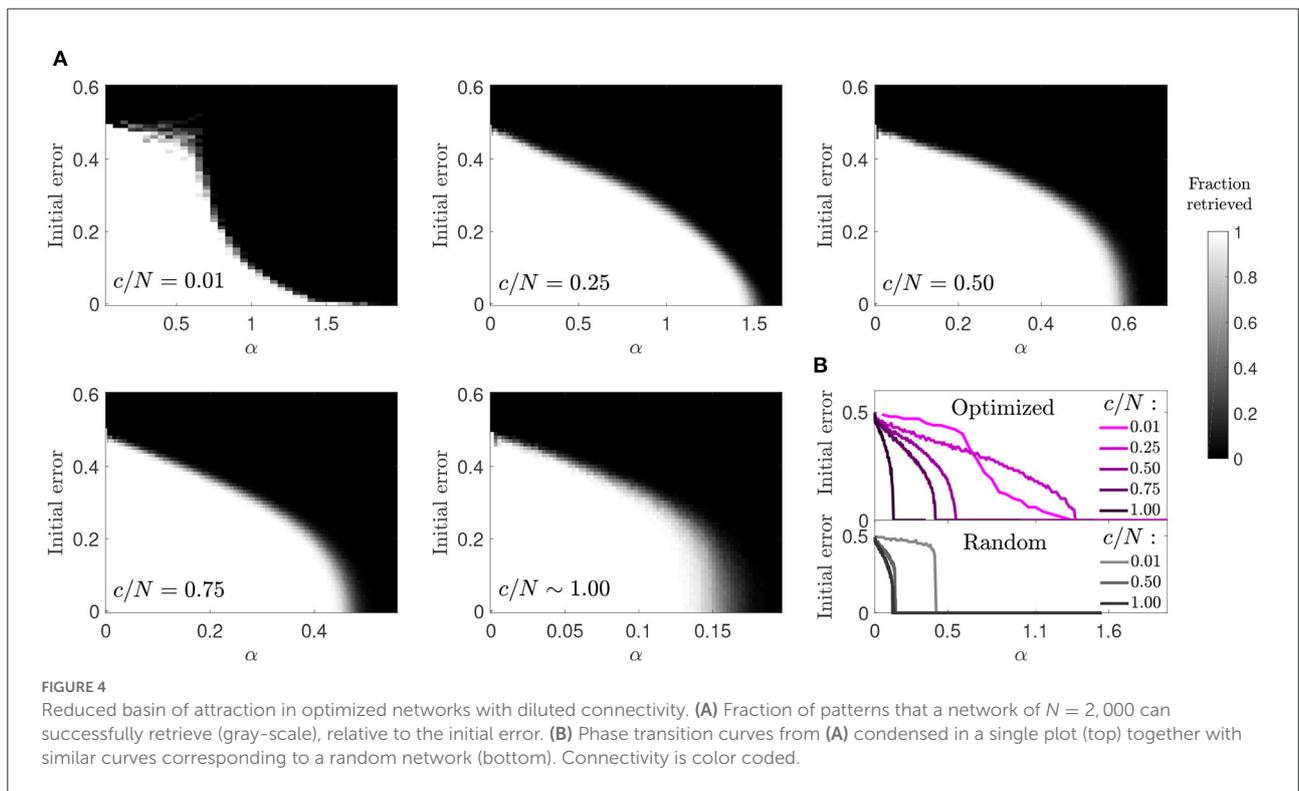
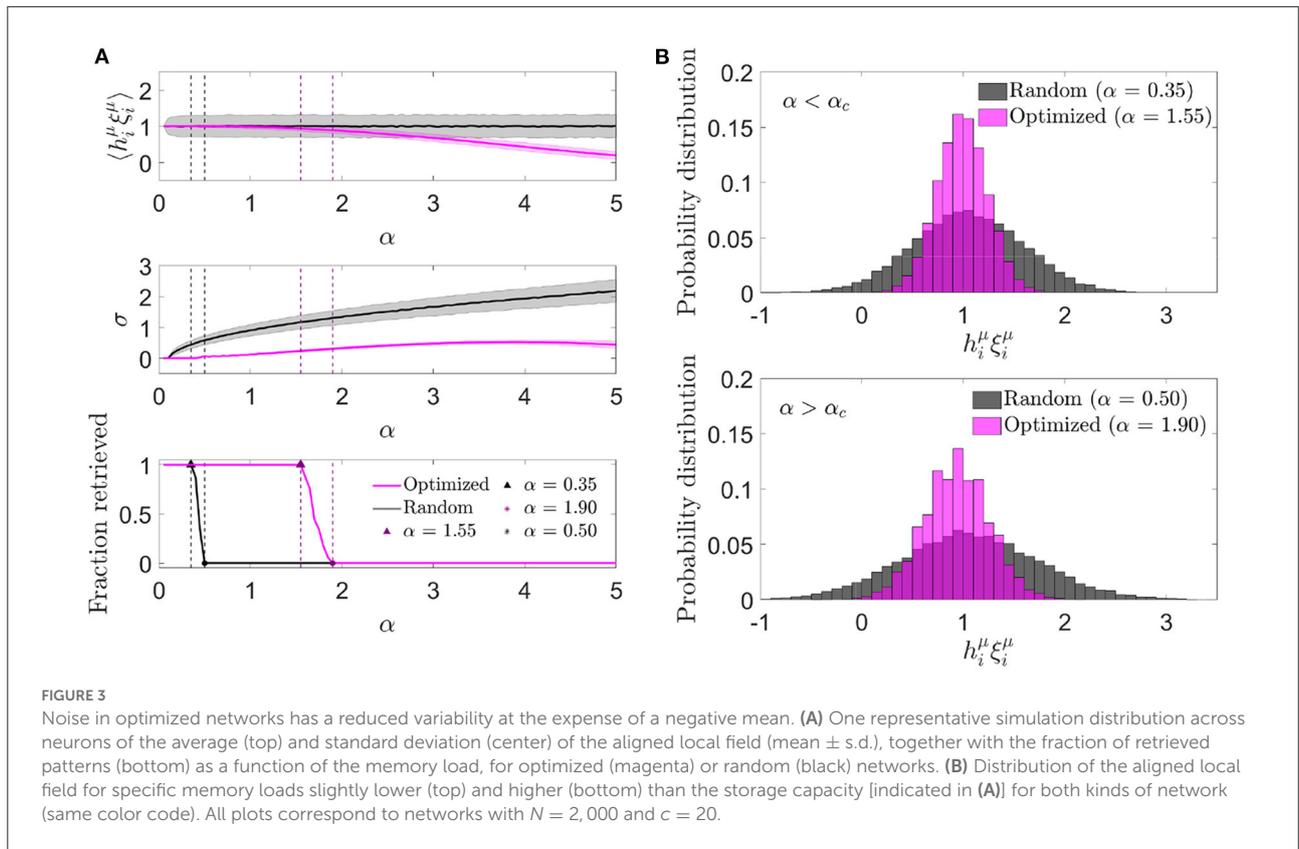
The unexpected finding that, in order to minimize the overall noise, the optimization tended to decrease the mean aligned field to values lower than 1 (implying a negative mean aligned noise), led us to ask if better cost functions would lead to a situation in which the noise reinforced the signal, thus improving the overall performance of the network. We explore this possibility in Section 3.2.

3.1.4. Basin of attraction

We next studied the basin of attraction in optimized networks to understand if an increase in storage capacity came

at the cost of a reduction in attractor strength (Figure 4). We observed that for most connectivity levels the basin of attraction was only moderately reduced. This reduction was comparable to the one observed in random networks, although the transition to the amnesic phase was smoother. The highest contrast with the behavior of random networks was observed in the case of diluted optimized networks ($c/N = 0.01$), where large basins of attraction were observed for values of the memory load α up to around 0.7, above which they tended to be very small. Close to the storage capacity limit, although α was high, the attractors were very weak, as a pattern could only be retrieved if the initial state was almost identical to it. However, for $c/N = 0.25$, a connectivity within the region where the maximum improvement ratio in storage capacity was achieved (Figure 1), attractors were more robust than in diluted optimized networks, and only slightly less robust than in random networks at a comparable distance from the critical storage capacity.

The increase in the size of the basin of attraction from $c/N = 0.01$ to $c/N = 0.25$ in optimized networks is counter intuitive, but analytic modeling of networks with random connectivity provides a potential explanation (Roudi and Treves, 2004). In a non-diluted, random connectivity network, the field of a given neuron has a noise term proportional to its activity, caused by feedback of the neuron's activity state through multiple-synaptic loops. The probability of existence of such connectivity loops,



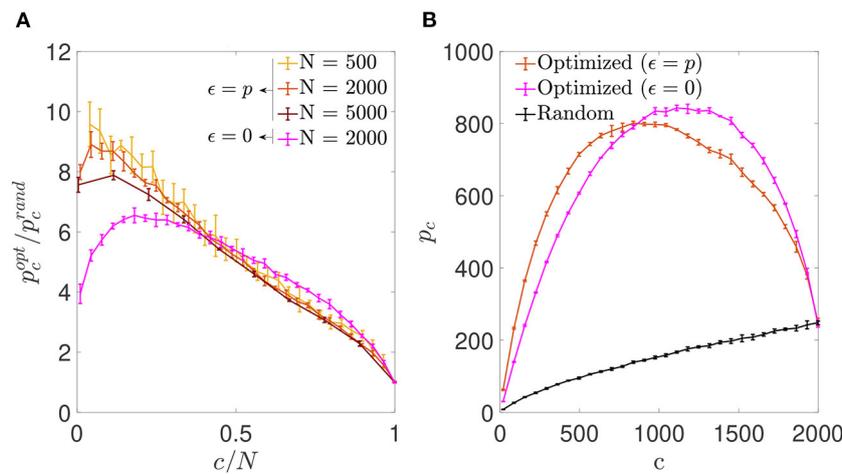


FIGURE 5 Signal reinforcement outperforms noise reduction and is optimal with highly diluted connectivity. **(A)** Ratio between the storage capacity for optimized (p_c^{opt}) and random (p_c^{rand}) connectivity as a function of the number of connections per neuron for networks of different N (shades of red; mean \pm s.d.). For comparison, the curve corresponding to noise reduction is included (magenta; $N = 2,000$). **(B)** Overall maximum number of patterns (not normalized by c) that can be stored and retrieved as a function of c in a network with $N = 2,000$ optimized by signal reinforcement (orange), optimized by noise reduction (magenta) or with random connectivity (black) (mean \pm s.d.). Note that, as in Figure 1, the maxima for p_c^{opt}/p_c^{rand} **(A)** and p_c^{opt} **(B)** occur for different connectivity levels. Pool of data corresponding to five simulations for each connectivity.

starting in the neuron and projecting back to it, decreases toward zero in diluted networks, and so thus the effect of this noise term. It is possible that the optimized network leverages from multi-synaptic feedback loops to increase the robustness of attractors, as the sudden decay in the size of the basin of attraction but not in overall storage capacity from $c/N = 0.25$ to $c/N = 0.01$ seems to suggest.

3.2. Signal reinforcement

3.2.1. Storage capacity

Inspired by the results in the previous section, we next assessed the possibility of optimization by reinforcement of the signal rather than noise reduction, by using the same optimization procedure with a different cost function (Equation 7). We observed that, as in the previous section, the optimization process increased the storage capacity of autoassociative networks (Figure 5A). However, in contrast to the $\epsilon = 0$ case, the improvement ratio p_c^{opt}/p_c^{rand} for the optimization with $\epsilon = p$ increased monotonically with network dilution for most of the c/N range, reaching its best performance $p_c^{opt}/p_c^{rand} \sim 10$ around the lowest connectivity values.

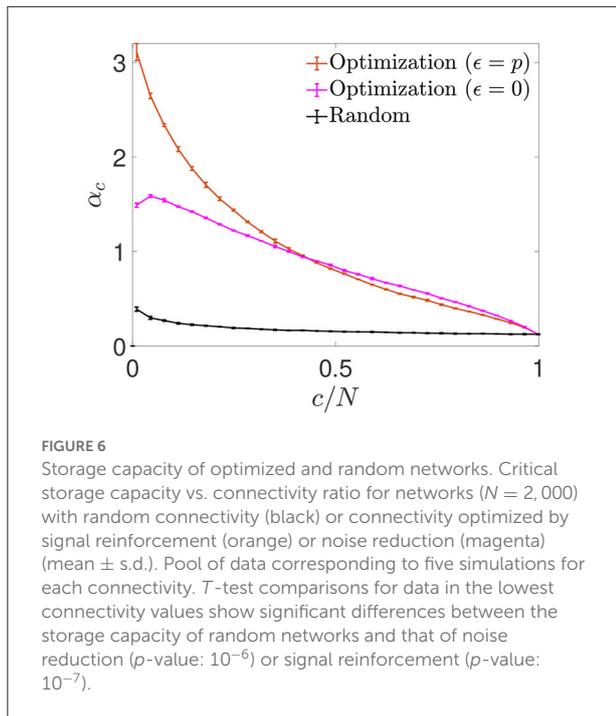
We also observed that the overall number of patterns p that a network of fixed N could store and retrieve was higher if optimized by minimizing E_i^p than by minimizing E_i^0 in the low connectivity range (Figure 5B), but the opposite was true for high connectivity. While the connectivity capable of storing more patterns for the optimization with $\epsilon = 0$ was near $c/N \sim$

0.59, for the optimization $\epsilon = p$ it was $c/N \sim 0.42$, with slightly fewer patterns in total. However, both of these limits are far from the biologically plausible limit of high dilution, where signal-reinforcing networks seem to outperform noise-minimization ones.

We next plotted the α_c curves as a function of c/N (Figure 6), which allowed for a better visualization of the fact that for a connectivity of $c/N = 0.01$, both optimizations achieved an enhanced storage capacity, in the case of signal-reinforcement networks up to an order of magnitude greater than obtained using random connectivity. When minimizing E_i^0 , the capacity was close to $\alpha_c \sim 1.49$, while in E_i^p minimization, the critical capacity more than doubled the latter, reaching $\alpha_c \sim 3.15$.

3.2.2. Structural connectivity analysis

We next asked if the criteria for selecting synapses in the signal reinforcement scenario were similar to those found in the previous section. We observed that the conditional distribution $P(C_{ij} = 1 | W_{ij})$ obtained in the low connectivity range increased monotonically with the absolute value of the synaptic weights, implying that, in contrast to what was previously observed, this time the optimization process favored connections with high absolute weight (Figure 7). As connectivity increased, networks tended to use all available synapses with high absolute value, which forced them to include lower modulus synapses. For an intermediate range of connectivity ($>c/N \sim 0.42$, where pattern storage peaked), close-to-zero weight synapses started being used, similarly to what happened in the noise reduction



scenario. For highly connected networks ($c/N \gtrsim 0.75$), the pool of available close-to-zero synapses also depleted and more intermediate-range synapses were included. This strategy of selecting synapses with a high degree of consensus across patterns seems opposite to the one described for noise reduction.

3.2.3. Signal-to-noise analysis

As previously, we plotted the mean and standard deviation of the aligned local field as a function of the memory load α (Figure 8). We observed that the mean aligned local field increased monotonically with α , as expected from the fact that the cost function was designed to make the noise reinforce the signal, reaching a substantial deviation from 1, its zero-noise value. The standard deviation also increased with α , although it remained lower than the one corresponding to an otherwise identical random network for values of α where retrieval was possible. At α_c , the standard deviation for the optimized network was noticeably higher than the one corresponding to the random network considered at its own α_c value, implying that the increase in storage capacity was due to an ability to make the mean aligned local field increase faster than its standard deviation for a limited range of values of α .

3.2.4. Basin of attraction

Similar to the previous optimization, and in contrast to random connectivity networks, the basin of attraction decreased progressively rather than abruptly with memory load (Figure 9).

However, when comparing the two optimization strategies, we observed that memories in the $\epsilon = p$ scenario were more robust in the connectivity range $c/N \lesssim 0.4$. This indicates that the new optimization strategy was better than the previous one due to a synergistic combination of increased storage capacity and larger basins of attraction. As connectivity increased, the basins of attraction tended to look like the ones observed in random networks, as in the previous optimization case.

3.3. Online algorithm

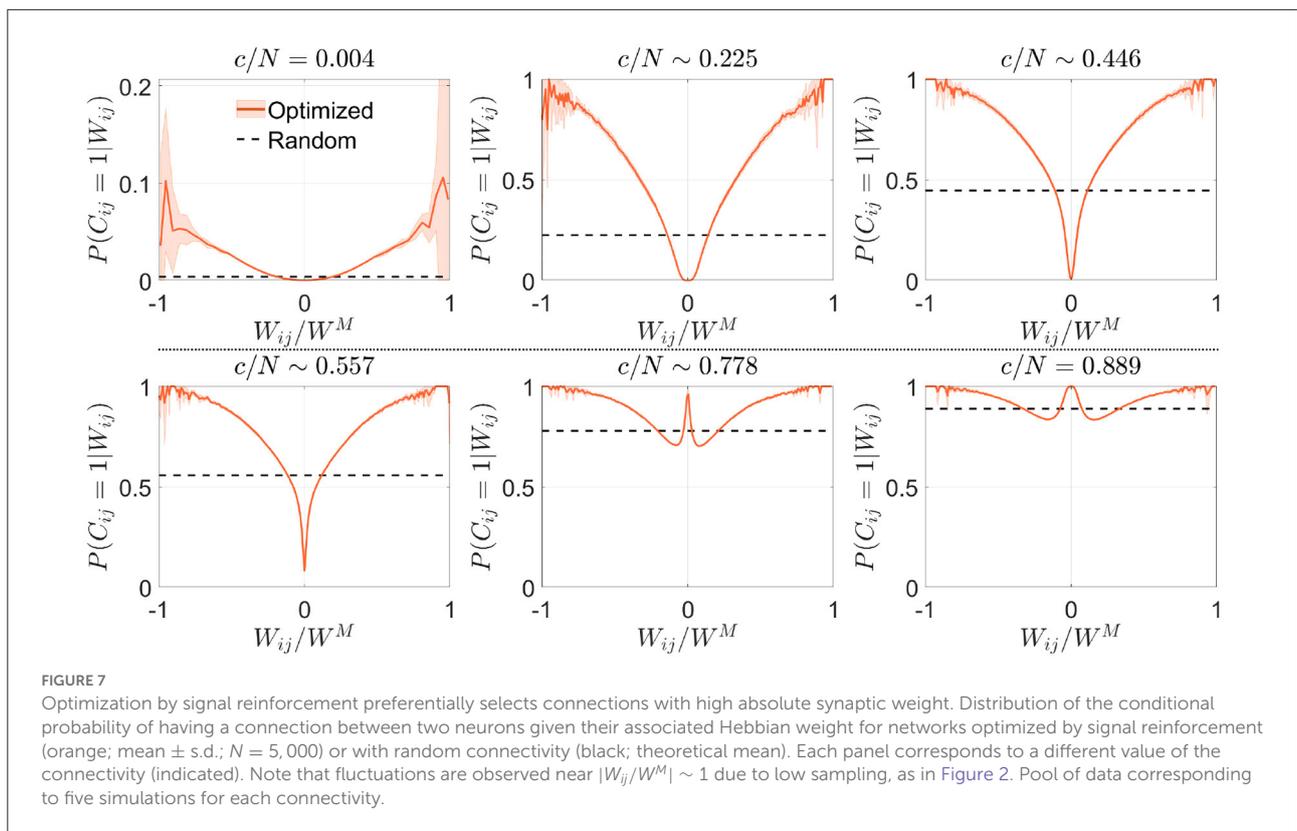
To simulate a situation similar to the development of the human brain, where connectivity changes as the subject learns, we explored an online learning scenario with dynamic generation and elimination of synapses. Results shown in this section correspond to simulations of $N = 2000$ networks with different initial random connectivity values c_0 . The connectivity optimization was done by implementing Algorithm 1 with the cost function shown in Equation (7), and setting $\epsilon = N/2$.

3.3.1. Mean connectivity variation

We first studied the evolution of the mean number of connections per neuron along the learning process. As previously mentioned, the algorithm allowed each neuron to freely eliminate or generate connections, with minimization of their own cost function as the only constraint. We asked if the final number of connections, which was a priori unknown, depended on initial conditions (for example on the initial connectivity value c_0) or on the history of the learning process. We observed that the final average connectivity did not depend on initial conditions, stabilizing at a value $c_{est} \sim 445$ for all tested c_0 (this stability value depended on ϵ) (Figure 10). For networks with initial connectivity above c_{est} the average connectivity decreased monotonically, i.e., pruning predominated throughout the learning process. For networks with initial connectivity below c_{est} we observed a dynamic qualitatively analogous to the maturation of the human brain, with an initial creation-predominant period and a late pruning-predominant period.

3.3.2. Storage capacity

The online algorithm substantially increased the storage capacity of the autoassociative networks compared to what would be obtained without modification of the connectivity matrix. The improvement was comparable to the one obtained with the offline simulated annealing algorithm E_i^p (Figure 11). As observed with mean connectivity, the final effective storage capacity reached by the network did not depend on the initial connectivity, although it was strongly dependent on the ϵ parameter. Interestingly, when the initial connectivity was set



to highly diluted, the initial increase in p_{eff} occurred along the storage capacity limit for random networks, implying that the optimization was similar to what would be obtained by adding random connections, up to a connectivity level close to the final one. From that point on, optimization proper took place, and the optimized p_{eff} increased above the random limit in a process initially dominated by addition and later on by pruning. In contrast, the effect of selective pruning was never equivalent to that of a random one. Given the monotonically increasing relationship between p_c and c in random networks, random pruning necessarily results in a decay in p_c .

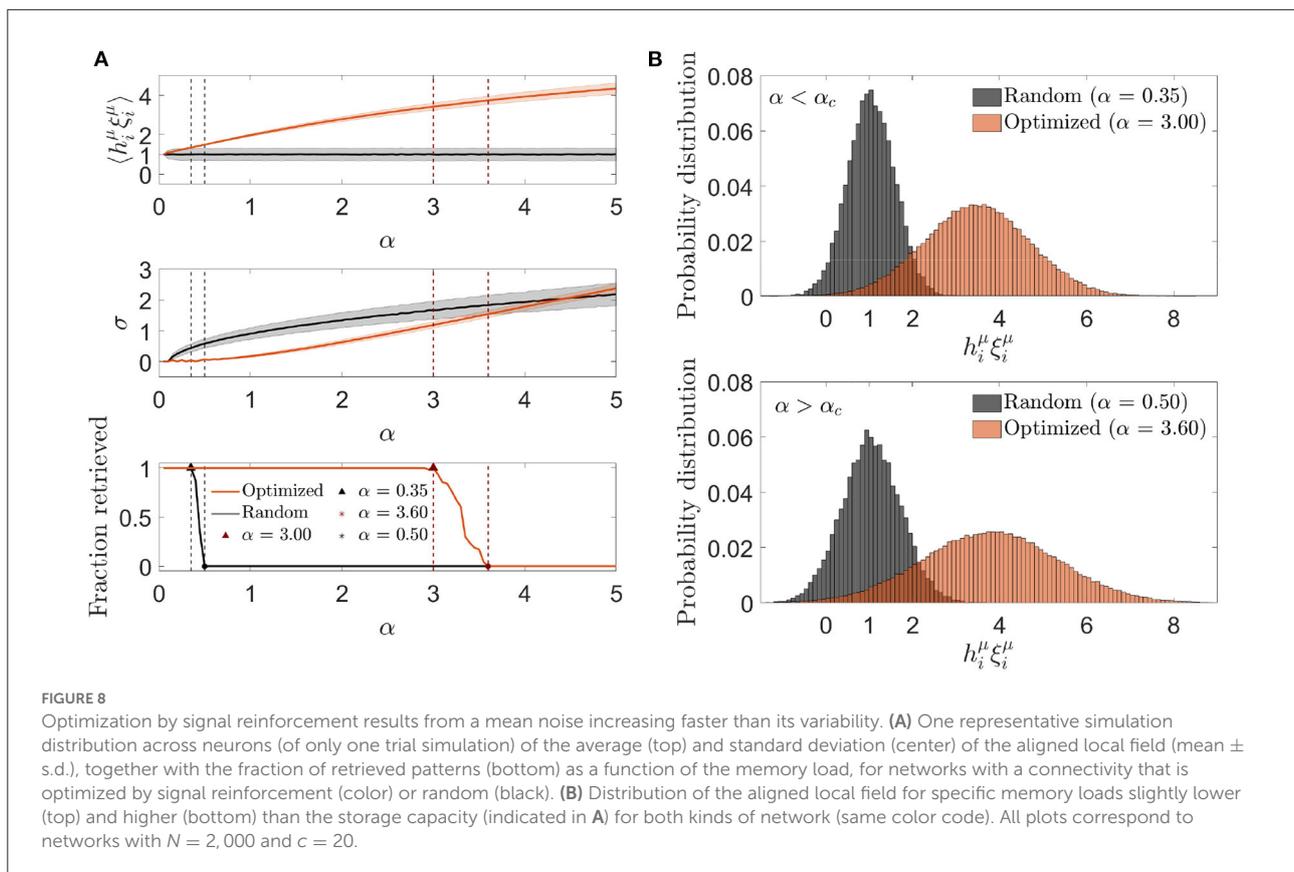
Put together, these results suggest that the algorithm was able to converge to a connectivity configuration that represented a global minimum in the space of solutions, with a unique intermediate dynamics that depended on initial conditions.

4. Discussion

Our main result is that it is possible to improve the storage capacity of an autoassociative network by optimizing its connectivity matrix. We found that if structural connectivity is optimized to minimize the noise present in the local field of each neuron, up to a seven-fold improvement in the storage

capacity can be obtained in comparison to random networks. This maximal improvement, however, occurs with a relatively dense connectivity close to $c/N \sim 0.20$, which is not typical in the mammalian brain. We next found that an even stronger improvement could be obtained by reinforcing the signal rather than minimizing the noise, this time reaching almost one full order of magnitude in the biologically plausible highly diluted limit. These results suggest that diluted autoassociative networks would benefit most from a mechanism where the connectivity configuration allows the noise term to reinforce the signal. The fact that all networks studied in this work have a similar matrix of synaptic weights, governed by the Hebbian principle, and yet levels of performance that span over one order of magnitude, stresses the importance of considering the remodeling of network architecture as a process potentially independent and complementary to synaptic weight modification.

We found that in the diluted connectivity region, the signal-reinforcement optimization was achieved by selecting connections with high consensus across patterns, with a probability that decreased with the modulus of the synaptic weight. Our results are in the same direction of previous work showing a substantial storage capacity increase in a highly diluted network that only retains the c strongest connections per neuron (Montemurro and Tamarit, 2001). This *extreme pruning* scheme was shown analytically to make the storage capacity diverge in the limit $c \ll \min\{\frac{N}{2p-1}, \ln(N)\}$, which, however, is

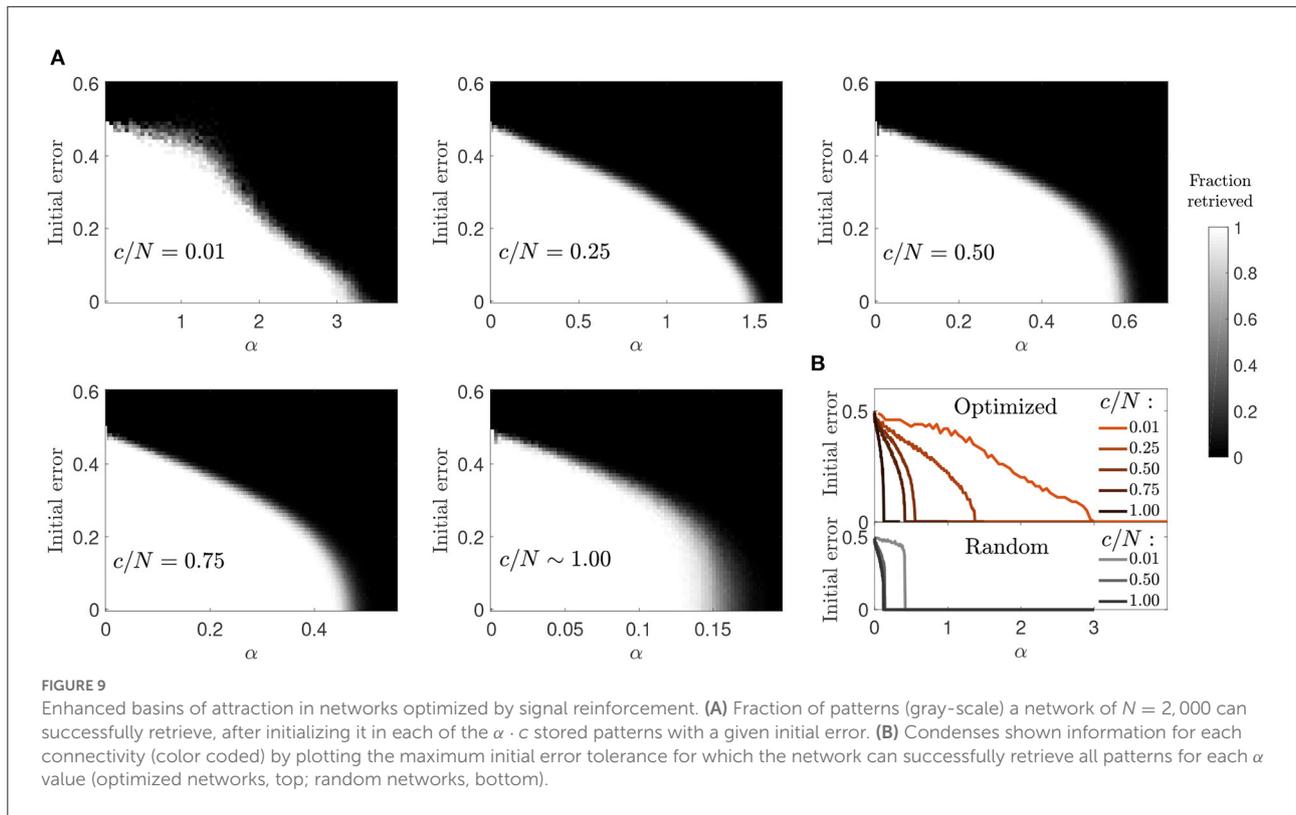


too diluted for real brains. Simulations under the more plausible constraint of $c \ll N$ resulted in a storage capacity smaller but in the range of the one we have found. This is consistent with our observation that, although optimized networks tend to prefer high absolute weights, they do not include only the highest ones, and instead recruit as well a number of intermediate weights. More generally, the idea of a connectivity scheme that reinforces the Hebbian principle was presented in the cited work and is also part of our results using the signal-reinforcement optimization.

Our model used rather simplistic units and architecture. Similar results could perhaps be replicated in models that include more realistic elements such as networks with graded response units and non binary patterns (Treves and Rolls, 1991). We speculate that graded response units may provide the network with one extra degree of freedom, since, for a given neuron, the demands for connectivity optimization could be weighted by its activity level in each stored pattern. Regarding architecture, since our network does not mimic the hierarchical organization of the brain (Fulvi Mari, 2004), it most suitably models phenomena taking place in short-range cortical networks, where remodeling of

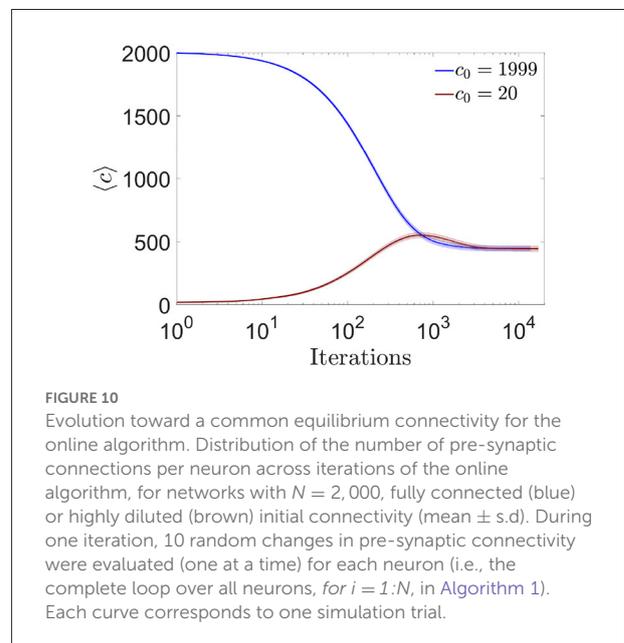
connections is most relevant and has been most extensively studied.

Brains do not have a fixed connectivity throughout the lifetime of humans and other mammals. Evidence indicates that connectivity levels increase during childhood, reaching a maximum around the age of 2, and then decrease, reaching a stable value in late adolescence or early adulthood (Huttenlocher, 1979; Lichtman and Colman, 2000; Hua and Smith, 2004; Navlakha et al., 2015), a period roughly coinciding with the one originating most of long lasting memories (Rubin et al., 1998). This implies that addition and pruning take place in parallel to the most substantial stages of learning. In a simplified model of this process, our online algorithm explores the evolution of connectivity and storage capacity in a network where neurons can freely add or delete connections as the network incorporates more patterns. We observed that if the network is initialized at a high connectivity level, it mainly prunes connections, with an immediate gain over random connectivity setups. However, if its initial connectivity is low, a process with several stages takes place. During the initial stage, connections are added with a gain in storage capacity equivalent to the addition of random connections, implying that a similar

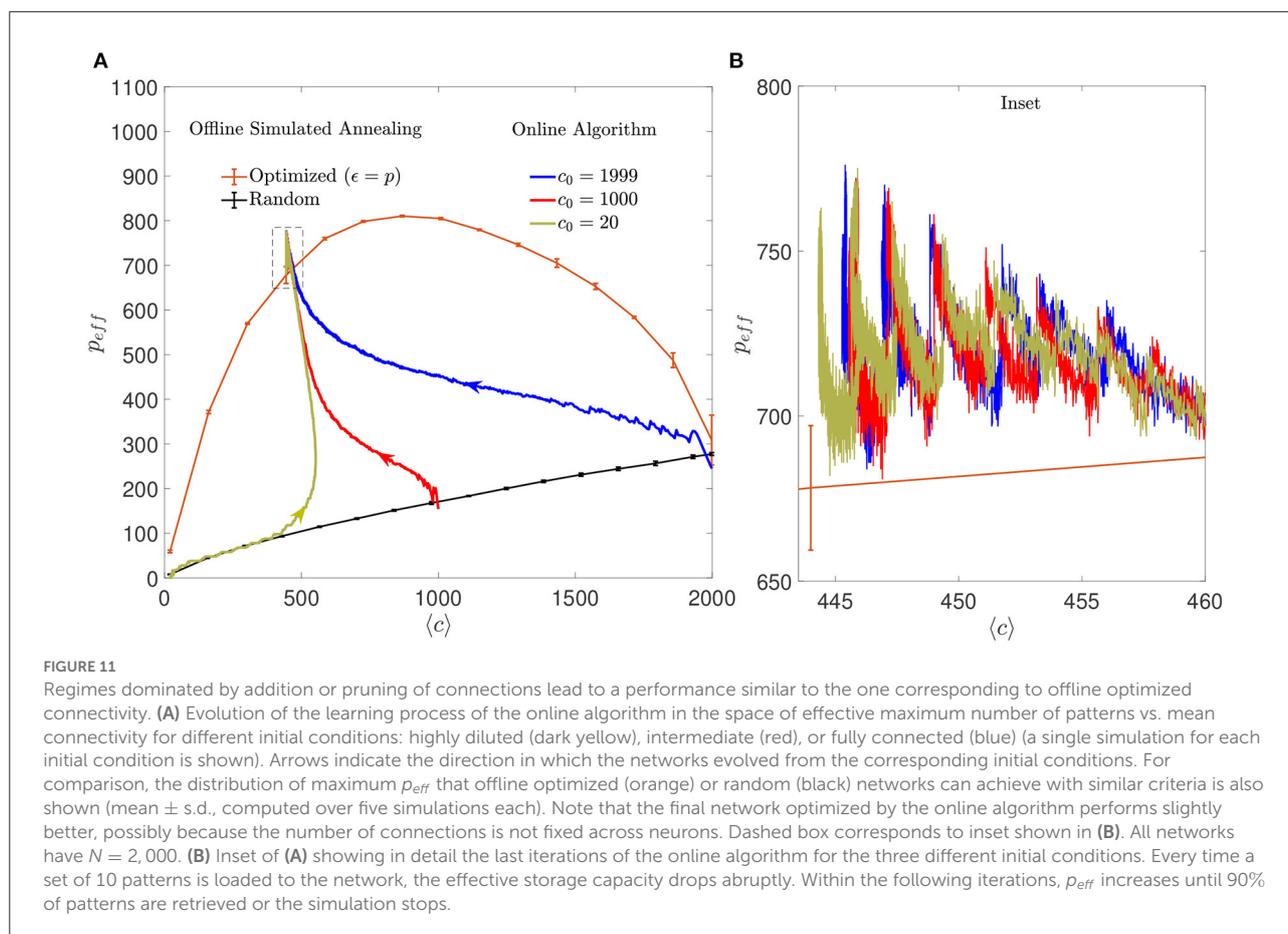


path would be followed by networks randomly initialized in a range of values of low connectivity. Equivalently, this first stage that could be associated with the initial development of the brain could add connections randomly instead of selectively, with a similar effect on performance. Once connectivity is close to the final one (unknown a priori but dependent on the parameter ϵ), the addition of connections becomes more meaningful, improving the storage capacity beyond the level of random networks. Remarkably, this process takes connectivity beyond the final value, reminiscent of what is actually observed during childhood in the human brain. In a third stage, increase in storage capacity is associated with a net loss of connections, until an equilibrium is reached and storage capacity is close to the global minimum of the cost function obtained with offline minimization. Although this dynamic is complex, the final state of the network is independent of initial conditions, suggesting a strong global minimum reached by the optimization process.

Throughout this work we have used an extremely simple model of autoassociative memory that, however, captures all the basic behaviors exhibited by this kind of network even in biologically plausible or experimental setups (Roudi and Latham, 2007; Carrillo-Reid et al., 2016). Using mean network activity of 50%, as in the original Hopfield model, helped making our simulations computationally efficient by keeping the storage capacity relatively low. The effects of connectivity



optimization on sparser networks, however, with a biologically plausible value around 5 or 10%, should be explored in the future. One possibility is that connectivity optimization becomes less effective because the pool of active neurons from which



to choose is more restricted. In contrast, we speculate that other modifications aiming to add biological plausibility could make connectivity optimization even more relevant than in the present work. For example, if memorized patterns were obtained from a distribution exhibiting a set of non-trivial correlations (Kropff and Treves, 2007), the initial stage of optimization could capture early the main statistical features of inputs, making corrections at a later stage less important or necessary. In addition, shared features across patterns might help in the reinforcement of the signal by the noise. It has been shown that in this scenario the selected deletion of neurons can yield a significant improvement in storage capacity, implying that a much more subtle correction in connectivity could lead to similar or higher performance levels (Kropff, 2007).

We have shown that in a simple setup, connectivity optimization can improve storage capacity, in some cases up to around one order of magnitude. Though considerable, this gain might still be not enough to explain the viability of attractor networks in the human brain, given all the drawbacks of biologically plausible networks. Further work on the effect

of optimized connectivity in networks that include some of these details of biological plausibility could perhaps place our estimation of storage capacity at least in the reasonable range between thousands and tens of thousands of memories, a limit compatible with the number of faces we remember or the number of words we use in our daily life.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://github.com/Facuemina/Selective-Connectivity-2022>.

Author contributions

FE and EK planned the simulations, analyzed the results, and wrote the manuscript. FE wrote the scripts and ran the simulations. Both authors contributed to the article and approved the submitted version.

Funding

This work had support from PICT-2019-2596 (Science Ministry, Argentina) and Human Frontiers Science Program RGY0072/2018 (EK).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

References

- Amit, D. (1989). *Modelling Brain Function*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511623257
- Arenzon, J. J., and Lemke, N. (1994). Simulating highly diluted neural networks. *J. Phys. A* 27:5161. doi: 10.1088/0305-4470/27/15/016
- Assaf, Y., Bouznach, A., Zomet, O., Marom, A., and Yovel, Y. (2020). Conservation of brain connectivity and wiring across the mammalian class. *Nat. Neurosci.* 23, 805–808. doi: 10.1038/s41593-020-0641-7
- Bullmore, E., and Sporns, O. (2012). The economy of brain network organization. *Nat. Rev. Neurosci.* 13, 336–349. doi: 10.1038/nrn3214
- Carrillo-Reid, L., Yang, W., Bando, Y., Peterka, D. S., and Yuste, R. (2016). Imprinting and recalling cortical ensembles. *Science* 353, 691–694. doi: 10.1126/science.aaf7560
- Chklovskii, D. B., Mel, B. W., and Svoboda, K. (2004). Cortical rewiring and information storage. *Nature* 431, 782–788. doi: 10.1038/nature03012
- Citri, A., and Malenka, R. C. (2008). Synaptic plasticity: multiple forms, functions, and mechanisms. *Neuropsychopharmacology* 33, 18–41. doi: 10.1038/sj.npp.1301559
- Derrida, B., Gardner, E., and Zippelius, A. (1987). An exactly solvable asymmetric neural network model. *Europhys. Lett.* 4:167. doi: 10.1209/0295-5075/4/2/007
- Forrest, B. M. (1988). Content-addressability and learning in neural networks. *J. Phys. A* 21:245. doi: 10.1088/0305-4470/21/1/029
- Fulvi Mari, C. (2004). Extremely dilute modular neuronal networks: neocortical memory retrieval dynamics. *J. Comput. Neurosci.* 17, 57–79. doi: 10.1023/B:JCNS.0000023871.60959.88
- Gardner, E. (1988). The space of interactions in neural network models. *J. Phys. A* 21:257. doi: 10.1088/0305-4470/21/1/030
- Herculano-Houzel, S., Mota, B., Wong, P., and Kaas, J. H. (2010). Connectivity-driven white matter scaling and folding in primate cerebral cortex. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19008–19013. doi: 10.1073/pnas.1012590107
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558. doi: 10.1073/pnas.79.8.2554
- Hua, J. Y., and Smith, S. J. (2004). Neural activity and the dynamics of central nervous system development. *Nat. Neurosci.* 7, 327–332. doi: 10.1038/nn1218
- Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex—developmental changes and effects of aging. *Brain Res.* 163, 195–205. doi: 10.1016/0006-8993(79)90349-4
- Janowsky, S. A. (1989). Pruning versus clipping in neural networks. *Phys. Rev. A* 39:6600. doi: 10.1103/PhysRevA.39.6600
- Kropff, E. (2007). “Full solution for the storage of correlated memories in an autoassociative memory,” in *Proceedings of the International Meeting on “Closing the Gap Between Neurophysiology and Behaviour: A Computational Modelling Approach”, Vol. 163* (Abingdon: Taylor and Francis), 195–205.
- Kropff, E., and Treves, A. (2005). The storage capacity of Potts models for semantic memory retrieval. *J. Stat. Mech.* 2005:P08010. doi: 10.1088/1742-5468/2005/08/P08010
- Kropff, E., and Treves, A. (2007). Uninformative memories will prevail: the storage of correlated representations and its consequences. *HFSP J.* 1, 249–262. doi: 10.2976/1.2793335
- Lichtman, J. W., and Colman, H. (2000). Synapse elimination and indelible memory. *Neuron* 25, 269–278. doi: 10.1016/S0896-6273(00)80893-4
- Montemurro, M. A., and Tamarit, F. A. (2001). An efficient dilution strategy for constructing sparsely connected neural networks. *Phys. A* 294, 340–350. doi: 10.1016/S0378-4371(01)00123-6
- Navlakha, S., Barth, A. L., and Bar-Joseph, Z. (2015). Decreasing-rate pruning optimizes the construction of efficient and robust distributed networks. *PLoS Comput. Biol.* 11:e1004347. doi: 10.1371/journal.pcbi.1004347
- Roudi, Y., and Latham, P. E. (2007). A balanced memory network. *PLoS Comput. Biol.* 3:e141. doi: 10.1371/journal.pcbi.0030141
- Roudi, Y., and Treves, A. (2004). An associative network with spatially organized connectivity. *J. Stat. Mech.* 2004:P07010. doi: 10.1088/1742-5468/2004/07/P07010
- Rubin, D. C., Rahhal, T. A., and Poon, L. W. (1998). Things learned in early adulthood are remembered best. *Mem. Cogn.* 26, 3–19. doi: 10.3758/BF03211366
- Treves, A., and Rolls, E. T. (1991). What determines the capacity of autoassociative memories in the brain? *Network* 2, 371–397. doi: 10.1088/0954-898X_2_4_004
- Tsodyks, M. (1989). Associative memory in neural networks with the Hebbian learning rule. *Modern Phys. Lett. B* 3, 555–560. doi: 10.1142/S021798498900087X
- Tsodyks, M. V., and Feigel'man, M. V. (1988). The enhanced storage capacity in neural networks with low activity level. *Europhys. Lett.* 6:101. doi: 10.1209/0295-5075/6/2/002
- Yang, X.-S. (2014). “Chapter 4: Simulated annealing,” in *Nature-Inspired Optimization Algorithms*, ed X.-S. Yang (Amsterdam: Elsevier), 67–75. doi: 10.1016/B978-0-12-416743-8.00004-X

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.