Enhanced storage capacity with errors in scale-free Hopfield neural networks: An analytical study

Do-Hyun Kim[1]*, Jinha Park[2], Byungnam Kahng[2]*

1 Department of Physics, Sogang University, Seoul, Korea, **2** CCSS, CTP and Department of Physics and Astronomy, Seoul National University, Seoul, Korea

* dohyunkim@sogang.ac.kr (D-HK); bkahng@snu.ac.kr (BK)

Abstract

The Hopfield model is a pioneering neural network model with associative memory retrieval. The analytical solution of the model in mean field limit revealed that memories can be retrieved without any error up to a finite storage capacity of $O(N)$, where N is the system size. Beyond the threshold, they are completely lost. Since the introduction of the Hopfield model, the theory of neural networks has been further developed toward realistic neural networks using analog neurons, spiking neurons, etc. Nevertheless, those advances are based on fully connected networks, which are inconsistent with recent experimental discovery that the number of connections of each neuron seems to be heterogeneous, following a heavy-tailed distribution. Motivated by this observation, we consider the Hopfield model on scale-free networks and obtain a different pattern of associative memory retrieval from that obtained on the fully connected network: the storage capacity becomes tremendously enhanced but with some error in the memory retrieval, which appears as the heterogeneity of the connections is increased. Moreover, the error rates are also obtained on several real neural networks and are indeed similar to that on scale-free model networks.

Introduction

Human neuroscience has attracted increasing attention through various studies. Among such activities, the retrieval or recall of associative memory in neural networks is a historically noticeable issue [1, 2]. Associative memory means the ability to link, learn and remember the relationship between independent and unrelated items such as one man's name (e.g., Albert Einstein) and his previous achievement (e.g., $E = mc^2$). Neural network models of associative memory have been used to explain how the brain stores and recalls long-term memories. These models incorporate the so-called Hebbian rule for a cell assembly, a group of excitatory neurons mutually coupled by strong synapses [3]: Memory storage occurs when a cell assembly is created by Hebbian synaptic plasticity, and memory retrieval or recall occurs when the neurons in the cell assembly are activated by a stimulus. Neural network models of associative memory assume that information exists alternatively as neural activity or as synaptic

connections. When novel information first enters into the brain, it is encoded in a pattern of neural activity. If this information is stored as memory, the neural activity leaves synaptic connections modified. The stored information can be retrieved when the modified connections again become active [4].

A Hopfield model was introduced in the year 1982 [5], following which there have been enormous researches based on the premise that retrieval of associative memory occurs by way of pattern recognition in collective excitations of associated neurons. The Hopfield model has been thus accepted as a paradigmatic neural network model for associative memory retrieval. A Hopfield network is composed of Ising-type neurons with two discrete states, that is, an excitation pattern of each neuron is in a state either +1 or −1, representing excited and rest states for transmitting or not transmitting a signal, respectively. Each neuron is supposed to be connected to all other neurons and then the synaptic weight is updated by the Hebbian rule. An associative memory was introduced as an activity pattern by collective excitations of the associated neurons. A simple example of excitation updated by the Hebbian rule is presented in Section 1 of the supporting information (S1 File).

In this model study, a mathematical quantity called *energy* was introduced to each memory pattern. This quantity decreases as the retrieval activity proceeeds until the system reaches a stable state, at which the retrieved pattern is consistent to a stored pattern. We will show why such behavior occurs in the Hopfield model in Section 2 of the S1 File. This behavior is similar to the dynamic process of a thermodynamic system toward an equilibrium state, at which a free energy becomes minimum.

In associative neural networks, the number of patterns that can be stored in synaptic connections is a central quantity to calibrate their performance. The maximal number of patterns that can be stored before having total confusion divided by the total number of neurons in the network is called the *storage capacity* of the network. If the cell assemblies share no neurons in common, the number of patterns that can be stored is as many as the total number of neurons in the network. If the cell assemblies share some neurons, however, interference may occur among those cell assemblies. If too many patterns are stored in the network, the stored patterns can be interfered and the quality of memory recall can be lowered. Therefore, interference effect induces errors in memory retrieval and reduces the retrievability [4]. Furthermore, stochastic processes in any real network system can occur. We thus consider a temperature T, which is not to be understood as a physical temperature, but as a noise strength for the stochastic process. In the limit of zero temperature ($T = 0$), the deterministic model is recovered. The Hopfield model treated such noise effect in terms of temperature, and invoke the formalism of Boltzmann statistical mechanics to obtain various properties of the retrievability as a function of both the number of stored patterns and temperature, i.e., noise strength. Amit *et al.* calculated the storage capacity of the Hopfield model on a fully connected network at finite temperature using a method in statistical physics [6, 7].

Recent experimental results using the functional magnetic resonance imaging (fMRI) technique have revealed that functional neural networks in resting state are not fully connected networks but the number of connections of each coarse-grained neuron are heterogeneous following a heavy-tailed distribution [8, 9]. For further discussion, the number of connections of a neuron is referred to as degree using a term in graph theory. Even though it is not clear yet how functional connections of an individual neuron is related to its anatomical connections, it becomes more acceptable that neurons with similar patterns of connection tend to exhibit similar functions [8]. This suggests that neural networks need not necessarily be a fully connected network. Indeed, there exists supporting evidences: a structural neural network of the worm *Caenorhabditis elegans* has a power-law tail in the degree distribution [10]. Mathematically, this is expressed as $P_d(k) \sim k^{-\gamma}$, where $P_d(k)$ is the degree distribution, k denotes degree and γ

is the degree exponent. Such networks are referred to as scale-free (SF) networks. Recent electrophysiological data [11] also revealed that the distribution of synaptic connections strength follows a log-normal distribution, which has a heavy tail similar to that in the power-law distribution. Moreover, it is known that the brain damages of schizophrenia [12, 13] or comatose patients [14] are caused by the malfunctioning of hub neurons. Thus, hub neurons that have many connections is likely to exist in a brain network.

This paper aims to investigate the properties of the retrieval patterns created by the Hopfield model on SF networks at finite temperature analytically. We also compare obtained results with previous numerical results obtained from fully connected networks and diverse SF networks such as the Barabási-Albert model with $\gamma = 3$ [15] and the Molloy-Reed model with several values of γ [16]. We also use the Chung-Lu model [17, 18] to construct uncorrelated SF networks over the entire range of γ. The details of the Chung-Lu model are presented in the Method section. Particularly we consider the limit $\gamma \rightarrow 2$, which is the case observed in the fMRI data [19–21].

The main results of our studies are presented in Figs 1 and 2. In Fig 1, we present properties of the retrieval pattern for various degree exponents γ as a fuction of temperature T and storage rate a. When $\gamma < 2.04$ (Fig 1(e) and 1(f)), the retrieval phase spans most of the low-temperature (noise) region; thus memory retrieval in the system is appreciably enhanced compared with the one of the original Hopfield model [6, 7]. In Fig 2, the error rate, a fraction of neurons which fails memory retrieval, is obtained as a function of storage rate $a \equiv p/N$, the number of stored patterns per neuron, at zero temperature $T = 0$. Remarkably when $\gamma = 2.01$, the error rate is almost zero when the number of stored patterns p is small, and gradually increases but is less than 0.3 even when p is increased to the total number of neurons N. This implies that the storage capacity becomes tremendously enhanced as the SF network becomes extremely heterogeneous in structure, but there occur some errors. We remark that the previous solution of the Hopfield model on fully connected networks [6, 7] revealed that the error rate is zero up to a certain threshold, but beyond which it becomes one-half and the system falls into a total confusion state. Based on these results, we think that the solution of the Hopfield model on SF networks reflects more a normal brain in the point that it provides the case of imperfect memory retrieval with some error even when its storage capacity is small. Moreover, the result is timely in accord with a recent experimental discovery that storage capacity of brain is in the petabyte range, as much as entire web, ten times more than previously thought [22]. We hope that our result will provide some theoretical development for modeling associative memory networks in neuroscience.

Results

Model development

Hopfield model on a given network. In the Hopfield model, each neuron at node i of a given neural network (denoted as G) has an Ising spin S_i with two states, $S_i = +1$ and $S_i = -1$ representing excited and rest states for transmitting or not transmitting a signal, respectively [5]. The Hamiltonian (corresponding to energy of the system) is introduced as

$$\mathcal{H} = -\sum_{(i,j)\in G} J_{ij} S_i S_j, \tag{1}$$

where the coupling strength J_{ij} between two connected nodes $(i, j) \in G$, known as the synapse
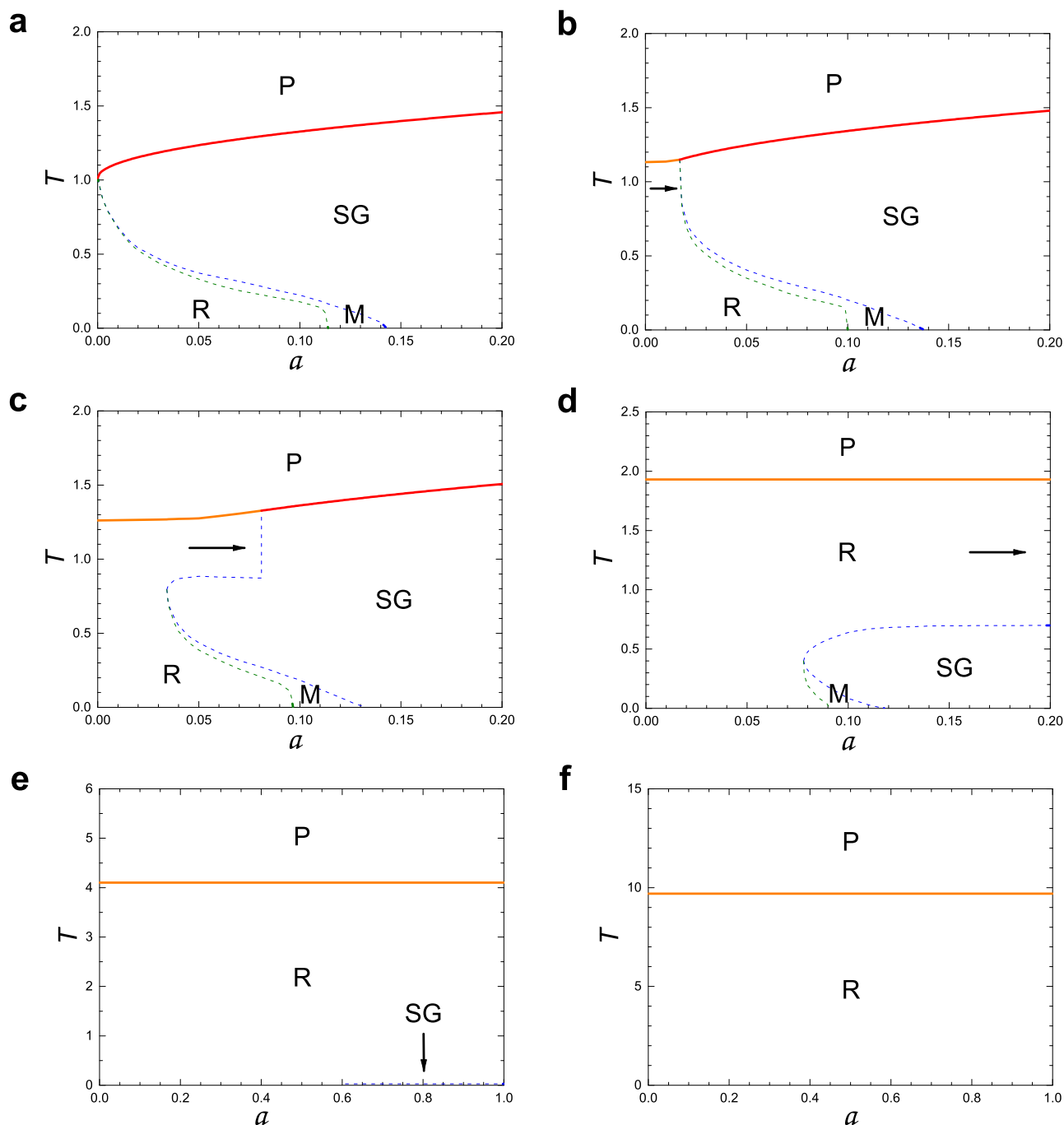
Fig 1. Phase diagram of the Hopfield model in the plane of (T, a). Here T and a denote temperature and storage rate, respectively. Degree exponent γ is infinity in **a**, 5.0 in **b**, 4.0 in **c**, 3.0 in **d**, 2.04 in **e**, and 2.01 in **f**. **P** represents the paramagnetic phase, in which $m = 0$, $q = 0$, and $r = 0$ because of thermal fluctuations. Here, m, q, and r are given by Eqs (29-31) of the S1 File, respectively. **SG** does the spin-glass phase, in which $m = 0$, $q > 0$, and $r > 0$. In the P and SG phases, the retrieval of stored patterns is impossible. Thus, they are often referred to as the confusion phase. The retrieval phase is denoted as **R**, in which $m > 0$, $q > 0$, and $r > 0$. The retrieval of stored memory is possible. Finally, **M** does the mixed phase, in which the features of both the retrieval and the spin-glass phases coexist. As the degree exponent γ is decreased from infinity in **a** through $\gamma = 2.01$ in **f**, the retrieval phase not only intrudes into the region of the SG phase, but also raises the boundary of the phase **P** to a higher temperature region. Eventually the SG phase remains on the $T = 0$ axis when $\gamma = \gamma_c \simeq 2.04$, in which the phase **R** spans most of the low-temperature region. Thus, memory retrieval is enhanced. The phase boundary was obtained by performing numerical calculations for the Chung-Lu SF networks with the system size $N = 1000$ and mean degree $K = 5.0$. Solid and dotted lines or curves indicate the second-order and first-order transitions, respectively. We note that the case **a** on ER network is nearly the same as that in mean field limit obtained in the original Hopfield model.

https://doi.org/10.1371/journal.pone.0184683.g001
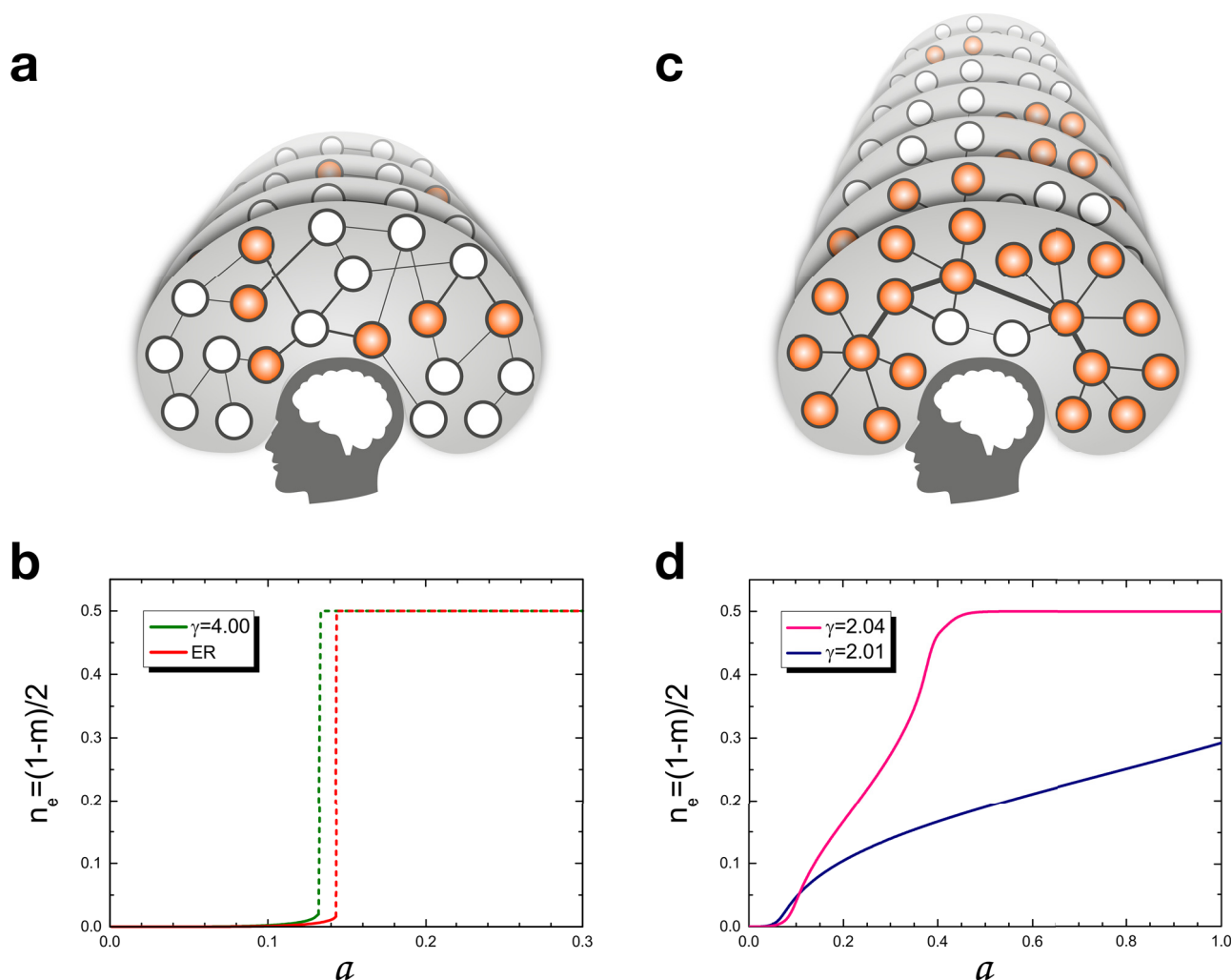
Fig 2. Conceptual figures of the storage capacities and the error rates. a for an ER random network and **c** for a SF network. **b** and **d**
Plot of the error rate $n_e \equiv (1 - m)/2$ vs storage rate a for several γ values of the Chung-Lu model at $T = 0$. Here, numerical values are
obtained using $N = 1000$ and $K = 5.0$. The dotted lines for $\gamma \gg 2.0$ indicate the sudden jumps from small error rates to the state of $n_e = 0.5$. (**a**
and **c**, Figure courtesy of Joonwon Lee.)

https://doi.org/10.1371/journal.pone.0184683.g002

efficacy, takes the Hebbian form,

$$J_{ij} = \frac{1}{K} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu. \qquad (2)$$

K is the mean degree, i.e., the average number of edges of the network G of size N. ξ_i^μ is another
quantity assigned to node i, which also has either +1 or −1. A collective quantity $\{\xi_i^\mu\}$ repre-
sents a memory pattern denoted by μ that is stored in the system. The index μ runs $\mu = 1, \ldots,$
p, which means that the number of memory patterns stored is p. Whereas ξ_i^μ is fixed through-
out the dynamics. Starting from some initial values of $\{S_i(t = 0)\}$, the state of each spin is
updated asynchronously as

$$S_i(t+1) = \mathrm{sgn}\left(\sum_j J_{ij} S_j(t) \right). \qquad (3)$$

When $\Sigma_j J_{ij} S_j(t)$ becomes zero, $S_i(t + 1) = +1$ is assigned definitely.

As the updating is repeated, the energy given by (1) reduces and the system reaches a local minumum state. In this state, each spin S_i becomes equivalent to ξ_i^μ for a given pattern μ, and the stored memory is retrieved. The underlying mechanism for such converging behavior is explained in Section 2 of the S1 File. In real-world systems, however, the repeated dynamics may not be deterministic as Eq 3, but it may include some noise. To take into account of noise effect, the original study [6] of Hopfield model invoked the formalism developed in equilibrium statistical mechanics at finite temperature, in which temperature represents noise strength.

Ensemble average of the Hopfield model over different network configurations and stored patterns.   Thus far, we consider the Hopfield model on a given network. However, connection profiles of SF networks can be different from sample to sample even though they follow the same degree distribution. Thus, we need to take average of physical quantities over the ensemble of different network configurations. To proceed, we consider the probability that a given SF network G exists in the ensemble. That is given as

$$P_K(G) = \prod_{(i,j)\in G} f_{ij} \prod_{(i,j)\notin G} (1 - f_{ij}),$$

(4)

where f_{ij} is the probability to connect a link between two nodes i and j. It was derived that $f_{ij} = 1 - \exp(-NKw_iw_j)$ [17, 18]. The factor w_i is a weight of node i reflecting heterogeneous degrees of a SF network. The explicit form of the weight factor is presented in the Method section. Then the ensemble average over different networks for any given physical quantity A is taken as

$$\langle A \rangle_K = \sum_G P_K(G)A(G),$$

(5)

where $\langle \cdots \rangle_K$ denotes the average over different graph configurations. $A(G)$ represents any physical quantity obtained in a graph G.

Next, we consider a situation in which stored patterns are not deterministic, but stochastically generated. This case is considered for the purpose of testing the efficiency of the algoritumc. Specifically, each pattern μ is created with the population of $\xi_i^\mu = 1$ with probability 1/2 and that of $\xi_i^\mu = -1$ with probability 1/2 for each node i. Then, the probability that a pattern μ is created is given as

$$P(\{\xi_i^\mu\}) = \prod_{i=1}^{N} \left[\frac{1}{2}\delta(\xi_i^\mu - 1) + \frac{1}{2}\delta(\xi_i^\mu + 1) \right].$$

(6)

We take the average of any given physical quantity A over the ensemble of different stored patterns as

$$\langle A \rangle_\xi = \int d\xi_i^\mu P(\{\xi_i^\mu\})A(\{\xi_i^\mu\}),$$

(7)

where $\langle \cdots \rangle_\xi$ is an average over the quenched disorder of ξ_i^μ.

Analytic solutions

The order parameters.   We characterize the phases of the Hopfield model by three quantities, which are often called the order parameters in statistical physics as follows: i) The overlap parameter defined as $m_\alpha^\mu \equiv \sum_i^N w_i \langle \xi_i^\mu S_i^\alpha \rangle$, which represents the extent of which the μ-th pattern of memory ξ^μ and the α-th state of the system S^α overlap with each other. Thus, this quantity

measures the retrieval success rate of the μ-th memory. We remark that the factor w_i is required to take into account heterogeneous degrees. w_i is proportional to the degree k_i of neuron i [23–26]. Actually this weight w_i is the same as that used in the Chung-Lu model introduced in Methods. ii) The so-called spin glass order parameter $q_{\alpha\beta} \equiv \sum_i^N w_i \langle S_i^\alpha S_i^\beta \rangle$ representing the extent of which the two states α and β of the replica overlap each other. The replica is a quantity introduced in the spin-glass theory to resolve the technical difficulties in taking the ensemble average $\langle \cdots \rangle_K$ and $\langle \cdots \rangle_\xi$ introduced in Eqs (5) and (7). iii) A new quantity is introduced as $r_{\alpha\beta} \equiv (N/p) \sum_{\mu=2}^p m_\alpha^\mu m_\beta^\mu$. This quantity is necessary to derive the first two order parameters, and is interpreted as the sum of the effects of each non-retrieval pattern.

Next, we take the replica-symmetric solution by setting $q_{\alpha\beta} = q$ and $r_{\alpha\beta} = r$ for all $\alpha \neq \beta$ and $m_\alpha^1 = m$ for all α. Then the order parameters m, q and r at zero temperature are obtained as

$$m = \sum_{i=1}^N w_i \ \mathrm{erf}\left(\sqrt{\frac{Nw_i}{2ar}} m \right) \tag{8}$$

$$q = 1 - \frac{1}{NK} \sum_{i=1}^N \sqrt{\frac{2Nw_i}{\pi ar}} \ \exp\left(-\frac{Nw_i}{2ar} m^2 \right) \tag{9}$$

$$r = \frac{q}{(1 - K + Kq)^2}, \tag{10}$$

which are used for obtaining the error rate. Here, storage rate a is defined as the number of existing memory patterns p divided by the total number of neurons in a given network, i.e., the network size N. Therefore, the storage capacity a_c is the maximum value of storage rate a. Detailed calculations for those parameters are presented in Section 4 of the S1 File.

Phase diagram and error rates.   The phase diagram we obtained are shown in Fig 1 in the $T - a$ plane for different degree exponents γ but with fixed $N = 1000$ and $K = 5.0$. First, in Fig 1(a), we consider the case of the Erdős and Rényi (ER) network, equivalent to the limit $\gamma \to \infty$. This phase diagram is nearly the same as that obtained on fully connected network in Ref. [7]. There exist four different phases: i) The paramagnetic phase denoted as P, in which $m = 0$, $q = 0$, and $r = 0$ because of thermal fluctuations. ii) The spin-glass phase denoted as SG, in which $m = 0$, $q > 0$, and $r > 0$. In those phases P and SG, the retrieval of stored patterns is impossible. Thus, they are referred to as the confusion phases. iii) The retrieval phase denoted as R, in which $m > 0$, $q > 0$, and $r > 0$, and in which the retrieval of stored memory is possible. Finally, the mixed phase denoted as M, in which the properties of both the retrieval and the spin-glass phases coexist.

The free energy of the system in each phase has also been investigated in Ref. [7]. In phase R, a stored pattern, for instance, $\{\xi_i^\mu\}_i$ is matched to a state of the system $\{S_i^\alpha\}_i$ at the global minimum of free energy. However, in phase M, a stored pattern is matched to a state of the system in a metastable state, while the SG state lies at the global minimum.

For $a = 0$, which occurs when $p \sim \mathcal{O}(N)$ in the limit $N \to \infty$, there exist two phases, P and R. As a is increased slightly from $a = 0$, the SG and M phases appear between the phases P and R. The transition between the P and the SG phases is a second-order transition, whereas that between SG and M is a first-order transition. Likewise, the transition between M and R is also first order. At $T = 0$, the transition between R and M occurs at $a_m \simeq 0.114$, and the transition between M and SG occurs at $a_c \simeq 0.143$. Therefore, as temperature is lowered from a sufficiently high value, successive transitions occur following the steps P $\to$ SG $\to$ M $\to$ R for

$a < a_m$. For $a_m < a < a_c$, successive transitions occur following the steps P → SG → M, and for $a > a_c$, a transition from P → SG occurs.

Second, we obtain the phase diagram for finite γ values in Fig 1(b)–1(e). That undergoes drastic changes depending on γ. Remarkably, the R phase intrudes into the region of the SG phase, but it also raises the boundary of the P phase to a higher-temperature region. As γ approaches 2.04 in Fig 1(e), we observe that the R phase prevails, whereas the SG phase shrinks until it only exists at $T = 0$ in the region $a > 0.6$. Such changes in the $T − a$ diagrams can be understood analytically by examining the phase boundary between the P and SG phases. The details are presented in Section 5 of the S1 File.

Third, we consider a particular case, the noiseless case $T = 0$ in Fig 2. The order parameters m, q and r are given in Eqs (8–10). The figures show the behavior of the error rates n_e as a function of a with $N = 1000$ and $K = 5.0$. The error rate means the relative error of the neural networks and is defined as $n_e \equiv (1 − m)/2$. This figure is obtained analytically for various degree exponent values including $\gamma \simeq 2.0$. We first consider the case in which $\gamma \to \infty$, i.e., the ER limit. The result of this case is almost the same as that obtained in Ref. [7]. When a is less than $a_c \simeq 0.143$, n_e is very small, such that the error rate is negligible and the system is almost in the error-free state. The obtained value a_c is close to $a_c \simeq 0.138$, which was obtained on the fully connected network in Ref. [7]. As a reaches a_c, n_e suddenly jumps to 0.5 as shown in Fig 2(b). This means that for $a > a_c$, the system is in the error-full state (i.e., the state of complete confusion). As γ decreases, this behavior persists up to $\gamma_c \simeq 2.7$, and no longer holds for $\gamma \le \gamma_c$. Note that the value of a_c slightly decreases with decreasing γ, but their dependences are almost negligible.

Next, when γ approaches 2.0, n_e noticeably changes from the step-function-like shape to a monotonously increasing one as shown in Fig 2(d). As γ is lowered further and approaches 2.0, the range of a for the state of complete confusion, with $n_e = 0.5$, disappears and a_c cannot be defined anymore. For instance, at $\gamma = 2.01$, the error rate becomes less than 0.3 for the entire range of a. Therefore, when γ is lowered to as small as 2.0, while the range of a for the state of $n_e = 0$ (the state of perfect retrieval without error) is reduced, the range of a for the state of $n_e = 0.5$, the region of the state of complete confusion disappears. The details are presented in Section 6 of the S1 File. These behaviors have never been observed yet in previous studies [15, 16, 27]. Fig 2 thus implies that as the network changes from the ER network to the extremely heterogeneous SF network with degree exponent two, the storage capacity becomes tremendously enhanced, but some error occurs. This result suggests that hubs play central roles in memory retrieval.

Conclusion and discussion

We obtained the results that as the network changes from a hub-absent network to a SF network with degree exponent just above two, the storage capacity becomes tremendously enhanced, but some error occurs. These features seem to be in accordance with what we experience in everyday life and with a recent discovery of enormously high storage capability in the human brain. Thus hubs, i.e., neurons with a large number of synapses, and other nodes with heterogeneous degrees in neural networks play a central role in enhancing the storage capacity. It is interesting that a normal human brain has such a structure, even though detailed structural properties such as modularity and degree correlation are not yet known. We consider the effect of degree correlation near $\gamma = 2.0$ by performing a similar analysis of the static model [23–26], the degree correlation of which is disassortative between $2 < \gamma < 3$. We find that correlated SF networks near $\gamma = 2.0^+$ have lower values of n_e than the uncorrelated ones. The details are presented in Section 8 of the S1 File. We also checked the error rate n_e on several

real neural networks, having heavy-tailed degree distributions but with undetermined degree exponents due to the small network size. The obtained patterns of the error rates are indeed similar to the theoretical prediction. The details are presented in Section 9 of the S1 File.

Our results might provide some guidelines for constructing an artificial neural network, provided that it is constructed on the basis of the Hopfield model: If we want to construct an artificial neural network that is capable of perfect memory recall with a large value of a_c, as a basic topology of the artificial neural network, it may be more appropriate to choose an ER-type random network or a fully connected one. However, if we prefer to construct an artificial neural network that supports an extended range of storage rate and tolerate a small range of errors, it may be more appropriate to choose an SF-type network with the degree exponent $\gamma \simeq 2.0$. In this case, we can construct an artificial neural network with relatively low cost compared with a fully-connected one which needs very many synapses of $O(N^2)$.

Methods

Construction of scale-free networks

To construct uncorrelated SF networks, we use the Chung-Lu (CL) model [17, 18]: We start with a fixed number of N vertices. Each vertex i ($i = 1, 2, \ldots, N$) is assigned a weight

$$w_i = \frac{(i + i_0 - 1)^{-\nu}}{\sum_{j=1}^{N} (j + i_0 - 1)^{-\nu}}, \tag{11}$$

where ν is a control parameter in the range [0, 1), and i_0 is constant given by

$$i_0 = \begin{cases} [10\sqrt{2}(1 - \nu)]^{1/\nu} N^{1-1/2\nu} & (1/2 < \nu < 1), \\ 1 & (0 \leq \nu < 1/2). \end{cases} \tag{12}$$

A pair of vertices (i, j) is chosen with the probabilities w_i and w_j, respectively, and they are connected with an edge, unless the pair is already connected. This process is repeated $NK/2$ times. Then, the resulting network becomes an uncorrelated SF network following a power-law degree distribution, $P_d(k) \sim k^{-\gamma}$, where k denotes the degree and γ is the degree exponent with $\gamma = 1 + 1/\nu$. In such random networks, the probability that a given pair of vertices (i, j) ($i \neq j$) is not connected by an edge, as denoted by $1 - f_{ij}$, is given by $(1 - 2w_i w_j)^{NK/2} \simeq \exp(-NKw_i w_j)$, while the connection probability $f_{ij} = 1 - \exp(-NKw_i w_j)$.

As a particular case, when we choose i_0 as 1 for all the values of ν, the CL model reduces to the static model [23–26], which has correlations for the range of $1/2 < \nu < 1$ [23, 24]. Therefore, the weights for the static model are given by

$$w_i = \frac{i^{-\nu}}{\zeta_N(\nu)} \tag{13}$$

where ν is a control parameter in the range [0, 1), and $\zeta_N(\nu) \equiv \sum_{j=1}^{N} j^{-\nu} \simeq N^{1-\nu}/(1 - \nu)$. Note that $f_{ij} \simeq NKw_i w_j$ for finite K, however, $f_{ij} \simeq 1$ for $2 < \gamma < 3$ and $ij \ll N^{3-\gamma}$.

For the Erdős-Rényi (ER) graph [28–30], ν becomes 0 and the weights of both models become $w_i = 1/N$, independent of the index i. Since $w_i w_j = 1/N^2$, the fraction of bonds present becomes $f_{ij} \simeq K/N$ and the total number of the connected edges L is $NK/2$. So K becomes the mean degree in the ER graph.

When K approaches N, this network becomes a fully-connected one or a regular lattice with infinite-range interaction. Note that previous studies on the Hopfield model focused mainly on such extreme cases of $K \to N$ [2, 6, 7, 31, 32].

Supporting information

S1 File. Supporting information for the main paper. This supporting information contains the detailed calculations of the free energy and order parameters using the replica analysis.
(PDF)

S1 Fig. A simple example of simulation of the Hopfield model. (a) A network of size $N = 9$ and mean degree $K = \frac{5}{3}$. (b) Two patterns ξ^1 and ξ^2 are encoded by the Hebbian rule. In (c) and (d), we start from two random patterns. After some updates using Eq (1) of the S1 File, the patterns ξ^1 and ξ^2 are retrieved.
(EPS)

S2 Fig. Phase diagrams shown in Fig 1 of the main paper are redrawn with the Almeida-Thouless(AT) line (Eq (32) of the S1 File). Note that the dotted black line near zero temperature in each panel represents the AT line. Thus, we check that the replica-symmetric solution is valid over almost the entire region in the plane of (T, a).
(EPS)

S3 Fig. Phase diagrams of the Hopfield model in the plane of (T, N) on the Chung-Lu model of SF networks with (a) $\gamma = 5.00$ and (b) $\gamma = 3.00$. Here, a is fixed as 0.1.
(EPS)

S4 Fig. Phase diagrams of the Hopfield model in the plane of (T, a) on the static model of SF networks with (a) $\gamma = 2.35$ and (b) $\gamma = 2.01$. Those phase diagrams correspond to Fig 1(e) and 1(f) of the main paper for the CL model, respectively. **P** represents paramagnetic phase, **SG** spin glass phase, and **R** retrieval phase. The SG phase remains only on the axis $T = 0$ when $\gamma = \gamma_c \simeq 2.35$, then the R phase spans the entire region of a at the lower temperatures. Thus, the static model shows better memory retrieval than the CL model. To obtain the phase boundary, numerical calculations were performed for the static model with plugging $N = 1000$ and $K = 5.0$ into the formulas. Solid and dotted curves indicate the second-order and the first-order transitions, respectively. Note that black dotted line in each panel near the $T = 0$ line represents the AT line (Eq (32) of the S1 File). Thus, replica-symmetric solution is valid over almost the entire region of the phase space.
(EPS)

S5 Fig. Plot of the error rate $n_e \equiv (1 - m)/2$ vs storage rate a for $\gamma = 2.01$ and 2.35 for the static model at $T = 0$, which corresponds to Fig 2(b) of the main paper for the CL model. Here, numerical values are obtained using $N = 1000$ and $K = 5.0$.
(EPS)

S6 Fig. Plot of the error rate n_e vs storage rate a for (a) $\gamma = 2.01$ and (b) $\gamma = 2.04$ for the CL and the static model at $T = 0$. These figures provide the comparison of the error rate between for the CL model and for the static model. Here, numerical values are obtained using $N = 1000$ and $K = 5.0$.
(EPS)

S7 Fig. Plots of the number of neurons (nodes) with degree larger than k vs degree k of the real neural networks (left column) and of their error rates n_e vs storage rate a (right column): (a) and (b) for the visual cortex network of macaque monkey, with $N = 30$ and $L = 311$. (c) and (d) for the corticocortical connectivity network in the visual and sensorimotor area of macaque monkey, with $N = 47$ and $L = 505$. (e) and (f) for the cortex network of cat, with $N = 52$ and $L = 818$. The red lines of (a), (c) and (e) are guidelines with slope -1 for eye drawn to compare the data points with the scale-freeness of $\gamma = 2.0$. The red curves of (b), (d) and (f)

are drawn to compare the simulation data (•) with the analytic solution (red curve) using Eq (40) of the S1 File under the same conditions of N and L. Here, the γ values we used for w_i in Eq (40) of the S1 File were 2.0035 (b), 2.0050 (d), and 2.0035 (f), respectively.
(EPS)

S1 Data Set. Supporting information for the "S7 Fig". This is another supporting information which contains the data set necessary to replicate the "S7 Fig".
(XLS)

Acknowledgments

Author Contributions

Conceptualization: Do-Hyun Kim.

Data curation: Do-Hyun Kim, Jinha Park.

Formal analysis: Do-Hyun Kim, Jinha Park.

Funding acquisition: Do-Hyun Kim, Byungnam Kahng.

Investigation: Do-Hyun Kim, Jinha Park.

Methodology: Do-Hyun Kim, Jinha Park.

Project administration: Do-Hyun Kim, Byungnam Kahng.

Resources: Do-Hyun Kim, Jinha Park.

Software: Do-Hyun Kim, Jinha Park.

Supervision: Do-Hyun Kim, Byungnam Kahng.

Validation: Do-Hyun Kim, Byungnam Kahng.

Visualization: Do-Hyun Kim, Jinha Park.

Writing – original draft: Do-Hyun Kim, Byungnam Kahng.

Writing – review & editing: Do-Hyun Kim, Jinha Park, Byungnam Kahng.

References

1. Josselyn SA, Köhler S, Frankland PW. Finding the engram. Nat Rev Neurosci. 2015 Aug; 16(9):521–534. https://doi.org/10.1038/nrn4000 PMID: 26289572

2. Amit DJ. Modeling brain function: The world of attractor neural networks. Cambridge University Press; 1989.

3. Hebb DO. The Organization of Behavior. Wiley & Sons; 1949.

4. Kandel ER, Schwartz JH, Jessell TM, Siegelbaum SA, Hudspeth AJ. Principles of Neural Science. vol. 5. 5th ed. McGraw-hill New York; 2013.

5. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci USA. 1982 Apr; 79(8):2554–2558. https://doi.org/10.1073/pnas.79.8.2554 PMID: 6953413

6. Amit DJ, Gutfreund H, Sompolinsky H. Storing infinite numbers of patterns in a spin-glass model of neural networks. Phys Rev Lett. 1985 Sep; 55(14):1530. https://doi.org/10.1103/PhysRevLett.55.1530 PMID: 10031847

7. Amit DJ, Gutfreund H, Sompolinsky H. Statistical mechanics of neural networks near saturation. Ann Phys (NY). 1987 Jan; 173(1):30–67. https://doi.org/10.1016/0003-4916(87)90092-3

8. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009 Mar; 10(3):186–198. https://doi.org/10.1038/nrn2575 PMID: 19190637

9. van den Heuvel MP, Sporns O. Network hubs in the human brain. Trends Cogn Sci. 2013 Dec; 17(12):683–696. https://doi.org/10.1016/j.tics.2013.09.012 PMID: 24231140

10. Barabási AL, Albert R. Emergence of scaling in random networks. Science. 1999 Oct; 286(5439):509–512. https://doi.org/10.1126/science.286.5439.509 PMID: 10521342

11. Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB. Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol. 2005 Mar; 3(3):e68. https://doi.org/10.1371/journal.pbio.0030068 PMID: 15737062

12. Rubinov M, Bullmore E. Schizophrenia and abnormal brain network hubs. Dialogues Clin Neurosci. 2013 Sep; 15(3):339–349. PMID: 24174905

13. Steullet P, Cabungcal J, Monin A, Dwir D, O'Donnell P, Cuenod M, et al. Redox dysregulation, neuroinflammation, and NMDA receptor hypofunction: A "central hub" in schizophrenia pathophysiology? Schizophr Res. 2014 Jul https://doi.org/10.1016/j.schres.2014.06.021 PMID: 25000913

14. Achard S, Delon-Martin C, Vértes PE, Renard F, Schenck M, Schneider F, et al. Hubs of brain functional networks are radically reorganized in comatose patients. Proc Natl Acad Sci USA. 2012 Oct; 109(50):20608–20613. https://doi.org/10.1073/pnas.1208933109 PMID: 23185007

15. Stauffer D, Aharony A, da Fontoura Costa L, Adler J. Efficient Hopfield pattern recognition on a scale-free neural network. Eur Phys J B. 2003 Apr; 32(3):395–399. https://doi.org/10.1140/epjb/e2003-00114-7

16. Torres JJ, Munoz MA, Marro J, Garrido P. Influence of topology on the performance of a neural network. Neurocomputing. 2004 Mar; 58:229–234. https://doi.org/10.1016/j.neucom.2004.01.048

17. Chung F, Lu L. Connected components in random graphs with given expected degree sequences. Ann Comb. 2002 Nov; 6(2):125–145. https://doi.org/10.1007/PL00012580

18. Cho YS, Kim JS, Park J, Kahng B, Kim D. Percolation transitions in scale-free networks under the Achlioptas process. Phys Rev Lett. 2009 Sep; 103(13):135702. https://doi.org/10.1103/PhysRevLett.103.135702 PMID: 19905523

19. Eguiluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV. Scale-free brain functional networks. Phys Rev Lett. 2005 Jan; 94(1):018102. https://doi.org/10.1103/PhysRevLett.94.018102 PMID: 15698136

20. van den Heuvel MP, Stam CJ, Boersma M, Pol HH. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. NeuroImage. 2008 Aug; 43(3):528–539. https://doi.org/10.1016/j.neuroimage.2008.08.010 PMID: 18786642

21. Gallos LK, Makse HA, Sigman M. A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. Proc Natl Acad Sci USA. 2012 Feb; 109(8):2825–2830. https://doi.org/10.1073/pnas.1106612109 PMID: 22308319

22. Bartol TM, Bromer C, Kinney J, Chirillo MA, Bourne JN, Harris KM, et al. Nanoconnectomic upper bound on the variability of synaptic plasticity. eLife. 2015 Nov; 4:e10778. https://doi.org/10.7554/eLife.10778 PMID: 26618907

23. Goh KI, Kahng B, Kim D. Universal behavior of load distribution in scale-free networks. Phys Rev Lett. 2001 Dec; 87(27):278701. https://doi.org/10.1103/PhysRevLett.87.278701 PMID: 11800921

24. Lee JS, Goh KI, Kahng B, Kim D. Intrinsic degree-correlations in the static model of scale-free networks. Eur Phys J B. 2006 Feb; 49(2):231–238. https://doi.org/10.1140/epjb/e2006-00051-y

25. Kim DH, Rodgers G, Kahng B, Kim D. Spin-glass phase transition on scale-free networks. Phys Rev E. 2005 May; 71(5):056115. https://doi.org/10.1103/PhysRevE.71.056115

26. Kim DH. Inverse transitions in a spin-glass model on a scale-free network. Phys Rev E. 2014 Feb; 89(2):022803. https://doi.org/10.1103/PhysRevE.89.022803

27. Castillo IP, Wemmenhove B, Hatchett J, Coolen A, Skantzos N, Nikoletopoulos T. Analytic solution of attractor neural networks on scale-free graphs. J Phys A: Math Gen. 2004 Sep; 37(37):8789. https://doi.org/10.1088/0305-4470/37/37/002

28. Erdös P, Rényi A. On random graphs, I. Publ Math Debr. 1959; 6:290–297.

29. Erdös P, Rényi A. On the evolution of random graphs. Publ Math Inst Hung Acad Sci. 1960; 5(17-61):43.

30. Bollobás B. Random graphs. 2001. Cambridge University Press; 2001.

31. Mézard M, Parisi G, Virasoro MA. Spin glass theory and beyond. World Scientific; 1987.

32. Nishimori H. Statistical physics of spin glasses and information processing: an introduction. Oxford University Press; 2001.