Retrieval capabilities of hierarchical networks: From Dyson to Hopfield

Elena Agliari,[1] Adriano Barra,[1] Andrea Galluzzi,[2] Francesco Guerra,[1] Daniele Tantari,[2] and Flavia Tavani[3]

[1]*Dipartimento di Fisica, Sapienza Università di Roma, P.le A. Moro 2, 00185, Roma, Italy.*
[2]*Dipartimento di Matematica, Sapienza Università di Roma, P.le Aldo Moro 2, 00185, Roma, Italy.*
[3]*Dipartimento SBAI (Ingegneria), Sapienza Università di Roma, Via A. Scarpa 14, 00185, Roma, Italy.*

(Dated: October 1, 2014)

We consider statistical-mechanics models for spin systems built on hierarchical structures, which provide a simple example of non-mean-field framework. We show that the coupling decay with spin distance can give rise to peculiar features and phase diagrams much richer than their mean-field counterpart. In particular, we consider the Dyson model, mimicking ferromagnetism in lattices, and we prove the existence of a number of meta-stabilities, beyond the ordered state, which become stable in the thermodynamic limit. Such a feature is retained when the hierarchical structure is coupled with the Hebb rule for learning, hence mimicking the modular architecture of neurons, and gives rise to an associative network able to perform single pattern retrieval as well as multiple patterns retrieval, depending crucially on the external stimuli and on the rate of interaction decay with distance; however, those emergent multitasking features reduce the network capacity with respect to the mean-field counterpart. The analysis is accomplished through statistical mechanics, Markov chain theory, signal-to-noise ratio technique and numerical simulations in full consistency. Our results shed light on the biological complexity shown by real networks, and suggest future directions for understanding more realistic models.

PACS numbers: 07.05.Mh,87.19.L-,05.20.-y

In the last decade, extensive research on complexity in networks has evidenced (among many results [1, 2]) the widespread of modular structures and the importance of quasi-independent communities in many research areas such as neuroscience [3–5], biochemistry [6] and genetics [7], just to cite a few. In particular, the modular, hierarchical architecture of cortical neural networks has nowadays been analyzed in depths [8], yet the beauty revealed by this investigation is not captured by the statistical mechanics of neural networks, nor standard ones (i.e. performing single pattern retrieval) [9, 10] neither multitasking ones (i.e. performing multiple patterns retrieval) [11, 12]. In fact, these models are intrinsically mean-field, thus lacking a proper definition of metric distance among neurons.

Hierarchical structures have been proposed in the past as (relatively) simple models for ferromagnetic transitions beyond the mean-field scenario -the Dyson hierarchical model (DHM) [13]- and are currently experiencing a renewal interest for understanding glass transitions in finite dimension [14, 15]. Here we investigate their retrieval capabilities when performing as associative networks: we start studying the DHM mixing the Amit-Gutfreund-Sompolinsky ansatz approach [10] (to select eligible retrievable states) with the interpolation technique (to check their thermodynamic stability) and we show that, as soon as ergodicity is broken, beyond the pure ferromagnetic state (largely discussed in the past, see e.g., [16, 17]), a number of metastable states suddenly appear and become *stable* in the thermodynamic limit. The emergence of such states implies the breakdown of classical (mean-field) self-averaging and stems from the weak ties connecting distant neurons, which, in the thermodynamic limit, effectively get split into detached modules (see Fig. 1). As a result, if the latter are initialized with opposite magnetizations, they remain stable.

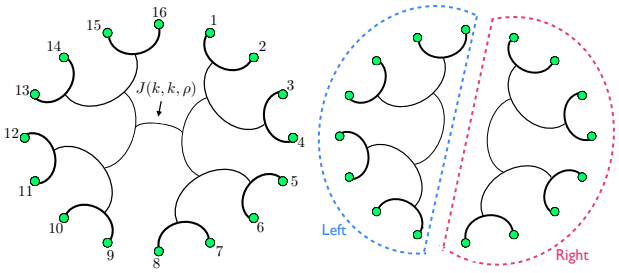

FIG. 1: Sketch of the hierarchical topology underlying the model (here $k = 4$, $N = 2^k = 16$). Nodes represent spins; the larger the distance between two spins and the weaker their coupling. On the right we highlight how the graph breaks down in two equivalent components (referred to as "left" and "right") as the weakest link $J(k, k, \rho)$ (see Eq. 2) is cut.

This is a crucial point because, once implemented the Hebbian prescription [9, 10] to account for multiple pattern storage, it allows proving that the system not only is able to retrieve any single pattern at a time as a standard Hopfield model [9, 10], but its communities can perform autonomously, hence making the simultaneous retrieval of multiple patterns feasible too. We stress that this feature is essentially due to the notion of metric the system is endowed with, differently from the multiple retrieval performed by the mean-field multitasking networks which require blank pattern entries [11, 12].

Therefore, the hierarchical neural network is able to perform both as a serial processor and as a multitasking processor. We corroborate this scenario merging results

from statistical mechanics, Markov chain theory, signal-to-noise ratio technique and extensive numerical simulations as explained hereafter.

In the DHM the mutual interaction between 2^{k+1} Ising spins $\sigma_i = \pm 1$, with $i = 1, ..., 2^{k+1}$, is described by the following Hamiltonian, defined recursively as

$$H_{k+1}(\vec{\sigma}) = H_k(\vec{\sigma_1}) + H_k(\vec{\sigma_2}) - \frac{J}{2^{2\rho(k+1)}} \sum_{i<j=1}^{2^{k+1}} \sigma_i \sigma_j, \quad (1)$$

where $J > 0$ and $\rho \in (1/2, 1[$ tune the interaction strength; $\vec{\sigma_1} \equiv \{\sigma_i\}_{1 \leq i \leq 2^k}$, $\vec{\sigma_2} \equiv \{\sigma_j\}_{2^k+1 \leq j \leq 2^{k+1}}$ and $H_0(\vec{\sigma}) = 0$. This model is explicitly non-mean-field as we implicitly introduced a distance: Two spins i and j turn out to be at distance $d_{ij} = d$ if, along the recursive construction, they first get connected at the d-th iteration; of course d ranges in $[1, k]$ (see also Fig. 1). It is possible to re-write the Hamiltonian (1) in terms of d_{ij} as $H_{k+1}(\vec{\sigma}) = -\sum_{i<j} J_{ij}\sigma_i\sigma_j$, where

$$J_{ij} = \sum_{l=d_{ij}}^{k} \left(\frac{J}{2^{2\rho l}}\right) = J(d_{ij}, k, \rho) = J\frac{4^{\rho - d_{ij}\rho} - 4^{-(k+1)\rho}}{4^{\rho} - 1}. \quad (2)$$

Set the noise level $\beta = 1/T$ in proper units, we are interested in an explicit expression of the infinite volume limit of the mathematical pressure $\alpha(\beta, J, \rho) = -\beta f(\beta, J, \rho)$, (where f is the free energy) defined as

$$\alpha(\beta, J, \rho) = \lim_{k\to\infty} \frac{1}{2^{k+1}} \log \sum_{\vec{\sigma}} \exp[-\beta H_{k+1}(\vec{\sigma})],$$

whose maxima return the equilibrium states of the system. The latter are expressed in terms of the global magnetization $m_{k+1} = \frac{1}{2^{k+1}} \sum_i^{2^{k+1}} \sigma_i$ and of the set of k magnetizations $\vec{m}_1, ..., \vec{m}_k$, which quantify the state of each community, level by level; for instance, the two magnetizations related to the two largest modules (see Fig.1) read off as

$$m_k^{(1)} = m_{\text{left}} = \frac{1}{2^k} \sum_{i=1}^{2^k} \sigma_i, \quad m_k^{(2)} = m_{\text{right}} = \frac{1}{2^k} \sum_{i=2^k+1}^{2^{k+1}} \sigma_i.$$

We approach the investigation of the DHM metastabilities exploiting the interpolative technology introduced in [15], that allows obtaining bounds beyond the mean-field paradigm (as fluctuations are not completely discarded). This procedure, assuming the homogeneity of the sub-magnetizations (i.e., $m_{\text{left}} = m_{\text{right}} = m$), returns the following expression (see [15, 20] for details)

$$\alpha(\beta, J, \rho) \geq \sup_m \left[\log 2 + \log\cosh(\beta J m C_{2\rho}) - \frac{\beta J m^2}{2} C_{2\rho}\right], \quad (3)$$

where $C_y = 2^y/[(2^y - 1)(2^y - 2)]$. However, if we remove the hypothesis that the two main modules display the same magnetization, namely if we call independently

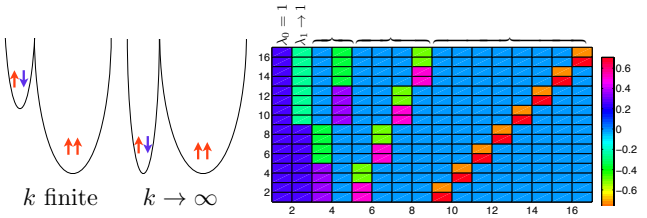

FIG. 2: Left panel: Schematic representation of the free energy for the DHM, where minima correspond to equilibrium states. At finite size the pure ferromagnetic state (denoted by aligned arrows) is a global minimum, while the mixed state (denoted by misaligned arrows) is a local minimum. In the thermodynamic limit both states are global minima and the system can relax to any of them (although the latter displays a smaller attraction basin). Right panel: Representation of the eigenstates of W for a system with $k = 4$ and $\rho = 0.75$. Each column represents a different eigenstate, eigenstates pertaining to the same degenerate eigenvalue are highlighted. Different colors represent different entries in the eigenstate, as shown by the colormap on the right.

$m_{\text{left}} = m_1$ and $m_{\text{right}} = m_2$, under the ansatz of *mixed state* (i.e., $m_{\text{left}} = -m_{\text{right}}$), formula (3) is generalized as

$$\alpha(\beta, J, \rho) \geq \sup_{m_1, m_2} \Big\{ \ln 2 - \frac{\beta J}{2} C_{2\rho} \Big(\frac{m_1^2 + m_2^2}{2}\Big) + \frac{1}{2}[L(\beta m_1 C_{2\rho}) + L(\beta m_2 C_{2\rho})] \Big\}, \quad (4)$$

where $L(x) = \ln\cosh(x)$; of course, posing $m_1 = m_2 = m$, we recover (3). Requiring thermodynamic stability, and using brackets for Boltzmann averages, we obtain the following disentangled self-consistencies

$$\langle m_{1,2} \rangle = \langle \tanh(\beta J m_{1,2} C_{2\rho}) \rangle, \quad (5)$$

whose solution is different from the trivial paramagnetic one (i.e., $\langle m_{1,2} \rangle = 0$) below the critical temperature $T_c = JC_{2\rho}$. The main idea underlying the derivation of Eq. 4, representing the behavior of two quasi-non-interacting sub-communities, is that the interaction term in Eq. 1 is bounded by $2^{-(k+1)(2\rho-1)}$ that is vanishing in the thermodynamic limit [20]. As a consequence, the meta-stable mixed state and the stable ferromagnetic state display an (intensive) energy gap $\Delta E \propto 2^{-(k+1)(2\rho-1)}$ such that, as $k \to \infty$, both states become stable sharing the same intensive free-energy (see Fig. 2, left): as the paramagnetic solution becomes unstable, while thermodynamics is dominated by the ferromagnetic behavior, both the ferromagnetic (i.e. $m_{\text{left}} = m_{\text{right}}$) and the mixed, (i.e. $m_{\text{left}} = -m_{\text{right}}$) solutions appear. We can use this argument iteratively splitting the system in smaller and smaller blocks: after M iterations, we need a set of different magnetizations (one for each 2^M emerging communities) until the $k + 1 - M$ level. The procedure keeps on working as far as the interaction term, bounded by $\sum_{l=k+1-M}^{k+1} 2^{l(1-2\rho)}$, remains vanishing, hence at most $M(k)/k \to 0$. The existence of states different from the

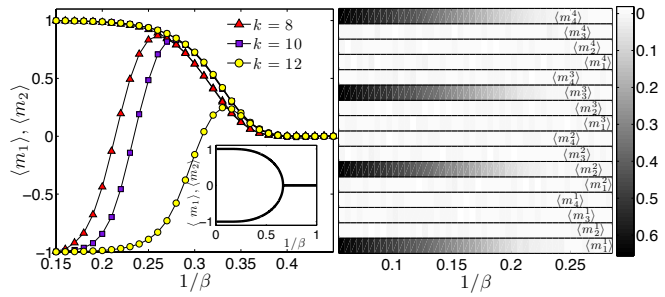

FIG. 3: (Color on line) Left panel: Magnetizations obtained by MC simulations of the DHM for different sizes (main figure) and by the theoretical curves given by Eq. 5 (inset) holding in the thermodynamic limit. In the main plot, the spontaneous switch between serial and mixed state is a finite-size effect. Right panel: Mattis magnetizations obtained by MC simulations of the HHM for $k = 12$ and $p = 4$. Consistently with the analytical picture (Eq. 13 and Eq. 5 give analogous solutions), the simultaneous retrieval of multiple patterns exists and is stable.

purely ferromagnetic one is confirmed by Monte Carlo simulations (see Fig. 3, left).

This point can be further understood from a different perspective: the DHM can be looked at as a ferromagnet embedded in a fully-connected topology, where the link connecting two arbitrary nodes, i and j, displays a weight J_{ij} decaying with the distance between i and j, and defined according to a suitable metric (e.g., the one based on recursion described above or the 2-adic metric $\tilde{d}_{ij} = 2^{-\mathrm{ord}_2(i-j)}$, in such a way that $J_{ij} \sim \tilde{d}_{ij}^{-2\rho}$ [21]). Moreover, the set of nodes is countable and weights are finite, i.e. $J_{\min} = 4^{-(k+1)\rho} \leq J_{ij} \leq J_{\max} = (1 - 4^{-(k+1)\rho})/(4^{\rho} - 1)$, thus, upon proper normalization of weights $J_{ij} \to W_{ij} = J_{ij}/w_i$, where $w_i = \sum_j J_{ij}$, the graph describes a Markov chain, where each node represents a state and W is the transition matrix [18]. The evolution of the random process is therefore provided by the master equation $p(t+1) = Wp(t) \to \dot{p}(t) = Wp(t) - p(t)$, whose stationary distribution, referred to as π, satisfies $\pi = W\pi$, that is, π coincides with the eigenvector ϕ_{λ_0} of W corresponding to eigenvalue $\lambda_0 = 1$ (that is just the Perron-Frobenius eigenvalue of W) and it is uniformly distributed as $\pi = \mathbf{e}/2^{(k+1)/2}$. The second-largest eigenvalue λ_1 and the related eigenstate are, respectively,

$$\lambda_1 = \sum_{j=1}^{2^k} W_{1j} - 2^k W_{12^{k+1}} \to 1 - \mathcal{O}(2^{-(2\rho-1)(k+1)}),$$
$$\phi_{\lambda_1} = (\underbrace{1, 1, \ldots, 1}_{2^{(k+1)/2}}, \underbrace{-1, -1, \ldots, -1}_{2^{(k+1)/2}})/2^{(k+1)/2}.$$

As λ_1 converges to 1 in the thermodynamic limit, ergodicity breaking for the stochastic process is expected. In fact, ϕ_{λ_0} and ϕ_{λ_1} generate a subspace where any vector is an eigenvector of W with the same eigenvalue $\lambda = 1$. In particular, we see that

$$\phi_{\lambda_0} + \phi_{\lambda_1} = (\underbrace{1, 1, \ldots, 1}_{2^{(k+1)/2}}, \underbrace{0, 0, \ldots, 0}_{2^{(k+1)/2}})\sqrt{2/2^{k+1}}, \quad (6)$$

$$\phi_{\lambda_0} - \phi_{\lambda_1} = (\underbrace{0, 0, \ldots, 0}_{2^{(k+1)/2}}, \underbrace{1, 1, \ldots, 1}_{2^{(k+1)/2}})\sqrt{2/2^{k+1}}, \quad (7)$$

correspond to stationary states localized on the left and on the right branch of the graph, respectively. Otherwise stated, there is no flow between the two main modules as if they were autonomous. The same holds as we split each branch in smaller sub-units iteratively (see Fig. 2, right), and mirrors the genesis of metastable states in the thermodynamics counterpart.

As a final perspective, we check the robustness of states through a signal-to-noise ratio analysis. To this aim we express the fields acting on the spins in (1) by writing $H_{k+1}(\vec{\sigma}) = -\sum_i h_i(\vec{\sigma}, \rho, k)\sigma_i$, being

$$h_i(\vec{\sigma}, \rho, k) = \sum_{d=1}^{k+1} \Big[\sum_{l=d}^{k+1} \Big(\frac{J}{2^{2l\rho}} \Big) 2^{d-1} m_{f(d,i)}^{d-1} \Big], \quad (8)$$

where $m_{f(d,i)}^{d-1}$ is the normalized magnetization of spins at distance d from the i-th one. The microscopic law governing the evolution of the system is a stochastic alignment with the local field $h_i(\vec{\sigma}, \rho, k)$, that is, $\sigma_i(t + \delta t) = \mathrm{sign}\{\tanh[\beta h_i(\tilde{\sigma}(t), \rho, k)] + \eta_i(t)\}$. In the noiseless limit, the stochasticity captured by the independent random numbers $\eta_i(t)$ (uniformly distributed over the interval $[-1, +1]$) is lost, and

$$\lim_{\beta \to \infty} \sigma_i(t + \delta t) = \mathrm{sign}\{h_i(\vec{\sigma}(t), \rho, k)\}. \quad (9)$$

Thus, if $\sigma_i h_i(\vec{\sigma}, \rho, k) > 0, \forall i \in [1, 2^{k+1}]$, the configuration $\vec{\sigma}$ is dynamically stable. Hereafter, we focus on the ferromagnetic/single pattern case and on the mixed/multitasking case only, referring again to [21] for an extensive treatment.

In the former case, $\sigma_i = +1, \forall i \in [1, 2^{k+1}] \Rightarrow h_i(\vec{\sigma}, \rho, k) > 0 \ \forall k, \rho \in (1/2, 1]$. Therefore, the ferromagnetic/single pattern case state is stable for $\beta \to \infty$.

In the latter case, $\sigma_i = +1, \forall i \in [1, 2^k]$ and $\sigma_i = -1, \forall i \in [2^k + 1, 2^{k+1}] \Rightarrow \lim_{k\to\infty} h_i(\vec{\sigma}, \rho, k) = 1/(2^{1-2\rho} + 4^{\rho} - 3) > 0 \ \forall \rho \in (1/2, 1]$. Therefore, the mixed/multitasking case is stable for $\beta \to \infty$.

Now, retaining the outlined perspective, we recursively define the hierarchical Hopfield model (HHM) by the following Hamiltonian

$$H_{k+1}(\vec{\sigma}) = H_k(\vec{\sigma_1}) + H_k(\vec{\sigma_2}) - \frac{1}{2} \frac{1}{2^{2\rho(k+1)}} \sum_{\mu=1}^{p} \sum_{i,j=1}^{2^{k+1}} \xi_i^{\mu} \xi_j^{\mu} \sigma_i \sigma_j \quad (10)$$

with $H_0(\vec{\sigma}) = 0$, $\rho \in (1/2, 1[$; beyond 2^{k+1} dichotomic neurons σ_i, also p quenched patterns ξ^{μ}, $\mu \in (1, ..., p)$ are introduced. Their entries $\xi_i^{\mu} = \pm 1$ are drawn with the same probability 1/2 and are averaged by $\mathbb{E}_{\xi}$.

Again, we can write the Hamiltonian of the HHM in terms of the distance d_{ij}, obtaining $H_{k+1}(\vec{\sigma}) = -\sum_{i<j} \widetilde{J_{ij}}\sigma_i\sigma_j$, where

$$\widetilde{J_{ij}} = \frac{4^{\rho - d_{ij}\rho} - 4^{-k\rho}}{4^{\rho}-1} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \tag{11}$$

hence the Hebbian kernel $\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$ is tuned by the distance-dependent weight $J(d_{ij}, k, \rho)$.

Once introduced suitably Mattis overlaps, both global $m_\mu = \sum_{i=1}^{2^{k+1}} \xi_i^\mu \sigma_i / 2^{k+1}$, and community restricted, as

$$m^\mu_{\text{left}} = \frac{1}{2^k}\sum_{i=1}^{2^k} \xi_i^\mu \sigma_i, \; m^\mu_{\text{right}} = \frac{1}{2^k}\sum_{j=2^k+1}^{2^{k+1}} \xi_j^\mu \sigma_j, \tag{12}$$

the statistical-mechanical route, as in the Dyson model, returns a non-mean field approximation for the pressure of the single pattern retrieval state as $\alpha(\beta, p, \rho) \sim \sup_m \{\log 2 - \frac{\beta}{2}\sum_{\mu=1}^{p} m_\mu^2 C_{2\rho} + \mathbb{E}_\xi \log\cosh[\sum_{\mu=1}^{p}(\beta m_\mu C_{2\rho})\xi^\mu]\}$. Assuming two different families of Mattis magnetizations $\{m^\mu_{1,2}\}_{\mu=1}^{p}$ for the two largest communities (left and right), we get a non-mean-field approximation for the multiple patterns retrieval pressure

$$\alpha(\beta,p,\rho) \sim \sup_{\{m^\mu_{1,2}\}} \{\ln 2 - \frac{\beta}{2} C_{2\rho} \sum_{\mu=1}^{p} \frac{(m_1^\mu)^2 + (m_2^\mu)^2}{2}$$
$$+ \frac{1}{2}\mathbb{E}_\xi [L\big(\sum_{\mu=1}^{p}(\beta m_1^\mu C_{2\rho} + \xi^\mu)\big) + L\big(\sum_{\mu=1}^{p}(\beta m_2^\mu C_{2\rho})\big)]\},$$

whose disentangled optimal order parameters satisfy

$$\langle m^\mu_{1,2}\rangle = \mathbb{E}_\xi \langle \xi^\mu \tanh[\beta \sum_{\nu=1}^{p} C_{2\rho} m^\nu_{1,2} \xi^\nu]\rangle, \tag{13}$$

returning again $T_c = C_{2\rho}$. The behavior of $\langle m^\mu_{1,2}\rangle$ resulting from Eq. 13 is consistent with the simulation outcomes (see Fig. 3, right panel). Again, we can iteratively repeat the procedure as far as the interaction term among the sub-communities remains vanishing, hence getting $M(k) = o(k)$. Therefore, if we want the system to handle p patterns simultaneously, we need p blocks and this, for $k \to \infty$, establishes the bound $p = o(2^k)$.

This picture is confirmed by the signal-to-noise ratio analysis: we start from the single pattern state, i.e., $\sigma_i = \xi_i^\mu$, and check its stability writing $\sigma_i h_i(\vec{\sigma}, \rho, k)$ as a signal term S plus a noise term R and then comparing their amplitudes:

$$\xi_i^\mu h_i(\vec{\sigma}, \rho, k) = \xi_i^\mu \sum_{\nu=1}^{p} \xi_i^\nu \sum_{d=1}^{k} J(d,k,\rho) \sum_{j:d_{ij}=d} \xi_j^\nu \xi_j^\mu = S + R(\xi), \tag{14}$$

where $S = \sum_{d=1}^{k} J(d,k,\rho) 2^{d-1} \geq 0$, while

$$R(\xi) = \xi_i^\mu \sum_{\nu\neq\mu}^{p} \xi_i^\nu \sum_{d=1}^{k} J(d,k,\rho) \sum_{j:d_{ij}=d} \xi_j^\nu \xi_j^\mu.$$

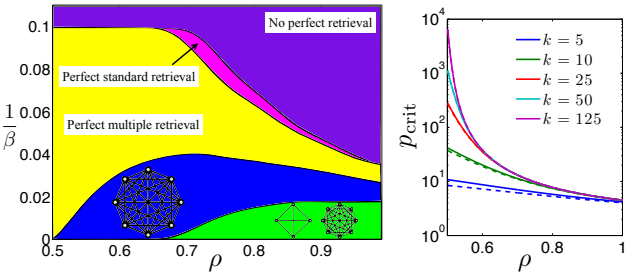

FIG. 4: Left panel: Phase diagram for the DHM as derived from the signal-to-noise ratio analysis. The curves separating different phases are obtained by solving numerically the transcendental equation $\tanh[\beta h_i(\vec{\sigma}, \rho, k)] \approx 1$ as a function of β and ρ. Here we fixed $k = 7$ and we focused on four different configurations (single pattern state, multitasking state and states where sub-communities made of four and eight spins, respectively, are misaligned with respect to the bulk). Right panel: $p^{\text{single}}_{\text{crit}}$ (solid line) and $p^{\text{multi}}_{\text{crit}}$ (dashed line), as a function of ρ and for several choices of k, as explained by the legend. Notice that as ρ increases the capacity of the system (for both single and multiple retrieval) decreases, in fact, the magnitude of couplings decreases exponentially with ρ, hence the overall storage capacity is also reduced [19, 20].

As clearly $\langle R(\xi)\rangle_\xi = 0$, we need to evaluate when the ratio $S/|\sqrt{\langle R(\xi)^2\rangle_\xi} \to 1$: the latter returns the maximum load $p^{\text{single}}_{\text{crit}}(k,\rho)$ storable by the network before the noise prevails over the signal and retrieval fails.
As for the stability of the multiple patterns retrieval, forcing $\sigma_i = \xi_i^\mu \; \forall i \in [1, 2^k]$ and $\sigma_i = \xi_i^\gamma \; \forall i \in [2^k+1, 2^{k+1}]$ for $\mu \neq \gamma$, and splitting again $\sigma_i h_i(\vec{\sigma}, \rho, k)$ in a signal plus a noise term, we can check again the maximum load $p^{\text{multi}}_{\text{crit}}(k,\rho)$ storable by the network. Both $p^{\text{single}}_{\text{crit}}(k,\rho)$ and $p^{\text{multi}}_{\text{crit}}(k,\rho)$ are monotonically decreasing functions of ρ, and they converge to the finite value $(4^\rho - 1)(4^{2\rho} - 2)/(4^{2\rho} - 3 \times 4^\rho + 2)^2 + 1$ as k gets larger (see also Fig. 4 and [21]).

Summarizing, beyond classical retrieval, the network is able to safely handle multiple patterns too. Unlike multitasking mean-field networks [11] (where multiple retrievals are independent), here the weak ties connecting different modules may account also for correlated retrieval (as the retrieved pattern of one module may drive retrieval in other modules), better mimicking the behavior of biological systems. However, there is a cost in terms of capacity: as we can neglect links among upper levels - those that become effectively negligible in the thermodynamic limit - globally the network loses a significant amount of bit-storing synapses. Thus a new compromise appears in non-mean-field cognitive systems: increasing multitasking capabilities diminishes the processor capacity, the trigger between them being ruled mainly by the rate of interaction decay ρ.

Sapienza University, GNFM-INdAM and INFN are acknowledged for financial support.

[1] M. Newman, *Networks*, Oxford University Press (2010).
[2] A. Vespignani, G. Caldarelli, *Large Scale Structure and Dynamics of Complex Networks*, World Sci. Publ. (2007).
[3] E. Bullmore, O. Sporns, *Nat. Rev. Neurosci.* **13**, 366 (2012).
[4] A. Kumar, et al., *Trends Neurosci.* **36**, 579 (2013).
[5] M. Zhao, et al., *Phys. Rev. E* **84**, 016109 (2011).
[6] L. K. Gallos, et al., *Proc. Natl. A. Soc.* **104**, 7746 (2007).
[7] C. Conaco, et al., *Proc. Natl. A. Soc.* **109**, 10612 (2012).
[8] P. Moretti, M.A. Munoz, *Nature Comm.* **4**, 2521 (2013).
[9] J.J. Hopfield, *Proc. Natl. Acad. Sci.* **79**, 2554 (1982).
[10] D.J. Amit, *Modeling brain function*, Cambridge University Press (1992).
[11] E. Agliari, et al., *Phys. Rev. Lett.* **109**, 268101 (2012).
[12] P. Sollich, et al., submitted, available at arXiv:1404.3654.
[13] F.J. Dyson, *Comm. Math. Phys.*, **12**, 91 (1969).
[14] M. Castellana, et al., *Phys. Rev. Lett.* **104**, 127206 (2010).
[15] M. Castellana, et al., *J. Stat. Phys.* **155**, 211, (2014).
[16] G. Gallavotti, S. Miracle-Sole, *Comm. Math. Phys.*, **5**, 317 (1967).
[17] P.M. Bleher, J.G. Sinai, *Comm. Math. Phys.* **33**, 23 (1973).
[18] J.R. Norris, *Markov Chains*, Cambridge Press, (1994).
[19] H. Sompolinsky, *Phys. Rev. A* **34**, 2571 (1986)
[20] E. Agliari, et al., available at arXiv:1407.5176.
[21] E. Agliari, et al., available at arXiv:1409.0227.