Journal of Statistical Physics (2026) 193:116
https://doi.org/10.1007/s10955-026-03679-y

# Dense Associative Memory with Biased Patterns: A Replica Symmetric Analysis

**Linda Albanese**[1,2] · **Andrea Alessandrelli**[1,2] · **Federico Carella**[3,4]

Received: 10 April 2026 / Accepted: 31 July 2026 / Published online: 20 August 2026
© The Author(s) 2026

**Abstract**

We investigate dense higher-order associative memories in the high storage regime when the stored patterns are biased, namely when the entries of the patterns are not symmetrically distributed around zero. In this setting, the standard Hebbian prescription must be modified by recentering and rescaling the pattern entries, and an additional term must be introduced in the Hamiltonian to enforce consistency between the average activity of the network and that of the stored patterns. As a first step, we perform a signal-to-noise analysis in the zero-temperature limit and show that the bias reduces the effective storage capacity through a multiplicative correction factor $(1-b^2)^P$, while preserving the superlinear scaling with the system size. We then derive the quenched statistical pressure within the Replica Symmetric framework by means of Guerra's interpolation method and obtain the corresponding self consistency equations for the relevant order parameters. The analytical treatment confirms the heuristic prediction of the signal-to-noise argument, taking to an attenuation of the signal or equivalent renormalization of the effective noise. Finally, we discuss the resulting phase behavior of the model and its implications for retrieval performance in the model.

**Keywords** Associative memory · Biased patterns · Mathematical methods · Guerra interpolation

Communicated by Chiara Cammarota.

Andrea Alessandrelli and Federico Carella have contributed equally to this work.

✉ Linda Albanese
linda.albanese@unisalento.it

Andrea Alessandrelli
andrea.alessandrelli@unisalento.it

Federico Carella
federico.carella2@unibo.it

[1] Dipartimento di Matematica e Fisica "Ennio De Giorgi", Università del Salento, Collegio Fiorini, 73100 Lecce, Italy

[2] Sezione di Lecce, Istituto Nazionale di Fisica Nucleare, Collegio Fiorini, 73100 Lecce, Italy

[3] Fondazione CMCC - Centro Euro-Mediterraneo sui Cambiamenti Climatici, Via Marco Biagi 5, 73100 Lecce, Italy

[4] Dipartimento di Fisica e Astronomia "Augusto Righi" - DIFA, Università di Bologna, Via Zamboni 33, 40126 Bologna, Italy

## 1 Introduction

Associative memory models constitute one of the most paradigmatic frameworks for understanding collective information processing in neural networks and disordered systems. Since Hopfield's seminal contribution [21], these models have provided a simple yet powerful account of how a network of interacting binary units can store a collection of patterns as stable attractors of the dynamics and retrieve them from partial or corrupted inputs. Beyond their original neuroscientific motivation, associative memories have come to play a central role in statistical mechanics [13, 21], machine learning [30], and the study of complex systems [2, 18, 29], where they serve as a natural bridge between spin glass methodologies and information processing [27, 28].

A significant line of research in this area focuses on enhancing storage and retrieval performance by introducing higher-order interactions. In dense $P-$spin generalisations of the Hopfield model [9, 23, 24], standard pairwise Hebbian couplings are replaced by interaction tensors of a higher order, enabling the network to encode correlations between multiple points among the stored patterns. Such architectures are known to display a relevant increase in storage capacity [15]. Whereas the classical Hopfield model can only store a number of patterns that scales linearly with system size $N$, dense higher-order models can operate in a high-storage regime, where the number of stored patterns $K$ increases as $K \sim O(N^{P-1})$ [15, 17]. This makes them a natural framework for investigating associative memory beyond the conventional low storage regime [5, 6, 10, 19].

However, most of the existing Literature assumes that the stored patterns are unbiased, meaning their entries are independent, identically distributed and symmetric Rademacher variables [13, 21]. Under this assumption, the standard Hebbian learning rule is naturally centred and the retrieval problem can be analysed using a well-established theoretical framework [19]. Anyhow, in many realistic scenarios, the statistics of the stored information are not balanced [4, 14, 25]. Patterns may exhibit non-zero mean activity, meaning the probabilities of entries taking $+1$ or $-1$ values are no longer equal. This creates an intrinsic bias in the stored memories, influencing both the learning rule and the retrieval dynamics. The presence of biased patterns gives rise to several non trivial issues. Firstly, the Hebbian interaction tensor must be appropriately centred and scaled to eliminate the contribution arising from the non-zero mean and retain the relevant fluctuations [31]. Secondly, even after this transformation, the network dynamics may not converge to configurations that exhibit the same average activity as the stored patterns. The issue can be addressed by adding an extra constraint term to the Hamiltonian [4, 14]. This soft mechanism ensures the desired level of average activation is reached. Therefore, the resulting model depends not only on the storage load and temperature, but also on the degree of bias and the strength of the activation constraint.

In this study, we examine a dense higher-order Hopfield model featuring biased patterns within the high storage regime. We analyse the thermodynamic behaviour of the system within the Replica Symmetric ansatz and our primary objective is to understand the influence of pattern bias on both the retrieval threshold and the storage capacity. To this end, we first conduct a signal-to-noise analysis at zero temperature. This heuristic approach suggests that, while there is a bias-dependent reduction in the effective load, the dense architecture preserves its superlinear storage capacity, captured by a multiplicative factor $(1-b^2)^P$, with $b$ patterns bias. In this way we show that bias induces a universal renormalization of the effective load, preserving the superlinear scaling while quantitatively modifying the retrieval threshold. We then perform a full thermodynamic analysis based on Guerra's interpolation method [20].

This approach allows us to derive the quenched statistical pressure of the model in the high storage regime, as well as obtain the corresponding self consistency equations for the relevant order parameters: the Mattis magnetization, the mean activity and the two-replicas overlap. The analytical solution corroborates the heuristic prediction obtained from the signal-to-noise analysis and this agreement provides a consistent picture of how bias quantitatively affects retrieval performance while maintaining the qualitative scaling behaviour of the model.

The paper is organized as follows. In Sect. 2, we define the model and introduce the biased pattern distribution together with the corresponding recentered Hebbian prescription and activity constraint. In Sect. 3, we present the signal-to-noise analysis and derive the associated estimate for the retrieval threshold at zero temperature. In Sect. 4, we develop the Guerra interpolation scheme and derive the RS expression of the quenched free energy, along with the self consistency equations for the order parameters and we discuss the resulting phase diagram and the retrieval properties of the model as functions of the bias, the storage load, the temperature, and the strength of the activation constraint.

## 2 Generalities

In this section we introduce the model under analysis, namely the *dense biased associative memory*.

Before doing this, let us analyze the *standard* dense generalization of associative memory with Hebbian synaptic couplings [9, 19]. In this case, one considers a collection of $K$ Boolean patterns of length $N$, denoted by $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^K$, whose entries are i.i.d. Rademacher random variables, namely

$$\mathcal{P}(\xi_i^\mu) = \frac{1}{2}\delta(\xi_i^\mu - 1) + \frac{1}{2}\delta(\xi_i^\mu + 1), \qquad \forall\, i = 1, \dots, N, \quad \mu = 1, \dots, K. \tag{1}$$

These random variables are used to construct the $P$-body Hebbian interaction tensor

$$J^{(P)}_{i_1,\dots,i_P}(\boldsymbol{\xi}) = \frac{1}{N^{P-1}} \sum_{\mu=1}^{K} \xi_{i_1}^\mu \cdots \xi_{i_P}^\mu, \tag{2}$$

which encodes the $P$-points correlations among the stored patterns and models the effective interactions among neurons in the network.

The neural configuration is described by

$$\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N) \in \Omega, \qquad \Omega = \{-1, +1\}^N, \tag{3}$$

and the system is governed by the Hamiltonian

$$H_N^{(P)}(\boldsymbol{\sigma} \mid \boldsymbol{\xi}) = -\frac{1}{P!} \sum_{i_1,\dots,i_P=1}^{N} J^{(P)}_{i_1 \cdots i_P}(\boldsymbol{\xi})\, \sigma_{i_1} \cdots \sigma_{i_P}. \tag{4}$$

The retrieval quality of pattern $\boldsymbol{\xi}^\mu$ is quantified by the corresponding *Mattis magnetization*

$$m_\mu(\boldsymbol{\sigma}|\boldsymbol{\xi}) = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \sigma_i, \qquad \mu = 1, \dots, K. \tag{5}$$

It is well known [19, 23] that these higher-order generalizations of the standard Hopfield model can retrieve patterns with almost perfect accuracy even when the number of stored

patterns grows superlinearly with the system size, namely

$$K \sim \mathcal{O}(N^{P-1}). \tag{6}$$

Despite the extensive Literature on this class of models, the case in which the pattern entries are not balanced has received much less attention. In particular, a systematic investigation is still lacking when the distribution of the pattern entries is biased, i.e., when the probabilities of taking the values $+1$ and $-1$ are no longer equal.

To model this situation, we replace (1) with the biased distribution

$$\mathcal{P}(\xi_i^\mu) = \frac{1+b}{2}\delta(\xi_i^\mu - 1) + \frac{1-b}{2}\delta(\xi_i^\mu + 1), \qquad \forall\, i = 1, \dots, N, \quad \mu = 1, \dots, K, \tag{7}$$

where $b \in (-1, +1)$ is a size-independent tunable parameter measuring the degree of bias, with the case $b = 0$ that encodes the unbiased situation. In the biased case, it is known [12] that the Hebbian prescription (2) must be suitably modified by recentering and rescaling the pattern entries. One therefore introduces

$$J^{(P)}_{i_1,\dots,i_P} = \frac{1}{N^{P-1}} \sum_{\mu=1}^{K} \eta^\mu_{i_1} \cdots \eta^\mu_{i_P}, \tag{8}$$

where

$$\eta_i^\mu = \frac{\xi_i^\mu - b}{\sqrt{1-b^2}}. \tag{9}$$

This change is natural since $\mathbb{E}[\eta_i^\mu] = 0$ and $\mathbb{E}[(\eta_i^\mu)^2] = 1$, so the unbiased variables $\boldsymbol{\eta}$ play the same role as the unbiased variables in the standard setting, having the same first two moments.

However, this modification alone is not sufficient. Indeed, inserting (8) into (4) does not guarantee that the neural configuration converges toward one of the original stored patterns, which now have average activation equal to $b$ [4, 14]. To enforce this property, one needs to introduce an additional term in the Hamiltonian, playing the role of a *pseudo-Lagrange multiplier*, in order to constrain the network toward configurations with the same mean activation as the stored patterns. For this reason, the Hamiltonian considered in our analysis is

$$H^{(P)}_{g,N}(\boldsymbol{\sigma} \mid \boldsymbol{\xi}) = -\frac{1}{P!}\frac{1}{N^{P-1}} \sum_{\mu=1}^{K} \sum_{i_1,\dots,i_P=1}^{N} \eta^\mu_{i_1} \cdots \eta^\mu_{i_P}\, \sigma_{i_1} \cdots \sigma_{i_P} + gN \left( \frac{1}{N} \sum_{i=1}^{N} \sigma_i - b \right)^2, \tag{10}$$

where $g$ is a tunable parameter controlling the strength of the constraint on the mean neural activity.

We have represented a didactic example of how this network works in Fig. 1.

For later convenience, and in view of the mathematical analysis, we introduce a *modified Mattis magnetization* computed with respect to the new unbiased patterns $\boldsymbol{\eta}$:

$$\tilde{m}_\mu(\boldsymbol{\sigma}|\boldsymbol{\xi}) = \frac{1}{N} \sum_{i=1}^{N} \eta_i^\mu \sigma_i = \frac{1}{\sqrt{1-b^2}} \Big( m_\mu(\boldsymbol{\sigma}|\boldsymbol{\xi}) - bM(\boldsymbol{\sigma}) \Big), \tag{11}$$

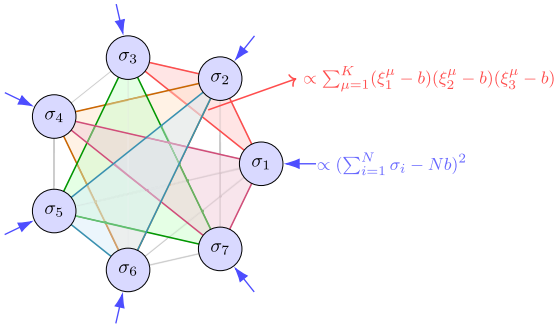


**Fig. 1** Representation of a toy example of the network under investigation with $P = 3$ grade of interactions, $N = 7$ spins. Vertices correspond to spins $\sigma_i$, and each triangle encodes a term $\sigma_{i_1}\sigma_{i_2}\sigma_{i_3}$ weighted by the biased patterns $\boldsymbol{\eta}^\mu = \boldsymbol{\xi}^\mu - b$. Highlighted triangles represent selected contributions, while gray ones denote the remaining possible interactions. The global term proportional to $(\sum_{i=1}^{N}\sigma_i - Nb)^2$ is also indicated

where we remember that $\eta_i^\mu = \eta_i^\mu(\boldsymbol{\xi})$ and we have also used the normalized mean activation of the network defined as,

$$M(\boldsymbol{\sigma}) = \frac{1}{N}\sum_{i=1}^{N}\sigma_i. \tag{12}$$

In terms of these order parameters, the Hamiltonian (10) can be rewritten as

$$H_{g,N}^{(P)}(\boldsymbol{\sigma}\,|\,\boldsymbol{\xi}) = -\frac{1}{P!}\frac{N}{2}\sum_{\mu=1}^{K}\left(\tilde{m}_\mu(\boldsymbol{\sigma}|\boldsymbol{\xi})\right)^P + \frac{g}{2}N\left(M(\boldsymbol{\sigma}) - b\right)^2. \tag{13}$$

Writing the Hamiltonian as in (13) makes explicit its linear scaling with the system size $N$ of the network.

As usual, the configuration space $\Omega$ is endowed with the random Boltzmann-Gibbs measure

$$\mathbb{P}_{g,N}^{(P)}(\boldsymbol{\sigma}|\boldsymbol{\xi}) = \frac{\exp\left(-\beta\, H_{g,N}^{(P)}(\boldsymbol{\sigma}\,|\,\boldsymbol{\xi})\right)}{Z_{g,N}^{(P)}(\beta\,|\,\boldsymbol{\xi})}, \tag{14}$$

where

$$Z_{g,N}^{(P)}(\beta\,|\,\boldsymbol{\xi}) = \sum_{\boldsymbol{\sigma}\in\Omega}\exp\left[-\beta\, H_{g,N}^{(P)}(\boldsymbol{\sigma}\,|\,\boldsymbol{\xi})\right] \tag{15}$$

is the partition function. Accordingly, for any observable $f(\boldsymbol{\sigma})$, the Boltzmann-Gibbs average is defined as

$$\omega_{g,N,\boldsymbol{\xi}}^{(P)}(f(\boldsymbol{\sigma})) = \frac{1}{Z_{g,N}^{(P)}(\beta\,|\,\boldsymbol{\xi})}\sum_{\boldsymbol{\sigma}\in\Omega} f(\boldsymbol{\sigma})\,\exp\left(-\beta\, H_{g,N}^{(P)}(\boldsymbol{\sigma}\,|\,\boldsymbol{\xi})\right), \tag{16}$$

while the quenched average is

$$\langle f(\boldsymbol{\sigma})\rangle = \mathbb{E}_{\boldsymbol{\xi}}\left[\omega_{g,N,\boldsymbol{\xi}}^{(P)}(f(\boldsymbol{\sigma}))\right]. \tag{17}$$

Throughout this work, we focus on the *high-storage* regime, namely we assume that the load $\gamma$

$$\gamma = \lim_{N\to\infty}\frac{K}{N^{P-1}} > 0 \tag{18}$$

remains finite in the thermodynamic limit. Our goal is to study the asymptotic behavior of the corresponding quenched statistical pressure, for fixed $P > 2$, under the Replica Symmetric (RS) *ansatz*, namely

$$\mathcal{A}^{(P)}(\beta, g) = \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\boldsymbol{\xi}} \log Z^{(P)}_{g,N}(\beta \mid \boldsymbol{\xi}). \tag{19}$$

More precisely, we aim to assess the collective retrieval capabilities of the network in the presence of biased patterns and superlinear storage (18).

To characterize the macroscopic behavior of the model, we introduce some relevant functions of the neural configuration $\boldsymbol{\sigma}$, denoted *order parameters*. Retrieval is monitored through the Mattis magnetization with respect to the selected pattern, namely the standard Mattis overlap (5) and its recentered counterpart (11). In addition, we consider the *two-replicas overlap*

$$q_{ab}(\boldsymbol{\sigma}) = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^a \sigma_i^b, \tag{20}$$

which quantifies the amount of cross-talk noise affecting the network.

Obtaining an explicit expression for the quenched free energy in terms of the relevant order parameters and control parameters of the theory –namely the bias level $b$, the constraint strength $g$, the inverse temperature $\beta$, and the load $\gamma$– allows us to derive the associated self consistency equations. Their numerical solution then provides the phase diagrams of the model, i.e., the regions in control parameter space corresponding to different computational behaviors, characterized by the values of the order parameters themselves.

Before turning to the full thermodynamic analysis, let us present a preliminary and well-known argument based on the signal-to-noise ratio. This approach allows us to assess, with minimal technical machinery and relying essentially on the update rule together with the Central Limit Theorem (CLT), whether in the zero-temperature limit ($\beta \to \infty$) a configuration initially prepared exactly on one of the stored patterns, say $\boldsymbol{\xi}^1$ without loss of generality, remains stable after one step of the dynamics. In other words, we want to determine whether the configuration stays aligned with the retrieved pattern or whether the cumulative interference generated by the other $K - 1$ stored patterns is strong enough to drive the system away from it. This argument provides a first estimate of whether the introduction of bias in the pattern distribution modifies the storage capacity of the network.

## 3 Signal-to-Noise Analysis

Exploiting the mean-field structure of the model, the Hamiltonian can be conveniently rewritten as

$$H_{g,N}(\boldsymbol{\sigma}|\boldsymbol{\xi}) = -\frac{1}{P!} \sum_{i=1}^{N} \hat{h}_i(\boldsymbol{\sigma})\sigma_i, \tag{21}$$

where the effective local field acting on neuron $i$ is given by

$$\hat{h}_i(\boldsymbol{\sigma}) := \frac{1}{N^{P-1}} \sum_{\mu=1}^{K} \sum_{\substack{i_2,\dots,i_P=1 \\ \neq i}}^{N} \eta_i^\mu \eta_{i_2}^\mu \cdots \eta_{i_P}^\mu \sigma_{i_2} \cdots \sigma_{i_P} - gP! \left( \frac{1}{N} \sum_{\substack{j=1 \\ \neq i}}^{N} \sigma_j - b \right). \tag{22}$$

We now introduce a dynamics that drives the system toward lower-energy configurations. In the presence of thermal noise, controlled by the inverse temperature $\beta = T^{-1} \in \mathbb{R}^+$, the

neural configuration $\boldsymbol{\sigma}(t)$ at discrete time $t$ evolves synchronously according to

$$\boldsymbol{\sigma}(t+1) = \text{sign}\Big[\tanh\big(\beta\hat{\boldsymbol{h}}(\boldsymbol{\sigma}(t))\big) + \boldsymbol{u}(t)\Big], \tag{23}$$

where $\hat{\boldsymbol{h}}$ is the effective field defined in Eq. (22), and $\boldsymbol{u}(t) \overset{i.i.d.}{\sim} \mathcal{U}([-1,1]^N)$.
To perform the signal-to-noise analysis, we initialize the network in one of the stored patterns,

$$\boldsymbol{\sigma}(0) = \boldsymbol{\xi}^1, \tag{24}$$

and compute the effective local field at time $t = 0$. We then estimate the overlap between the updated configuration and the same pattern in order to determine whether the initial condition is stable under the dynamics [1].

Since we are interested in the zero-temperature limit, the stochastic update rule (23) reduces to the deterministic dynamics

$$\boldsymbol{\sigma}(t+1) = \text{sign}\Big[\hat{\boldsymbol{h}}(\boldsymbol{\sigma}(t))\Big]. \tag{25}$$

The alignment of the network state with the first pattern can then be monitored through the Mattis magnetization (5), which at time $t+1$ reads

$$m_1(t+1) = \frac{1}{N}\sum_{i=1}^{N}\xi_i^1 \,\text{sign}\Big[\hat{h}_i(\boldsymbol{\sigma}(t))\Big] = \frac{1}{N}\sum_{i=1}^{N}\xi_i^1\sigma_i(t)\,\text{sign}\Big[\hat{h}_i(\boldsymbol{\sigma}(t))\sigma_i(t)\Big]. \tag{26}$$

Under the initialization (24), one has $m_1(0) = 1$. Therefore, in order to assess whether the configuration remains aligned with $\boldsymbol{\xi}^1$ after one update, we have to evaluate

$$m_1(1) = \frac{1}{N}\sum_{i=1}^{N}\xi_i^1\sigma_i(0)\,\text{sign}\Big[\hat{h}_i(\boldsymbol{\sigma}(0))\sigma_i(0)\Big] = \frac{1}{N}\sum_{i=1}^{N}\text{sign}\Big[\hat{h}_i(\boldsymbol{\xi}^1)\xi_i^1\Big]. \tag{27}$$

Hence, the condition $m_1(1) \approx 1$ is translated into the requirement that, with high probability,

$$\text{sign}\Big(\hat{h}_i(\boldsymbol{\xi}^1)\xi_i^1\Big) = 1, \tag{28}$$

or equivalently,

$$\hat{h}_i(\boldsymbol{\xi}^1)\xi_i^1 > 0. \tag{29}$$

We now decompose the local field into the contribution of the retrieved pattern, the crosstalk generated by the remaining patterns, and the pseudo-Lagrange multiplier term. To this aim, let

$$\chi_i^\mu := \xi_i^1\eta_i^\mu = \frac{(\xi_i^\mu - b)\xi_i^1}{\sqrt{1-b^2}}. \tag{30}$$

Then, evaluating the field on $\boldsymbol{\sigma} = \boldsymbol{\xi}^1$, we can write

$$\hat{h}_i(\boldsymbol{\xi}^1)\xi_i^1 = S_i + N_i + G_i, \tag{31}$$

where

$$S_i = \chi_i^1\left(\frac{1}{N}\sum_{j\neq i}\chi_j^1\right)^{P-1} \tag{32}$$

is the signal generated by the retrieved pattern,

$$N_i = \sum_{\mu=2}^{K} \chi_i^\mu \left( \frac{1}{N} \sum_{j \neq i} \chi_j^\mu \right)^{P-1} \tag{33}$$

is the crosstalk contribution, and

$$G_i = -g P! \xi_i^1 \left( \frac{1}{N} \sum_{j \neq i} \xi_j^1 - b \right) \tag{34}$$

is the contribution of the pseudo-Lagrange multiplier. For fixed $g$, since

$$\frac{1}{N} \sum_{j \neq i} \xi_j^1 = b + \mathcal{O}(N^{-1/2}), \tag{35}$$

one has

$$G_i \underset{d}{\sim} \mathcal{O}(N^{-1/2}), \tag{36}$$

and this term is subleading in the large-$N$ limit.

The key point is that the Gaussian approximation should be applied to the crosstalk term conditionally on the retrieved pattern $\boldsymbol{\xi}^1$, rather than to the full unconditioned random variable $\hat{h}_i(\boldsymbol{\xi}^1)\xi_i^1$. Indeed, the retrieved pattern determines the signal, whereas the remaining patterns generate the crosstalk noise. Let us first analyze the signal. For $\mu = 1$, one has

$$\chi_i^1 = \xi_i^1 \eta_i^1 = \frac{1 - b\xi_i^1}{\sqrt{1 - b^2}}. \tag{37}$$

Therefore,

$$\mathbb{E}[\chi_i^1] = \sqrt{1 - b^2}, \qquad \mathrm{Var}(\chi_i^1) = b^2. \tag{38}$$

By the law of large numbers,

$$\frac{1}{N} \sum_{j \neq i} \chi_j^1 = \sqrt{1 - b^2} + \mathcal{O}(N^{-1/2}), \tag{39}$$

and therefore

$$S_i = (1 - b\xi_i^1)(1 - b^2)^{(P-2)/2} + \mathcal{O}(N^{-1/2}). \tag{40}$$

Thus the bias does not induce an additional Gaussian noise contribution in the signal. Rather, it splits the local signal into two positive amplitudes:

$$S_i \simeq (1 - b)(1 - b^2)^{(P-2)/2}, \qquad \text{if } \xi_i^1 = +1, \tag{41}$$

and

$$S_i \simeq (1 + b)(1 - b^2)^{(P-2)/2}, \qquad \text{if } \xi_i^1 = -1. \tag{42}$$

Both amplitudes are positive for every fixed $b \in (-1, 1)$.

We now consider the crosstalk term $N_i$ conditionally on $\boldsymbol{\xi}^1$. For every $\mu \geq 2$,

$$\mathbb{E}[\chi_i^\mu \mid \boldsymbol{\xi}^1] = \xi_i^1 \mathbb{E}[\eta_i^\mu] = 0, \tag{43}$$

and

$$\mathrm{Var}(\chi_i^\mu \mid \boldsymbol{\xi}^1) = (\xi_i^1)^2 \mathbb{E}[(\eta_i^\mu)^2] = 1. \tag{44}$$

Moreover, $\chi_i^\mu$ is independent of the empirical average $N^{-1}\sum_{j\neq i}\chi_j^\mu$. Hence, up to lower-order corrections in $N$,

$$\mathbb{E}[N_i \mid \boldsymbol{\xi}^1] = 0, \tag{45}$$

and

$$\mathrm{Var}(N_i \mid \boldsymbol{\xi}^1) = (K-1)\mathbb{E}\left[\left(\frac{1}{N}\sum_{j\neq i}\chi_j^\mu\right)^{2(P-1)}\middle|\boldsymbol{\xi}^1\right]. \tag{46}$$

Since the variables $\chi_j^\mu$, conditionally on $\boldsymbol{\xi}^1$, have zero mean and unit variance, the leading contribution is

$$\mathrm{Var}(N_i \mid \boldsymbol{\xi}^1) = (2P-3)!!\frac{K-1}{N^{P-1}}\left[1+\mathcal{O}(N^{-1})\right]. \tag{47}$$

Therefore, under the standard Central Limit Theorem approximation for the crosstalk term,

$$N_i \mid \boldsymbol{\xi}^1 \simeq \mathcal{N}\left(0, (2P-3)!!\frac{K-1}{N^{P-1}}\right). \tag{48}$$

With the normalization adopted in Eq. (22), the leading crosstalk variance is independent of the bias $b$. The dependence on $b$ enters through the local signal amplitudes in Eqs. (41)–(42).

Using the law of total probability, we write

$$\mathbb{P}\left[\hat{h}_i(\boldsymbol{\xi}^1)\xi_i^1 > 0\right] = \mathbb{E}_{\boldsymbol{\xi}^1}\left[\mathbb{P}\left[S_i + N_i + G_i > 0\middle|\boldsymbol{\xi}^1\right]\right]. \tag{49}$$

Neglecting the subleading term $G_i$, and using Eq. (48), we obtain

$$\mathbb{P}\left[\hat{h}_i(\boldsymbol{\xi}^1)\xi_i^1 > 0\middle|\boldsymbol{\xi}^1\right] \simeq \Phi\left(\frac{S_i}{\sqrt{(2P-3)!!\frac{K-1}{N^{P-1}}}}\right), \tag{50}$$

where $\Phi$ denotes the cumulative distribution function of a standard Gaussian random variable. Averaging over the two possible values of $\xi_i^1$, gives

$$\begin{aligned}\mathbb{P}\left[\hat{h}_i(\boldsymbol{\xi}^1)\xi_i^1 > 0\right] \simeq & \frac{1+b}{2}\Phi\left(\frac{(1-b)(1-b^2)^{(P-2)/2}}{\sqrt{(2P-3)!!\frac{K-1}{N^{P-1}}}}\right) \\ & + \frac{1-b}{2}\Phi\left(\frac{(1+b)(1-b^2)^{(P-2)/2}}{\sqrt{(2P-3)!!\frac{K-1}{N^{P-1}}}}\right).\end{aligned} \tag{51}$$

This expression is consistent with the limiting case $K = 1$: in the absence of crosstalk, the variance in Eq. (47) vanishes, and since the two signal amplitudes are positive for $b \in (-1, 1)$, the retrieval probability tends to one in the large-$N$ limit.

At fixed $P$ and fixed $b \in (-1, 1)$, Eq. (51) shows that the dense higher-order architecture retains a superlinear storage capacity in the system size. Indeed, the crosstalk variance remains of order one when

$$K = \mathcal{O}(N^{P-1}). \tag{52}$$

A conservative estimate of the bias-dependent prefactor is obtained by requiring that the weakest of the two local signal amplitudes remains comparable with the crosstalk standard deviation. Since

$$S_{\min} \sim (1 - |b|)(1 - b^2)^{(P-2)/2}, \tag{53}$$

this gives

$$K \sim \frac{(1 - |b|)^2 (1 - b^2)^{P-2}}{(2P - 3)!!} N^{P-1}. \tag{54}$$

Thus, in this one-step signal-to-noise estimate, the bias does not change the $N^{P-1}$ scaling of the capacity. Rather, it reduces the effective signal available at the weakest sites. With the centered and normalized pattern variables $\eta_i^\mu$, the leading crosstalk variance is independent of $b$, while the bias dependence enters through the local signal amplitudes.

## 4 Guerra's Interpolation Technique

The signal-to-noise argument presented in the previous section provides a first heuristic estimate of the retrieval threshold in the presence of biased patterns. We now show that the same bias dependent correction can be recovered analytically by means of *Guerra's interpolation method* [20], within the RS framework.

Our goal is to derive the quenched statistical pressure $\mathcal{A}^{(P)}(\beta, g)$ (19) in the high-storage regime. Since we are interested in the retrieval region, we isolate, without loss of generality, the first of the stored patterns, as the candidate retrieved pattern, and separate its contribution from that of the remaining $K - 1$ patterns. The latter act as a source of cross-talk noise. Accordingly, the partition function can be rewritten as

$$Z_{g,N}^{(P)}(\beta \mid \boldsymbol{\xi}) = \sum_{\boldsymbol{\sigma} \in \Omega} \exp\left[ \frac{\beta N}{2} \left(\tilde{m}_1(\boldsymbol{\sigma} \mid \boldsymbol{\xi})\right)^P - \frac{\beta g}{2} N \left(M(\boldsymbol{\sigma}) - b\right)^2 + \frac{\beta}{N^{P-1}} \sum_{\mu>1}^{K} \sum_{i_1, \cdots, i_P = 1}^{N} \eta_{i_1}^\mu \cdots \eta_{i_P}^\mu \, \sigma_{i_1} \cdots \sigma_{i_P} \right], \tag{55}$$

where we rescale $\beta/P! \to \beta/2$ and $g\,P! \to g$. The first term represents the signal associated with the candidate retrieved pattern $\boldsymbol{\xi}^1$, while the last term collects the cross-talk noise generated by all the non-retrieved patterns $\boldsymbol{\xi}^{\mu>1}$.

In order to implement Guerra's interpolation scheme, we approximate the contribution of the non-retrieved patterns by means of the CLT (please see Appendix 1 for the detailed computations). In the high-storage regime, the sum over $\mu > 1$ gives rise to an effective Gaussian random term, so that the partition function can be recast as

$$Z_{g,N}^{(P)}(\beta \mid \boldsymbol{\xi}^{\mathbf{1}}, \boldsymbol{\lambda}) = \sum_{\boldsymbol{\sigma} \in \Omega} \exp\left[ \frac{\beta N}{2} \left(\tilde{m}_1(\boldsymbol{\sigma} \mid \boldsymbol{\xi})\right)^P - \frac{\beta g}{2} N \left(M(\boldsymbol{\sigma}) - b\right)^2 + \frac{\beta \tilde{\mathcal{V}} \sqrt{K}}{2N^{P-1}} \sum_{i_1 < \cdots < i_P = 1}^{N} \lambda_{i_1 \cdots i_P} \, \sigma_{i_1} \cdots \sigma_{i_P} \right], \tag{56}$$

where $\lambda_{i_1 \cdots i_P} \sim \mathcal{N}(0, 1)$ i.i.d. for $i_1, \ldots, i_P = 1, \ldots, N$ and $\tilde{\mathcal{V}} = \sqrt{2P[P!(2P - 3)!!]}$.

Following [20], we introduce an interpolating parameter $t \in [0, 1]$ and define the interpolating partition function as

$$Z_{g,N}^{(P)}(\beta|\boldsymbol{\xi}^{\mathbf{1}}, \boldsymbol{\lambda}, \boldsymbol{J}; t) = \sum_{\boldsymbol{\sigma}\in\Omega} \exp\Bigg[ t\frac{\beta N}{2}(\tilde{m}_1(\boldsymbol{\sigma}|\boldsymbol{\xi}))^P + \beta g N b M(\boldsymbol{\sigma}) - \frac{\beta g}{2} N t (M(\boldsymbol{\sigma}))^2 - \frac{\beta g}{2} N b^2 + \sqrt{t}\frac{\beta\tilde{\mathcal{V}}\sqrt{K}}{2N^{P-1}} \sum_{i_1<\cdots<i_P=1}^{N} \lambda_{i_1\cdots i_P}\, \sigma_{i_1}\cdots\sigma_{i_P} + (1-t)\psi N \tilde{m}_1(\boldsymbol{\sigma}|\boldsymbol{\xi}) + (1-t)\tilde{\psi} N M(\boldsymbol{\sigma}) + \sqrt{1-t}\, A \sum_i J_i\sigma_i \Bigg], \tag{57}$$

where $t \in [0, 1]$, the constants $\psi$, $\tilde{\psi}$, and $A$ will be fixed *a posteriori*, and $\{J_i\}_{i=1,\ldots,N}$ are i.i.d. standard Gaussian random variables.

We then introduce the finite-volume interpolating quenched statistical pressure

$$\mathcal{A}_N^{(P)}(\beta, g; t) = \frac{1}{N}\mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda},\boldsymbol{J}} \ln Z_{g,N}^{(P)}(\beta \mid \boldsymbol{\xi}^{\mathbf{1}}, \boldsymbol{\lambda}, \boldsymbol{J}; t), \tag{58}$$

and denote by

$$\mathcal{A}^{(P)}(\beta, g; t) = \lim_{N\to\infty} \mathcal{A}_N^{(P)}(\beta, g; t) \tag{59}$$

its thermodynamic limit.

By construction, at $t = 1$ we recover the original interacting model, so that

$$\mathcal{A}^{(P)}(\beta, g; t = 1) = \frac{1}{N}\mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \log Z_{g,N}^{(P)}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t = 1) = \mathcal{A}^{(P)}(\beta, g). \tag{60}$$

At $t = 0$, on the other hand, the model reduces to an effective one-body system, which is explicitly solvable: the neurons decouple and interact only with suitably chosen external fields that reproduce the relevant low-order statistics of the original interacting problem.

From now on, whenever no confusion arises, we omit the explicit dependence of the order parameters on $\boldsymbol{\sigma}$ and $\boldsymbol{\xi}$.

The relation between the two endpoints of the interpolation follows from the Fundamental Theorem of Calculus:

$$\mathcal{A}^{(P)}(\beta, g; t = 1) = \mathcal{A}^{(P)}(\beta, g; t = 0) + \int_0^1 dt\, \frac{\partial\mathcal{A}^{(P)}(\beta, g; s)}{\partial s}\Big|_{s=t}. \tag{61}$$

We now evaluate the two contributions separately.

***One-body contribution.*** The first term is readily computed. Setting $t = 0$ in (57), we obtain

$$Z_{g,N}^{(P)}(\beta|\boldsymbol{\xi}^{\mathbf{1}}, \boldsymbol{\lambda}, \boldsymbol{J}; t = 0) = \sum_{\boldsymbol{\sigma}\in\Omega} \exp\Bigg[\beta g b N M - \frac{\beta g}{2} N b^2 + \psi N \tilde{m}_1 + \tilde{\psi} N M + A\sum_i J_i\sigma_i\Bigg]. \tag{62}$$

At this stage the spins decouple, and the partition function factorizes into a product of one site contributions. Performing the Gaussian average and using standard algebraic manipulations, one finds

$$Z_{g,N}^{(P)}(\beta|\boldsymbol{\xi}^{\mathbf{1}}, \boldsymbol{J}; t = 0) = \exp\Bigg[-\frac{\beta g}{2} N b^2\Bigg] 2^N \cosh^N\Bigg[\beta g b + \psi\eta^1 + \tilde{\psi} + AJ\Bigg], \tag{63}$$

and therefore

$$\mathcal{A}^{(P)}(\beta, g; t = 0) = \log 2 - \frac{\beta g b^2}{2} + \mathbb{E}_{x,\eta^1} \log\cosh\Bigg[\psi\eta^1 + \beta g b + \tilde{\psi} + x\sqrt{A^2}\Bigg]. \tag{64}$$

The details of this computation are reported in Appendix 1.

***Streaming term.*** At finite volume, the derivative of the interpolating quenched pressure with respect to $t$ is computed in Appendix 1 and reads

$$\begin{aligned}\partial_t \mathcal{A}_N^{(P)}(\beta, g; t) = \frac{\beta}{2}\Big[\langle(\tilde{m}_1)^P\rangle - \frac{2\psi}{\beta}\langle\tilde{m}_1\rangle\Big] - \frac{\beta g}{2}\Big[\langle M^2\rangle + \frac{2\tilde{\psi}}{\beta g}\langle M\rangle\Big] \\ + \frac{\beta^2\tilde{\mathcal{V}}^2 K}{8N^{P-1}P!}\Big[1 - \langle q_{12}^P\rangle\Big] - \frac{A^2}{2}\Big[1 - \langle q_{12}\rangle\Big].\end{aligned} \tag{65}$$

We now impose the RS ansatz, namely that the relevant order parameters self-average around deterministic values in the thermodynamic limit. More precisely, for any order parameter $X$ with limiting value $\bar{X}$, we assume

$$\lim_{N\to\infty} \mathbb{P}_N(X) = \delta(X - \bar{X}). \tag{66}$$

Under this assumption, the streaming term simplifies in the thermodynamic limit to

$$\partial_t \mathcal{A}^{(P)}(\beta, g; t) = -\frac{\beta(P-1)}{2}\bar{\tilde{m}}^P + \frac{\beta g}{2}\bar{M}^2 - \frac{\beta^2\gamma}{8}\Big[1 - P\bar{q}^{P-1} + (P-1)\bar{q}^P\Big], \tag{67}$$

provided that we have used the definition of the load (18) and the interpolating parameters are chosen as

$$\psi = \beta\frac{P}{2}\bar{\tilde{m}}^{P-1}, \qquad \tilde{\psi} = -\beta g\bar{M}, \qquad A^2 = \frac{\beta^2\gamma P}{2}\bar{q}^{P-1} \tag{68}$$

and

$$2P(2P-3)!!\frac{K}{N^{P-1}} =: \gamma \tag{69}$$

***RS statistical pressure and self consistency equations.*** Substituting (64) and (67) into (61), we obtain the RS expression for the quenched statistical pressure in the thermodynamic limit:

$$\begin{aligned}\mathcal{A}^{(P)}(\beta, g) = \log 2 + \mathbb{E}_{x,\xi^1}\log\cosh\left[\beta\frac{P}{2}\bar{\tilde{m}}^{P-1}\frac{(\xi^1 - b)}{\sqrt{1-b^2}} + \beta g(b - \bar{M}) + x\beta\sqrt{\frac{\gamma P}{2}\bar{q}^{P-1}}\right] \\ - \frac{\beta(P-1)}{2}\bar{\tilde{m}}^P - \frac{\beta g}{2}(b^2 - \bar{M}^2) - \frac{\beta^2\gamma}{8}\Big[1 - P\bar{q}^{P-1} + (P-1)\bar{q}^P\Big].\end{aligned} \tag{70}$$

The corresponding order parameters satisfy the following self-consistency equations

$$\begin{aligned}\bar{\tilde{m}} &= \mathbb{E}_{x,\xi^1}\left\{(\xi^1 - b)\tanh\left[\beta\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + \beta g(b - \bar{M}) + x\beta\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\bar{q}^{P-1}}\right]\right\}, \\ \bar{q} &= \mathbb{E}_{x,\xi^1}\left\{\tanh^2\left[\beta\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + \beta g(b - \bar{M}) + x\beta\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\bar{q}^{P-1}}\right]\right\}, \\ \bar{M} &= \mathbb{E}_{x,\xi^1}\left\{\tanh\left[\beta\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + \beta g(b - \bar{M}) + x\beta\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\bar{q}^{P-1}}\right]\right\}.\end{aligned} \tag{71}$$

We stress that here we have performed the rescalings $\bar{\tilde{m}} \to \bar{\tilde{m}}/\sqrt{1-b^2}$, $g(1-b^2) \to g$, and $\beta/(1-b^2)^{(P-2)/2} \to \beta$.

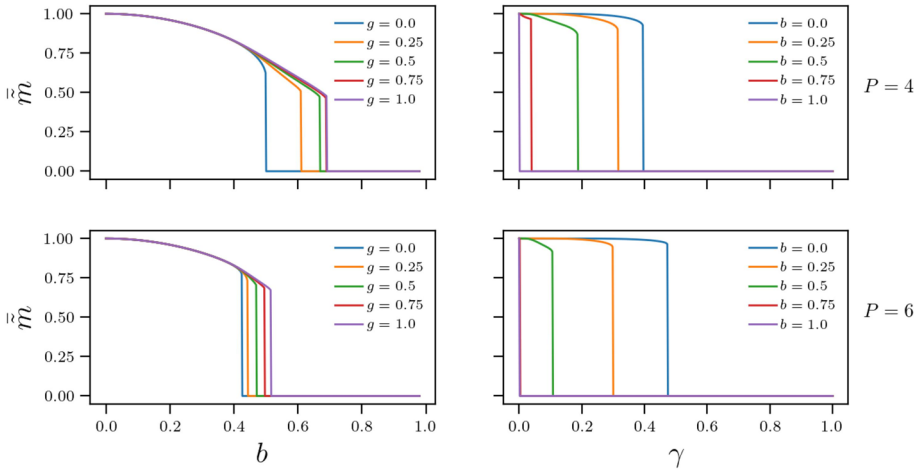


**Fig. 2** Numerical resolution of zero temperature self consistency equation is presented. The upper row corresponds to $P = 4$, and the lower row to $P = 6$. In the first column, $\gamma = 0.1$ is held fixed, while different values of $b$ and $g$ are explored. In the second column, $g = 1.0$ is fixed, and both the load $\gamma$ and bias $b$ are varied

Moreover exploiting (11) we can explicitly write down a self also for the Mattis Magnetization as

$$\bar{m} = \mathbb{E}_{x,\xi^1}\left\{\xi^1 \tanh\left[\beta\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + \beta g(b - \bar{M}) + x\beta\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}}\right]\right\}. \tag{72}$$

At this stage, it is worth emphasizing the full consistency between the heuristic signal-to-noise argument and the analytical Guerra interpolation scheme. Indeed, the variance of the effective Gaussian noise entering the self-consistency equations is renormalized by the same multiplicative factor $(1 - b^2)^P$ already predicted by the signal-to-noise estimate in (54). Therefore, the bias does not modify the superlinear scaling of the storage capacity, but reduces its prefactor in a quantitatively consistent way across the two approaches.

Finally, one can also analyze the zero-temperature limit of the previous equations which can be further simplified to

$$\begin{aligned}
\bar{\tilde{m}} &= \frac{1-b^2}{2}\left\{\mathrm{erf}\left[\frac{\frac{P}{2}\bar{\tilde{m}}^{P-1}(1-b) + g(b-\bar{M})}{\sqrt{\gamma P(1-b^2)^{P-2}}}\right] + \mathrm{erf}\left[\frac{\frac{P}{2}\bar{\tilde{m}}^{P-1}(1+b) - g(b-\bar{M})}{\sqrt{\gamma P(1-b^2)^{P-2}}}\right]\right\}, \\
\bar{M} &= \frac{1+b}{2}\mathrm{erf}\left[\frac{\frac{P}{2}\bar{\tilde{m}}^{P-1}(1-b) + g(b-\bar{M})}{\sqrt{\gamma P(1-b^2)^{P-2}}}\right] - \frac{1-b}{2}\mathrm{erf}\left[\frac{\frac{P}{2}\bar{\tilde{m}}^{P-1}(1+b) - g(b-\bar{M})}{\sqrt{\gamma P(1-b^2)^{P-2}}}\right].
\end{aligned} \tag{73}$$

These equations will be used to investigate the phase diagram and the retrieval performance of the model (see Appendix 1 for the calculation details).

To corroborate our results, we provide a numerical analysis. First, in Fig. 2 we plot the self consistency Eq. (73). Then, in Fig. 3, we use these equations to show how the critical load of the network depends on the bias parameters $b$ and $g$ for different interaction orders $P$.

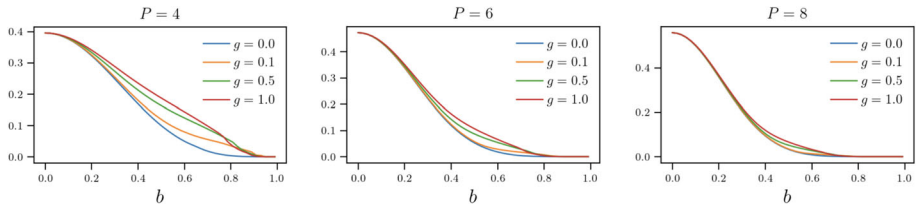


**Fig. 3** Behaviour of the critical load of the network at null temperature $T = 0$ for different values of $g$ and $b$. In each panel we change the grade of interactions of the network (from left to right, $P = 4, 6, 8$) and we analyze different values of $g$ from $g = 0$, when the term tuning the constraint on the activity of the network vanishes, to $g = 1$. We can see that $g$ plays the role of a light booster for the critical load, whose contribution decreases when the degree of interactions increases. Moreover, the critical load becomes smaller as the bias of the patterns increases

Furthermore, the behavior shown in Fig. 3 provides a clear physical intuition regarding the neural dynamics and the role of spurious states. In standard pairwise models, strong bias enlarges the attraction basins of spurious attractors, hindering correct thermalization. In our dense setting, the soft constraint $g$ prevents this dynamical failure at lower densities (e.g., $P = 4$, where $g > 0$ significantly improves $\gamma_c$). However, as $P$ increases ($P = 6, 8$), the highly non-linear dense interactions naturally suppress spurious states. Consequently, the network correctly thermalizes even without the auxiliary constraint, rendering $g$ largely uninfluential for high $P$ regimes.

In the next two figures (Figs. 4, 5), we compare, through Markov Chain Monte Carlo (MCMC) simulations, the retrieval performance of the biased pairwise Hopfield model with that of a dense biased associative memory with interaction order $P = 4$. We stress that the stored patterns are biased: specifically, the correlation between patterns is at least $b^2$, which corresponds to the random case baseline.

In Fig. 4, we initialize the MCMC dynamics from a state $\boldsymbol{\sigma}^{(0)}$ whose Mattis magnetization with respect to the first pattern is equal to 0.6, and we monitor the evolution of the Mattis magnetization $m_1(\sigma)$ and of the mean activity $M(\sigma)$ in the regime where the number of patterns scales as $K = O(N^2)$. The bias is fixed at $b = 0.3$ and the constraint at $g = 1.0$. The standard pairwise Hopfield model is unable to retrieve the patterns: indeed, after a few update steps, the magnetization decays to zero. By contrast, for $P = 4$ the pattern is successfully retrieved, as shown by the thermalized value of the Mattis magnetization, which approaches 1.0, while the mean activity stabilizes around $b$, consistently with the average activity of the biased patterns.

In Fig. 5, we further perform MCMC simulations close to the saturation load in order to test the stability of the biased patterns in a more challenging regime. We set $K = O(N^3)$, with $b = 0.3$, $g = 1.0$, and $\boldsymbol{\sigma}^{(0)} = \boldsymbol{\xi}^1$, and we let the system evolve while tracking both the Mattis magnetization with respect to the first pattern and the mean activity of the network. As can be seen, in the Hopfield model with $P = 2$ (top row) the stability of the pattern quickly deteriorates, since both the mean activity and the magnetization with respect to the first pattern decay to zero within a few steps. In contrast, the dense network with $P = 4$ (bottom row) continues to perform well, keeping both the mean activity and the Mattis magnetization of the first pattern close to $b$ and 1.0, respectively.

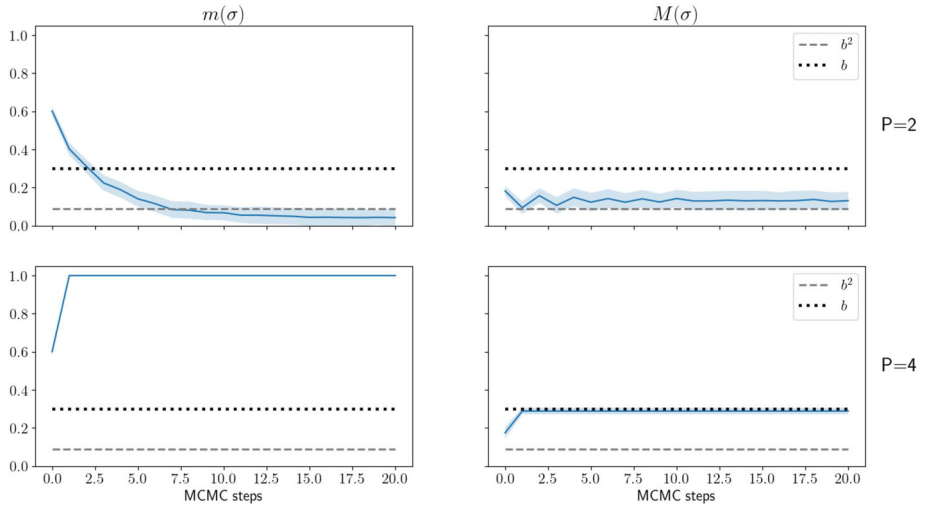


**Fig. 4** MCMC simulation, for $N = 1000$, $K = O(N^2)$, $b = 0.3$ and $g = 1.0$, of the Mattis magnetization $m(\sigma)$ and the mean activity $M(\sigma)$ with respect to different values of the degree of interactions $P$. Starting from a configuration which has an overlap with $\xi^1$ equal to 0.6, the standard Hopfield case ($P = 2$) is not able to retrieve the pattern, whereas the dense associative memory with $P = 4$ does this and, moreover, the mean activity stabilizes around the value of the bias $b$, as expected

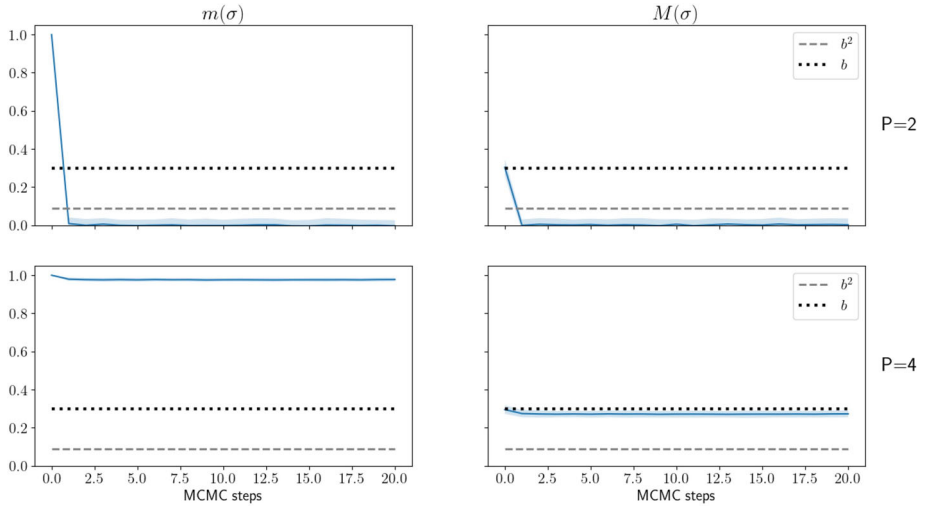


**Fig. 5** MCMC simulation for $N = 1000$, $K = 0.05N^3$, $b = 0.3$, $g = 1.0$. In the dense, low-load limit, the system successfully retrieves the stored pattern even in the presence of bias, yielding both $m(\sigma) = 1$ and $M(\sigma) = b$. Conversely, for $P = 2$ retrieval fails, as the load exceeds the network storage capacity. In all cases, the dynamics are initialized from a network configuration such that $\boldsymbol{\sigma}^{(0)} \cdot \boldsymbol{\xi}^1 \approx 1.0$. For $P = 2$, the trajectory rapidly departs from the initial condition, whereas for $P = 4$ the system remains trapped in the initial configuration, indicating stability

## 5 Conclusion

In this work, we investigated the retrieval properties of dense associative memories in the presence of biased patterns. The model we considered extends the standard $P-$spin Hopfield model architecture [23] to the case in which the stored patterns have non-zero mean activity. In this setting, the usual Hebbian prescription must be modified by recentering and rescaling the pattern entries, and a further, following [14], soft constraint on the global activity must be introduced in order to ensure that the network evolves toward configurations compatible with the average activation of the stored memories.

Our analysis combined two complementary perspectives. On the one hand, the signal-to-noise argument provided a simple and transparent zero-temperature estimate of the retrieval threshold. This heuristic computation showed that the effect of the bias is to reduce the effective storage capacity by a multiplicative factor $(1-b^2)^P$, while leaving unchanged the superlinear scaling $K \sim N^{P-1}$ that characterizes dense higher-order networks. On the other hand, the thermodynamic analysis based on Guerra's interpolation method allowed us to derive the quenched free energy within the RS framework and to obtain the corresponding self consistency equations for the relevant order parameters.

A central outcome of the paper is the full consistency between these two approaches. Indeed, the same bias dependent renormalization predicted by the signal-to-noise analysis emerges analytically in the RS equations. This agreement provides a clear physical interpretation of the role of bias: it does not alter the qualitative capacity scaling of the model, but quantitatively suppresses retrieval by increasing the effective difficulty of separating signal from noise.

More broadly, our results show that dense associative memories remain robust in the high storage regime even when the stored information is statistically unbalanced, provided that the learning rule and the network dynamics are modified consistently. The interplay among pattern bias, storage load, thermal noise, and activity constraint gives rise to a rich phase structure, whose numerical exploration can clarify the computational regions in which retrieval remains possible.

A further motivation for studying biased patterns comes from neuroscience-inspired associative memory models. Dense or higher-order associative architectures have recently been proposed as effective models for biological memory systems. For instance, dense associative memory ideas have been used in neuron-astrocyte associative memory models [22], where non-pairwise mechanisms may support enhanced storage, and in networks of grid and place cells [16], where supra-linear storage may be relevant for the representation of cognitive maps. These examples suggest that dense associative memories provide a useful framework for modelling high-capacity biological memory systems. In this context, biased patterns are particularly relevant because, in many neural systems, stored representations are not balanced but rather exhibit low levels of activity, with only a small fraction of neurons active in a given memory. This situation can be naturally related to the present Ising formulation by introducing Boolean variables $s_i^\mu \in 0, 1$ through $\xi_i^\mu = 2s_i^\mu - 1$, as already done for the Sherrington-Kirkpatrick model in [8]. If $a = \mathbb{P}(s_i^\mu = 1)$ denotes the Boolean activity level, then the Ising bias is $b = 2a - 1$, while the recentered variables used in this work can be written as

$$\eta_i^\mu = \frac{\xi_i^\mu - b}{\sqrt{1-b^2}} = \frac{s_i^\mu - a}{\sqrt{a(1-a)}}. \tag{74}$$

Thus, in the Boolean representation, the same parameter that fixes the average activity also controls the sparsity of the stored patterns: the low-activity regime $a \ll 1$ corresponds

to sparse neural codes. From this perspective, biased dense associative memories provide a natural statistical-mechanics framework for studying how sparse or low-activity neural representations can be stored and retrieved in high-capacity associative architectures.

Several directions deserve further investigation, both from the mathematical and from the neuroscience-inspired perspective. A first natural extension concerns the extension of the thermodynamic description beyond the Replica Symmetric ansatz. This would be relevant in regions of the phase diagram where glassy behavior may compete with retrieval and may require a treatment of possible Replica Symmetry Breaking effects, for instance through Guerra-type interpolation techniques. A second interesting problem is the characterization of the dynamical properties of the model beyond one step stability arguments, in particular the size of the basins of attraction and the convergence time toward biased memories. Another interesting extension of the present framework concerns hetero-associative architectures, where different networks interact and exchange information, potentially enhancing retrieval and generalization capabilities [7]. Finally, it would be worth exploring whether analogous bias-induced renormalizations persist in sparse or diluted higher-order architectures, as well as in modern Hopfield networks arising in machine learning [11, 26].

## Appendix A Universality of Noise in Dense Associative Memories

Following the step presented in [3], let us start by considering the noisy contribution generated by non-retrieved patterns

$$\frac{1}{N^{P-1}} \sum_{\mu>1}^{K} \sum_{i_1,\cdots,i_P=1}^{N} \eta_{i_1}^{\mu} \cdots \eta_{i_P}^{\mu} \sigma_{i_1} \cdots \sigma_{i_P} = N \sum_{\mu>1}^{K} \left(\tilde{m}_{\mu}\right)^P . \tag{A1}$$

Under the assumption of single-pattern retrieval of $\boldsymbol{\xi}^1$, it is easy to show that for $\mu > 1$

$$\tilde{m}_{\mu} \sim \mathcal{N}\left(0, \frac{1}{N}\right).$$

With a straightforward application of Cramer's lemma for large deviation theory, it can be shown that the variable $s_{\mu,l} = |\tilde{m}_{\mu}|^l$, is described by the probability distribution

$$\mathcal{P}(s_{\mu,l}) \propto s_{\mu,l}^{-(1-l^{-1})} \exp\left(-\frac{N}{2} s_{\mu,l}^{2/l}\right). \tag{A2}$$

For $l > 2$, the probability distribution has support on $]0; +\infty[$, with first and second moments given by

$$\mathbb{E}[s_{\mu,l}] = \frac{2^{\frac{l}{2}} N^{-\frac{l}{2}} \Gamma\left(\frac{l}{2} + \frac{1}{2}\right)}{\sqrt{\pi}}, \quad \mathbb{E}[s_{\mu,l}^2] = \frac{2^l N^{-l} \Gamma\left(l + \frac{1}{2}\right)}{\sqrt{\pi}}. \tag{A3}$$

For $l = P$ even

$$\mathbb{E}s_{\mu,P} = \left(\tfrac{1}{N}\right)^{P/2} (P-1)!!, \quad \mathbb{E}s_{\mu,P}^2 = \left(\tfrac{1}{N}\right)^{P} (2P-1)!!, \tag{A4}$$

so the variance is

$$\mathrm{Var}[s_{\mu,P}] = \left(\frac{1}{N}\right)^{P} \left[(2P-1)!! - ((P-1)!!)^2\right], \tag{A5}$$

Since the $s_{\mu,P} = m_\mu^P$'s are i.i.d. and their variance is finite, the sum (A1) converges as $K \to \infty$ to a Gaussian distributed variable by virtue of the CLT, in particular,

$$\frac{1}{N^P} \sum_{\mu>1}^{K} \sum_{i_1,\cdots,i_P=1}^{N} \eta_{i_1}^{\mu} \cdots \eta_{i_P}^{\mu} \sigma_{i_1} \cdots \sigma_{i_P} = \sum_{\mu>1}^{K} \left(\tilde{m}_\mu\right)^P \underset{d}{\to} \mathcal{N}\left(\mu_1, \mu_2 - \mu_1^2\right) \tag{A6}$$

with

$$\begin{aligned} \mu_1 &= K \left(\frac{1}{N}\right)^P (2P-1)!! \\ \mu_2 - \mu_1^2 &= K \left(\frac{1}{N}\right)^P \left[(2P-1)!! - ((P-1)!!)^2\right] \end{aligned} \tag{A7}$$

This means that the noise contribution $N \sum_{\mu\geq 2} m_\mu^P \sim N(\mu_1 + z\sqrt{\mu_2 - \mu_1^2})$, where $z \sim \mathcal{N}(0,1)$. Notice that the contribution of the expectation value can be neglected from the whole partition function as it would contribute to the free energy with a factor of order $\mathcal{O}(1/N)$.
In the spirit of Gaussian universality of noise generated by non-retrieved patterns, we consider the random variable

$$N^{-P} \mathcal{V} \sum_{i_1<\ldots<i_P} \lambda_{i_1\cdots i_P} \sigma_{i_1} \cdots \sigma_{i_P} \tag{A8}$$

with $\lambda_{i_1,\ldots,i_P} \underset{i.i.d.}{\sim} \mathcal{N}(0,1)$ and $\mathcal{V} > 0$ a free parameter to be suitably tuned. In the limit of large $N$ by the CLT it is easy to show that (A8) converge in distribution to $\mathcal{N}(0, \tilde{\sigma}^2)$ with $\tilde{\sigma} = N^{-P}\mathcal{V}\sqrt{\binom{N}{P}}$.

We can now tuning $\mathcal{V}$ in order to mach the two Gaussian distribution (A6) and (A8), obtaining

$$\mathcal{V}^2 = K N^P \frac{1}{\binom{N}{P}} \left[(2P-1)!! - ((P-1)!!)^2\right] \underset{N\gg 1}{\approx} K 2P(2P-3)!!\ P! = K\tilde{\mathcal{V}}^2 \tag{A9}$$

where in the last approximation we use $\frac{N!}{P!(N-P)!} \underset{N\gg 1}{\approx} \frac{N^P}{P!}$ and $\left[(2P-1)!! - ((P-1)!!)^2\right] \approx 2P(2P-3)!!$.

To conclude, we see that, in the $N, K \to \infty$ limit,

$$\frac{1}{N^{P-1}} \sum_{\mu>1}^{K} \sum_{i_1,\cdots,i_P=1}^{N} \eta_{i_1}^{\mu} \cdots \eta_{i_P}^{\mu} \sigma_{i_1} \cdots \sigma_{i_P} \underset{d}{\to} \tilde{\mathcal{V}}\sqrt{K} \sum_{i_1<\ldots<i_P} \lambda_{i_1,\ldots,i_P} \sigma_{i_1} \cdots \sigma_{i_P}, \tag{A10}$$

with $\lambda_{i_1,\ldots,i_P} \underset{i.i.d.}{\sim} \mathcal{N}(0,1)$.

## Appendix B Computation of t-Derivatives and One Body Terms of the Quenched Statistical Pressure

Let us start from the computation of the derivative with respect to $t$ of the quenched statistical pressure (58).

$$d_t \mathcal{A}_N(\beta, g; t) = \frac{1}{N}\mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t) \Bigg[ \frac{\beta N}{2} (\tilde{m}_1(\boldsymbol{\sigma}|\boldsymbol{\xi}))^P - \frac{\beta g}{2} N (M(\boldsymbol{\sigma}))^2 - \tilde{\psi} N M(\boldsymbol{\sigma}) - \frac{A}{2\sqrt{1-t}} \sum_i J_i \sigma_i + \frac{\beta \tilde{\mathcal{V}} \sqrt{K}}{2\sqrt{t} N^{P-1}} \sum_{i_1 < \cdots < i_P = 1}^{N} \lambda_{i_1 \cdots i_P} \sigma_{i_1} \cdots \sigma_{i_P} - \psi N \tilde{m}_1(\boldsymbol{\sigma}|\boldsymbol{\xi}) \Bigg] \tag{B11}$$

where $B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)$ is the Boltzmann weight, namely

$$B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t) = \exp \Bigg[ t \frac{\beta N}{2} (\tilde{m}_1(\boldsymbol{\sigma}|\boldsymbol{\xi}))^P + \beta g N b M(\boldsymbol{\sigma}) - \frac{\beta g}{2} N t (M(\boldsymbol{\sigma}))^2 - \frac{\beta g}{2} N b^2 \tag{B12}$$

$$+ \sqrt{t} \frac{\beta \tilde{\mathcal{V}} \sqrt{K}}{2N^{P-1}} \sum_{i_1 < \cdots < i_P = 1}^{N} \lambda_{i_1 \cdots i_P} \sigma_{i_1} \cdots \sigma_{i_P} \tag{B13}$$

$$+ (1-t) \psi N \tilde{m}_1(\boldsymbol{\sigma}|\boldsymbol{\xi}) + (1-t) \tilde{\psi} N M(\boldsymbol{\sigma}) + \sqrt{1-t}\, A \sum_i J_i \sigma_i \Bigg]. \tag{B14}$$

We can apply the definition of the quenched average (17) on the trivial terms in order to get

$$d_t \mathcal{A}_N(\beta, g; t) = \frac{\beta}{2} \langle (\tilde{m}_1(\boldsymbol{\sigma}|\boldsymbol{\xi}))^P \rangle - \frac{\beta g}{2} \langle (M(\boldsymbol{\sigma}))^2 \rangle - \tilde{\psi} \langle M(\boldsymbol{\sigma}) \rangle - \psi \langle \tilde{m}_1(\boldsymbol{\sigma}|\boldsymbol{\xi}) \rangle + \frac{1}{N}\mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t) \frac{\beta \tilde{\mathcal{V}} \sqrt{K}}{2\sqrt{t} N^{P-1}} \sum_{i_1 < \cdots < i_P}^{N} \lambda_{\boldsymbol{i}} \sigma_{i_1} \cdots \sigma_{i_P} - \frac{1}{N}\mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t) \frac{A}{2\sqrt{1-t}} \sum_i J_i \sigma_i. \tag{B15}$$

In order to compute the last two terms, we need to apply the *Stein's lemma* (also known as *Wick's theorem*) which states that for any smooth function $f(\boldsymbol{J})$ of a centered unit Gaussian variable $\boldsymbol{J}$ for which the two expectations $\mathbb{E}(\boldsymbol{J} f(\boldsymbol{J}))$ and $\mathbb{E}(\partial_{\boldsymbol{J}} f(\boldsymbol{J}))$ both exist the relation

$$\mathbb{E}_{\boldsymbol{J}}(\boldsymbol{J} f(\boldsymbol{J})) = \mathbb{E}_{\boldsymbol{J}} \left( \frac{\partial}{\partial \boldsymbol{J}} f(\boldsymbol{J}) \right), \tag{B16}$$

holds. Since the last two terms of (B15) have analogous computations, we will show the computations only for the first one.

$$\frac{\beta \tilde{\mathcal{V}} \sqrt{K}}{2\sqrt{t} N^P} \sum_{i_1 < \cdots < i_P = 1}^{N} \mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \left[ \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t) \lambda_{i_1 \cdots i_P} \sigma_{i_1} \cdots \sigma_{i_P} \right]$$

$$= \frac{\beta\tilde{\mathcal{V}}\sqrt{K}}{2\sqrt{t}N^P} \sum_{i_1<\cdots<i_P}^{N} \mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \left[ \partial_{\lambda_i} \left( \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)\sigma_{i_1}\cdots\sigma_{i_P} \right) \right]; \tag{B17}$$

$$\partial_{\lambda_{i_1\cdots i_P}} \left( \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)\sigma_{i_1}\cdots\sigma_{i_P} \right) =$$

$$\partial_{\lambda_{i_1\cdots i_P}} \left( \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \right) \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)\sigma_{i_1}\cdots\sigma_{i_P}$$

$$+ \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} \partial_{\lambda_{i_1\cdots i_P}} \left( B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)\sigma_{i_1}\cdots\sigma_{i_P} \right); \tag{B18}$$

$$\partial_{\lambda_{i_1\cdots i_P}} \left( \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \right) = -\frac{\beta\tilde{\mathcal{V}}\sqrt{Kt}}{2N^{P-1}} \left( \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \right)^2$$

$$\cdot \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)\sigma_{i_1}\cdots\sigma_{i_P}; \tag{B19}$$

$$\partial_{\lambda_{i_1\cdots i_P}} \left( B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)\sigma_{i_1}\cdots\sigma_{i_P} \right) = \sqrt{t}\frac{\beta\tilde{\mathcal{V}}\sqrt{K}}{2N^{P-1}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t) \left(\sigma_{i_1}\cdots\sigma_{i_P}\right)^2. \tag{B20}$$

Putting all the terms together we get

$$\frac{\beta^2 K\tilde{\mathcal{V}}^2}{2N^{2P-1}P!} \sum_{i_1,\cdots,i_P=1}^{N} \left\{ \mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \left[ \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)(\sigma_{i_1}\cdots\sigma_{i_P})^2 \right] \right.$$

$$\left. - \mathbb{E}_{\boldsymbol{\xi}} \left( \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)\sigma_{i_1}\cdots\sigma_{i_P} \right)^2 \right\}$$

$$\frac{\beta^2 K\tilde{\mathcal{V}}^2}{2N^{2P-1}P!} \sum_{i_1,\cdots,i_P=1}^{N} \mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \left[ \omega(\sigma_{i_1}\cdots\sigma_{i_P})^2 - \omega^2(\sigma_{i_1}\cdots\sigma_{i_P}) \right] = \frac{\beta^2 K\tilde{\mathcal{V}}^2}{2N^{P-1}P!} \left( 1 - \langle q_{12}^P \rangle \right). \tag{B21}$$

where in the last passage we have applied the definition of the two-replicas overlap (20) and the approximation $\sum_{i_1<\ldots<i_P} \approx (P!)^{-1} \sum_{i_1,\ldots,i_P}$.

In the same way one can compute the last term in (B15) getting

$$-\frac{A}{2\sqrt{1-t}} \sum_{i=1}^{N} \mathbb{E}_{\boldsymbol{\xi}^1,\boldsymbol{\lambda}} \frac{1}{Z_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t)} \sum_{\boldsymbol{\sigma}} B_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t) J_i\sigma_i = \frac{A^2}{2} \left(1 - \langle q_{12} \rangle\right). \tag{B22}$$

Finally, inserting (B21) and (B22) in (B15) and applying the definition of the quenched average (17), we get the expression of the derivative of the interpolating quenched statistical pressure with respect to $t$ at finite size $N$ (65), namely

$$\partial_t \mathcal{A}_N^{(P)}(\beta, g; t) = \frac{\beta}{2} \left[ \langle (\tilde{m}_1)^P \rangle - \frac{2\psi}{\beta} \langle m_1 \rangle \right] - \frac{\beta g}{2} \left[ \langle M^2 \rangle + \frac{2\tilde{\psi}}{\beta g} \langle M \rangle \right]$$

$$+ \frac{\beta^2 K\tilde{\mathcal{V}}^2}{8N^{P-1}P!} \left[ 1 - \langle q_{12}^P \rangle \right] - \frac{A^2}{2} \left[ 1 - \langle q_{12} \rangle \right]. \tag{B23}$$

In the thermodynamic limit $N \to +\infty$ we apply RS assumption (66). This means that all the centered moments higher than second order vanishes. This allows us to write for any order parameter $x$ whose equilibrium value is $\bar{x}$

$$\langle x^P \rangle - P\,\bar{x}^{P-1}\langle x \rangle = -(P-1)\bar{x}^P + \sum_{k=2}^{P} \binom{P}{k} \langle (x-\bar{x})^k \rangle \bar{x}^{P-k}, \tag{B24}$$

where the last term vanishes in RS assumption.
With some algebraic manipulation and putting the constants $\psi$, $\tilde{\psi}$, $A$ as in (68) we get the thermodynamic streaming term.

As far as one body terms concerns, one can rewrite the interpolating partition function (62) at $t = 0$ as

$$\begin{aligned} Z^{(P)}_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t=0) &= \sum_{\boldsymbol{\sigma}\in\Omega} \exp\left[\beta g b \sum_{i=1}^N \sigma_i + \psi \sum_{i=1}^N \eta_i^1 \sigma_i + \tilde{\psi} \sum_{i=1}^N \sigma_i + A \sum_{i=1}^N J_i \sigma_i\right] \\ &= \sum_{\boldsymbol{\sigma}\in\Omega} \prod_{i=1}^N \exp\left[\beta g b \sigma_i - \frac{\beta g}{2} N b^2 + \psi \eta_i^1 \sigma_i + \tilde{\psi}\sigma_i + A J_i \sigma_i\right] \\ &= \exp\left(-\frac{\beta g}{2} N b^2\right) \prod_{i=1}^N \sum_{\sigma_i \in \{-1,+1\}} \exp\left[\sigma_i \left(\beta g b + \psi \eta^1 + \tilde{\psi} + A J_i\right)\right] \\ &= \exp\left(-\frac{\beta g}{2} N b^2\right) \prod_{i=1}^N 2\cosh\left(\beta g b + \psi \eta^1 + \tilde{\psi} + A J_i\right). \end{aligned}$$

Since the role of each $J_i$ is the same for all index $i$, we can consider just one $J$ and, in this way, we have

$$Z^{(P)}_{g,N}(\beta|\boldsymbol{\xi}^1, \boldsymbol{\lambda}, \boldsymbol{J}; t=0) = \exp\left(-\frac{\beta g}{2} N b^2\right) 2^N \cosh^N\left(\beta g b + \psi \eta^1 + \tilde{\psi} + A J\right). \tag{B25}$$

Inserting (B25) in the definition of the interpolating quenched statistical pressure (58) and applying the properties of the logarithm we get the desired one body terms.

## Appendix C Null-Temperature Limit Computations

Inspired by Amit's approach in [12], let us start by introducing the term $\beta y$ in the argument of the hyperbolic tangent in the self-consistency equations (73), one obtains:

$$\begin{aligned} \bar{\tilde{m}} &= \mathbb{E}_{x,\xi^1}\left\{(\xi^1 - b)\tanh\left[\beta\left(\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + x\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}} + y\right)\right]\right\}, \\ \bar{q} &= \mathbb{E}_{x,\xi^1}\left\{\tanh^2\left[\beta\left(\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + x\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}} + y\right)\right]\right\}, \\ \bar{M} &= \mathbb{E}_{x,\xi^1}\left\{\tanh\left[\beta\left(\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + x\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}}\right)\right]\right\}. \end{aligned} \tag{C26}$$

where $x$ is a standard Gaussian variable. We observe that in the limit $T \to 0$ (or equivalently $\beta \to \infty$), we get $\bar{q} \to 1$, allowing us to introduce the reparameterization

$$\bar{q} = 1 - \frac{\delta\bar{q}}{\beta}, \qquad \beta \to +\infty. \tag{C27}$$

Therefore, we get

$$\begin{aligned}
\bar{\tilde{m}} &= \mathbb{E}_{x,\xi^1}\left\{(\xi^1 - b)\tanh\left[\beta\left(\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + x\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}} + y\right)\right]\right\},\\
1 - \frac{\delta\bar{q}}{\beta} &= \mathbb{E}_{x,\xi^1}\left\{\tanh^2\left[\beta\left(\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + x\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}} + y\right)\right]\right\},\\
\bar{M} &= \mathbb{E}_{x,\xi^1}\left\{\tanh\left[\beta\left(\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + x\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}}\right)\right]\right\}.
\end{aligned} \tag{C28}$$

Now we highlight that if we derive the equation for $\bar{m}$ with respect to $y$ we get

$$\frac{\partial\bar{m}}{\partial y} = \beta\left[1 - \left(1 - \frac{\delta\bar{q}}{\beta}\right)\right] = \delta\bar{q}, \tag{C29}$$

namely

$$\begin{aligned}
\bar{\tilde{m}} &= \mathbb{E}_{x,\xi^1}\left\{(\xi^1 - b)\tanh\left[\beta\left(\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + x\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}} + y\right)\right]\right\},\\
\bar{M} &= \mathbb{E}_{x,\xi^1}\left\{\tanh\left[\beta\left(\frac{P}{2}\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + x\sqrt{\frac{\gamma P}{2}(1-b^2)^{P-2}\,\bar{q}^{P-1}}\right)\right]\right\},\\
\delta\bar{q} &= \frac{\partial\bar{m}}{\partial y}.
\end{aligned} \tag{C30}$$

For $\beta \to +\infty$ and for any $A,\ B$ constants we can say that,

$$\begin{aligned}
&\int_{-\infty}^{\infty} dz\,\frac{\exp\left(-\frac{1}{2}z^2\right)}{\sqrt{2\pi}}\tanh\left[\beta(Az+B)\right]\\
&\qquad = \int_{-\infty}^{\infty} dz\,\frac{\exp\left(-\frac{1}{2}z^2\right)}{\sqrt{2\pi}}\operatorname{sign}\left[\beta(Az+B)\right] = \operatorname{erf}\left(\frac{B}{\sqrt{2}A}\right).
\end{aligned} \tag{C31}$$

Therefore, this means for $\beta \to +\infty$

$$\bar{M} = \mathbb{E}_{\eta^1}\left[\operatorname{erf}\left(\frac{\frac{P}{2}\,\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + y}{\sqrt{\gamma P(1-b^2)^{P-2}}}\right)\right] \tag{C32}$$

$$\bar{m} = \mathbb{E}_{\eta_1}\left[\eta^1\operatorname{erf}\left(\frac{\frac{P}{2}\,\bar{\tilde{m}}^{P-1}(\xi^1 - b) + g(b-\bar{M}) + y}{\sqrt{\gamma P(1-b^2)^{P-2}}}\right)\right] \tag{C33}$$

$$\bar{q} = 1 \tag{C34}$$

Putting $y = 0$ and expanding the average with respect to $t$ completes the proof.

**Author Contributions** All authors contributed equally to this work.

**Funding** Open access funding provided by Università del Salento within the CRUI-CARE Agreement. L.A. acknowledges funding from the project "Patto Territoriale del Sistema Universitario Pugliese" (CUP F61B23000370006). L.A. acknowledges the PRIN 2022 grant *Statistical Mechanics of Learning Machines* number 20229T9EAT funded by the Italian Ministry of University and Research (MUR) in the framework of European Union - Next Generation EU. L.A. acknowledges funding also by the PRIN 2022 grant *"Stochastic Modeling of Compound Events (SLIDE)"* n. P2022KZJTZ funded by the Italian Ministry of University and Research (MUR) in the framework of European Union - Next Generation EU. A.A. acknowledges BULBUL "Brain-inspired ULtra-Fast & Ultra-sharp neural networks" for support via post-Lauream research fellowships "Statistical mechanics of hetero-associative neural networks" (CUP F85F21006230001, D.D. n. 325/29-09-2025). L.A. and A.A. are members of the GNFM group of INdAM which is acknowledged too.

**Data Availability** Data sharing is not applicable to this article as no datasets were generated or analyzed.

## Declarations

**Conflict of interest** The authors declare that they have no known competing financial or non-financial interests that could have appeared to influence the work reported in this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

## References

1. Agliari, E., De Marzo, G.: Tolerance versus synaptic noise in dense associative memories. Eur. Phys. J. Plus **135**(11), 1–22 (2020)
2. Agliari, E., Alemanno, F., Barra, A., Fachechi, A.: Dreaming neural networks: rigorous results. J. Stat. Mech: Theory Exp. **2019**(8), 083503 (2019)
3. Agliari, E., Fachechi, A., Marullo, C.: Nonlinear pdes approach to statistical mechanics of dense associative memories. J. Math. Phys. **63**(10), 103304 (2022)
4. Agliari, E., Leonelli, F.E., Marullo, C.: Storing, learning and retrieving biased patterns. Appl. Math. Comput. **415**, 126716 (2022)
5. Agliari, E., Albanese, L., Alemanno, F., Alessandrelli, A., Barra, A., Giannotti, F., Lotito, D., Pedreschi, D.: Dense hebbian neural networks: a replica symmetric picture of supervised learning. Phys. A **626**, 129076 (2023)
6. Agliari, E., Albanese, L., Alemanno, F., Alessandrelli, A., Barra, A., Giannotti, F., Lotito, D., Pedreschi, D.: Dense hebbian neural networks: a replica symmetric picture of unsupervised learning. Phys. A **627**, 129143 (2023)
7. Agliari, E., Alessandrelli, A., Barra, A., Centonze, M.S., Ricci-Tersenghi, F.: Generalized hetero-associative neural networks. J. Stat. Mech: Theory Exp. **2025**(1), 013302 (2025)
8. Albanese, L., Alessandrelli, A.: Boolean sherrington–kirkpatrick model: L. albanese, a. alessandrelli. Bollettino dell'Unione Matematica Italiana, pp. 1–11 (2025)
9. Albanese, L., Alemanno, F., Alessandrelli, A., Barra, A.: Replica symmetry breaking in dense hebbian neural networks. J. Stat. Phys. **189**(2), 1–41 (2022)
10. Albanese, L., Alessandrelli, A., Annibale, A., Barra, A.: Replica symmetry breaking in supervised and unsupervised hebbian networks. J. Phys. A: Math. Theor. **57**(16), 165003 (2024)
11. Albanese, L., Alessandrelli, A., Barra, A., Sollich, P.: Yet another exponential Hopfield model. Phys. A **2025**, 131223 (2025)
12. Amit, D.J.: Modeling Brain Function: The World of Attractor Neural Networks. Cambridge University Press, Cambridge (1989)
13. Amit, D.J., Gutfreund, H., Sompolinsky, H.: Storing infinite numbers of patterns in a spin-glass model of neural networks. Phys. Rev. Lett. **55**, 1530–1533 (1985)

14. Amit, D.J., Gutfreund, H., Sompolinsky, H.: Information storage in neural networks with low levels of activity. Phys. Rev. A **35**(5), 2293 (1987)
15. Baldi, P., Venkatesh, S.S.: Number of stable points for spin-glasses and neural networks of higher orders. Phys. Rev. Lett. **58**, 1 (1987)
16. Barra, A., Centonze, M.S., Solazzo, M.M., Tantari, D.: On the supra-linear storage in dense networks of grid and place cells. Preprint at http://arxiv.org/abs/2511.02441 (2025)
17. Bovier, A., Niederhauser, B.: The spin-glass phase-transition in the Hopfield model with p-spin interactions. Adv. Theor. Math. Phys. **5**, 1001–1046 (2001)
18. Fachechi, A., Agliari, E., Barra, A.: Dreaming neural networks: forgetting spurious memories and reinforcing pure ones. Neural Netw. **112**, 24–40 (2019)
19. Gardner, E.: Multiconnected neural network models. J. Phys. A: Gen. Phys. **20**, 1 (1987)
20. Guerra, F.: Broken replica symmetry bounds in the mean field spin glass model. Commun. Math. Phys. **233**, 1–12 (2003)
21. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. U.S.A. **79**, 2554–2558 (1982)
22. Kozachkov, L., Slotine, J.-J., Krotov, D.: Neuron-astrocyte associative memory. Proc. Natl. Acad. Sci. **122**(21), e2417788122 (2025)
23. Krotov, D., Hopfield, J.J.: Dense associative memory for pattern recognition. Adv. Neural. Inf. Process. Syst. **2016**, 1180–1188 (2016)
24. Krotov, D., Hopfield, J.: Dense associative memory is robust to adversarial inputs. Neural Comput. **30**, 3151–3167 (2018)
25. Lowe, M.: On the storage capacity of the Hopfield model with biased patterns. IEEE Trans. Inf. Theory **45**(1), 314–318 (1999)
26. Lucibello, C., Mézard, M.: Exponential capacity of dense associative memories. Phys. Rev. Lett. **132**(7), 077301 (2024)
27. Mézard, M., Parisi, G., Virasoro, M.A.: Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications, vol. 9. World Scientific Publishing Company (1987)
28. Nishimori, H.: Statistical Physics of Spin Glasses and Information Processing: An Introduction, vol. 111. Clarendon Press (2001)
29. Peretto, P.: Collective properties of neural networks: a statistical physics approach. Biol. Cybern. **50**(1), 51–62 (1984)
30. Ramsauer, H., Schäfl, B., Lehner, J., Seidl, P., Widrich, M., Adler, T., Gruber, L., Holzleitner, M., Pavlović, M., Sandve, G.K.: Hopfield networks is all you need. Preprint at http://arxiv.org/abs/2008.02217 (2020)
31. Tsodyks, M.V., Feigel'man, M.V.: The enhanced storage capacity in neural networks with low activity level. EPL **6**(2), 101–105 (1988)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.