Effect of spatial correlations on Hopfield Neural Network and Dense Associative Memories

Giordano De Marzo\textsuperscript{1,2,3,4\textsuperscript{*}} and Giulio Iannelli\textsuperscript{1,5\textsuperscript{†}}

\textsuperscript{1}Centro Ricerche Enrico Fermi, Piazza del Viminale, 1, I-00184 Rome, Italy.
\textsuperscript{2}Dipartimento di Fisica Università “Sapienza”, P.le A. Moro, 2, I-00185 Rome, Italy.
\textsuperscript{3}Sapienza School for Advanced Studies, “Sapienza”, P.le A. Moro, 2, I-00185 Rome, Italy.
\textsuperscript{4}Complexity Science Hub Vienna, Josefstädter Strasse 39, 1080, Vienna, Austria, and
\textsuperscript{5}Dipartimento di Fisica, Università di Roma “Tor Vergata”, 00133 Roma, Italy.

(Dated: July 13, 2022)

Hopfield model is one of the few neural networks for which analytical results can be obtained. However, most of them are derived under the assumption of random uncorrelated patterns, while in real life applications data to be stored show non-trivial correlations. In the present paper we study how the retrieval capability of the Hopfield network at null temperature is affected by spatial correlations in the data we feed to it. In particular, we use as patterns to be stored the configurations of a linear Ising model at inverse temperature $\beta$. Exploiting the signal to noise technique we obtain a phase diagram in the load of the Hopfield network and the Ising temperature where a fuzzy phase and a retrieval region can be observed. Remarkably, as the spatial correlation inside patterns is increased, the critical load of the Hopfield network diminishes, a result also confirmed by numerical simulations. The analysis is then generalized to Dense Associative Memories with arbitrary odd-body interactions, for which we obtain analogous results.

Keywords: Hopfield neural network; dense associative memories; signal to noise analysis

I. INTRODUCTION

The Hopfield Neural Network (HNN) \cite{1} stands as a cornerstone in the history of Artificial Neural Networks (ANN) and the scientific attempt of modeling brain functions \cite{2,3}. In particular the HNN is a recurrent neural network composed of $N$ neurons or spins which interact pairwise and works as an Associative Memory (AM), meaning that it can store and then retrieve information.

A central point in the study of the HNN is understanding the maximal amount of data this system can successfully memorize \cite{4}. It has been shown that an HNN with $N$ neurons can store $K = \mathcal{O}(N)$ i.i.d. random patterns, each composed of $N$ i.i.d. Rademacher variables, if some extent of faults is tolerated \cite{5}. The storage capacity $\alpha = K/N$ of the HNN can thus be of order one and the neural network still keeps working. More precisely it has been shown \cite{5} that retrieval is possible up to $\alpha \leq \alpha_c$ with $\alpha_c \approx 0.14$. Heuristically this can be understood noticing that the interaction matrix of the HNN is made of $\mathcal{O}(N^2)$ independent components and so the maximal amount of bits which can be stored must be of the same order of magnitude.

In order to increase the storage capacity of such system some adaptations of the HNN has been proposed (e.g. \cite{6,7}). In recent years \cite{8,9} much attention has been devoted to generalized Hopfield networks with $p$-body interactions; this type of neural networks are called Dense Associative Memorises (DAMs). The interaction matrix of a DAM with $p$-body interactions contains $\mathcal{O}(N^p)$ components, so that, in principle, such network can store up to $N^{p-1}$ i.i.d. Rademacher patterns \cite{10,11}. These architectures have been proven not only to have a larger storage capacity, but also to be very robust against noise, faults and adversarial examples \cite{12–14}.

However patterns with random uncorrelated components, while representing an appealing standard for analytical studies, do not properly embody realistic memories. This point is very relevant since HNN and DAMs has been shown \cite{15–19} to be formally equivalent to Boltzmann machines \cite{20,21}, which are used in a number of machine learning applications where data are definitely not i.i.d. Rademacher patterns. Indeed in most real life situations patterns can be correlated one with the other or they may be structured, thus showing non trivial spatial correlations \cite{22–24}. This last situation occurs, for instance, if the patterns to be stored are texts, which can be represented as binary uni-dimensional sequences whose entries are correlated due to grammatical and semantic rules. Analogously black and white images are binary matrices with spatial correlations since a black (or white) pixel is more likely to be surrounded by pixels of the same kind. However, despite the importance of patterns with such kind of spatial correlations, very few studies considered how HNN behaves in this situation. In \cite{22} it has been shown that, if the spatial correlations arise from a simple two states Markov chain, than the HNN still exhibits a critical capacity of order one. More realistic correlations, arising from a two-dimensional Ising model, have been studied in \cite{25}, but for the case of exact retrieval of the patterns, thus providing only a lower bound to the true storage capacity. The effects of correlations on DAMs have been considered recently \cite{26} by using patterns generated by a Curie-Weiss model, but again only for a perfect retrieval of the patterns.

In the present work we considered how the HNN behaves in presence of patterns generated by a one dimensional Ising model (or Ising linear chain) at inverse temperature $\beta$. The spin configurations of this model can be regarded as prototypes of textual data. Note that by varying the temperature of the Ising model we can control the level...
of internal correlation of the patterns, while we set the temperature of the Hopfield model to zero, thus neglecting the effects of fast noise. In the rest of the work we refer to $\beta$ as the temperature parameter relative to the Ising model.

By using a Signal–to–Noise Analysis (SNA) [2] we assess the maximal storage capacity of the HNN as function of $\beta$, finding that for $\beta > 0$ the neural network can still store a number of patterns which grows linearly with the number of neurons $N$, even if the storage capacity is lower than in the case of i.i.d. Rademacher patterns. This result is non trivial since naively one would expect the storage capacity to increase due to the fact that internal correlations diminish the amount of information to be stored. Moreover, differently from previous studies, we obtain the true maximal storage capacity and not a lower bound. This allows to draw a phase diagrams in the space defined by the parameters $(\alpha, \beta)$, where a retrieval region and a fuzzy phase can be identified. The analysis is then generalized to the DAM version of the HNN with odd $p$–bodies interaction.

We show that the maximal storage capacity decreases as $\beta$ increases, but again the network can memorize $O(N^p)$ patterns as in the uncorrelated case, provided that $\beta > 0$.

The paper is structured as follows:

- In Sec. II we introduce the signal to noise technique and we apply it to the HNN and to the DAM with i.i.d. Rademacher patterns.
- In Sec. III we apply the SNA to the HNN feed with patterns generated by a one dimensional Ising model and we compare the analytical results with numerical simulations.
- In Sec. IV we generalize the results of the previous section to the case of a DAM with $p$–bodies interaction.
- Finally Sec. V is devoted to our conclusions and remarks.

II. SIGNAL TO NOISE TECHNIQUE FOR UNCORRELATED PATTERNS

The Hopfield Neural Network is made of $N$ neurons associated with binary state variables, i.e. $\sigma_i = \pm 1$ for $i = 1 \cdots N$, and interacting pairwise on a complete graph with coupling matrix $J_{ij}$. Denoting by $\xi^\mu_i$ with $\xi^\mu_i = \pm 1$ the $K$ binary patterns we want to store in the neural network, the coupling matrix is defined by the so called Hebbian rule, which reads

$$\forall i \neq j \quad J_{ij} = \frac{1}{N} \sum_{\mu=1}^{K} \xi^\mu_i \xi^\mu_j$$  (1)

and $J_{ii} = 0$ for all the diagonal elements, avoiding self interactions. The Hamiltonian of the system thus satisfies

$$H[\sigma] = -\sum_{i,j} J_{ij} \sigma_i \sigma_j = -\sum_{\mu=1}^{K} \sum_{i,j \neq i} \xi^\mu_i \xi^\mu_j \sigma_i \sigma_j. \quad (2)$$

In the following we will often avoid to explicitly write $i \neq j$, bearing in mind that the neurons have no self–interaction. If one neglects fast noise by considering the model at null temperature the evolution of the system is governed by the following dynamical rule

$$\sigma_i(t + 1) = \text{sign}[h_i(\sigma(t))], \quad (3)$$

where $h_i(\sigma(t))$ is the local field acting on the $i$th neuron at time $t$ and satisfying

$$h_i(\sigma(t)) = \frac{1}{N} \sum_{\mu=1}^{K} \sum_{j \neq i} \xi^\mu_i \xi^\mu_j \sigma_j(t). \quad (4)$$

Note that the Hamiltonian Eq. (2) can be rewritten in terms of these local fields as $H[\sigma] = -\sum_i h_i(\sigma)\sigma_i$ and the dynamics defined by Eq. (3) is equivalent to a steepest descent algorithm whose Lyapunov function is the Hamiltonian. Moreover from Eq. (3) it follows that the state of a given spin $\sigma_i$ is unaltered by the dynamics if it holds $\text{sign}[\sigma_i] = \text{sign}[h_i(\sigma)]$, that is if it holds $\sigma_i h_i(\sigma) > 0$. As a consequence a pattern, let us say $\xi_i^1$, is a fixed point of the dynamics and is thus dynamically stable provided that its components satisfy

$$\xi_i^1 h_i(\xi^1_i) = \frac{1}{N} \sum_{\mu=1}^{K} \sum_{j>1} \xi^\mu_i \xi^\mu_j \xi_i^1 \xi_j^1 > 0. \quad (5)$$

Here and in the following we focus, without loss of generality, on the first pattern. However this quantity, being a sum of random variables, is itself a random variable and so, in order to assess if it is positive, we have to determine whether its expectation is positive and larger than its standard deviation. The first quantity is the signal, while the latter is the noise affecting it. This explaining where the name “signal to noise” stems from. We denote by $\langle \cdot \rangle$ the expectation on the probability distribution of the $\xi_i^\mu$ for fixed $\mu$ and by $\text{Var}[\cdot]$ the variance computed over the same distribution. In these terms patterns are dynamically stable if it holds

$$\sqrt{C} \langle \xi_i^1 h_i(\xi^1_i) \rangle \geq \sqrt{\text{Var}[\xi_i^1 h_i(\xi^1_i)]}, \quad (6)$$

where $C$ is a numerical constant fixing the level of tolerated faults. Note that by explicitly writing the variance this condition is equivalent to

$$(C + 1) \langle \xi_i^1 h_i(\xi^1_i) \rangle^2 \geq \langle \xi_i^1 h_i(\xi^1_i) \rangle^2. \quad (7)$$

At this point we have all the ingredient for applying the Signal–to–Noise Analysis technique to the HNN fed with i.i.d. Rademacher patterns. First we compute the mean value appearing in Eq. (6)

$$\langle \xi_i^1 h_i(\xi^1_i) \rangle = \frac{1}{N} \sum_{\mu=1}^{K} \sum_{j>1} \langle \xi_i^\mu_j \xi_i^\mu_j \xi_i^1 \xi_j^1 \rangle =$$

$$= \frac{N - 1}{N} + \frac{1}{N} \sum_{\mu=2}^{K} \sum_{j>1} \langle \xi_i^\mu_j \langle \xi_j^\mu_j \rangle \langle \xi_i^1 \rangle \langle \xi_j^1 \rangle \rangle = 1, \quad (8)$$
where we implicitly took the limit of large $N$ and we used the fact that in the case under consideration, since there are no correlations, the expectation factorizes. Moreover, being $P(\xi_i^\mu = +1) = P(\xi_i^\mu = -1) = 1/2$, it holds $\langle \xi_i^\mu \rangle = 0$. Then, in order to obtain the variance, we first compute the expectation of $\langle \xi_i^\mu h_1(\xi^1) \rangle^2$. By noticing that

$$\langle \xi_i^\mu h_1(\xi^1) \rangle^2 = \left( \frac{N - 1}{N} + \frac{1}{N} \sum_{j > 1} \sum_{\mu, p > 1} \xi_i^\mu \xi_j^\nu \xi_i^\rho \xi_j^\sigma \right)^2$$

$$= 2 h_1^2 - 2 \sum_{j > 1} \sum_{\mu, p > 1} \xi_i^\mu \xi_j^\nu \xi_i^\rho \xi_j^\sigma$$

and by using Eq. (8) we get

$$\langle \xi_i^\mu h_1(\xi^1) \rangle^2 = 1 + \frac{1}{N^2} \sum_{j > 1} \sum_{\mu, p > 1} \xi_j^\nu \xi_i^\rho \xi_i^\sigma \xi_j^\rho \xi_i^\sigma = 1 + (K - 1)(N - 1) = 1 + \frac{K - 1}{N}.$$  

Here we used the fact that the only terms surviving the expectation are those for which $j = k$ and $\mu = \rho$. Combining this expression with Eq. (8) we can finally write the variance

$$\text{Var} [\xi_i^\mu h_0(\xi^1)] = \langle \xi_i^\mu h_1(\xi^1)^2 \rangle - \langle \xi_i^\mu h_1(\xi^1) \rangle^2 = \frac{K - 1}{N}.$$  

Plugging this relation and Eq. (8) into the condition for the dynamical stability Eq. (6) we thus obtain that the patterns are dynamically stable provided that $C \geq K - 1/N$ that is if

$$\alpha \leq C. $$  

Here we introduced the load $\alpha = K/N$ and we considered that, in the thermodynamic limit, $K \gg 1$. Since $C$ is a numerical constant we thus see that the network can successfully store up to $K = O(N)$ patterns before they stop to be fixed points of the dynamics, recovering, as expected, what previously proven [2, 5].

In the vary same way we can analyze also the corresponding DAM with $p$-bodies interaction, whose Hamiltonian, in analogy with Eq. (2), is

$$H^{(\rho)}[\sigma] = - \sum_{i_1, \ldots, i_p} J_{i_1 \ldots i_p} \sigma_{i_1} \ldots \sigma_{i_p} = - \frac{1}{N^{p-1}} \sum_{\mu = 1}^K \sum_{i_1 \ldots i_p} \prod_{k = 1}^p \xi_{i_k}^\mu \sigma_k.$$  

We can then define the field $h_1^{(\rho)}(\sigma)$ acting on the $i$th spin at time $t$ as

$$h_1^{(\rho)}(\sigma(t)) = - \frac{1}{N^{p-1}} \sum_{\mu = 1}^K \sum_{i_2, \ldots, i_p} \prod_{k = 2}^p \xi_{i_k}^\mu \sigma_k(t)$$  

and in these terms the dynamical rule determining the evolution on the neural network as the same form as Eq. (3), that is

$$\sigma_i(t + 1) = \text{sign} [h_1^{(\rho)}(\sigma(t))].$$

As a consequence all we have to do is to repeat the computation we did in the case of the HNN and use Eq. (6) to assess the dynamical stability of the patterns. The only difference is that this time the local field $h_1^{(\rho)}$ is given by Eq. (13) instead of Eq. (4). Focusing on $\xi_1^\mu$ and doing manipulations analogous to those already done for the HNN (corresponding to $p = 2$), we obtain

$$\langle \xi_1^\mu h_1^{(\rho)}(\xi^1) \rangle = 1,$$

and

$$\text{Var} [\xi_1^\mu h_1^{(\rho)}(\xi^1)] = (p - 1)! \frac{K - 1}{N^{p-1}}.$$  

Putting these equations into the condition for the dynamical stability we get $C \geq (p - 1)!K - 1/(N^{p-1})$, which means that Eq. (11) rewrites as

$$\alpha^{(p)} \geq \frac{C}{(p - 1)!}.$$  

where we defined the load of a $p$-bodies DAM as

$$\alpha^{(p)} = K/N^{p-1}.$$  

We thus see that in this case, as mentioned above and already proven, the neural network can store up to $O(N^{p-1})$ patterns.

III. HOPFIELD NEURAL NETWORK WITH CORRELATED PATTERNS

In the previous section we showed that when HNN behaves when fed with uncorrelated Rademacher patterns, but as aforementioned real data are structured and characterized by non trivial spatial correlations. Examples are text and images, both playing a central role in a huge number of machine learning applications.

It is then very natural to assess what happens to the HNN, and so also to the formally equivalent Boltzmann machine, when the patterns have internal correlations (but different patterns are uncorrelated). In order to do so, we exploit the one dimensional Ising model, that we use to generate one dimensional correlated data which can be regarded as prototypes of texts. More precisely, each pattern is obtained as a configuration of the Ising model at inverse temperature $\beta$. By tuning $\beta$ we can then easily control the level of correlation inside the data. This means that now the probability distribution over which we make the expectation is no more the trivial Rademacher distribution $P(\xi_i^\mu = +1) = P(\xi_i^\mu = -1) = 1/2$, but it is instead the Boltzmann one. More precisely it holds

$$P(\xi^\mu) = \frac{1}{Z} e^{-\beta H[\xi^\mu]}$$

where $Z$ is the partition function.
where $Z$ is the partition function and $H_I[\xi^\mu]$ is the one dimensional Ising Hamiltonian
\[ H_I[\xi^\mu] = -\frac{1}{2} \sum_{i,j \neq i} \xi^\mu_i \xi^\mu_j. \]

As a consequence, when taking the expectations, we now have to consider also correlations since the probability distribution does not factorizes. However, it is easy to show, for instance by using the transfer matrix formalism \cite{27}, that in the 1D Ising model only the two points correlator is non-trivial, while all correlations involving an odd number of points are null
\[ \langle \xi^\mu_i \xi^\mu_j \rangle = e^{-|i-j|/L}, \]
\[ \langle \xi^\mu_i \rangle = 0, \]
where $L = L(\beta)$ is the correlation length at inverse temperature $\beta$ and satisfies
\[ L = \frac{1}{\ln[\tanh(\beta)]}. \]

Having made these considerations, the procedure for assessing the dynamical stability of patterns is the same depicted in Sec. II and so we have to compute the mean value and the variance of $\xi^\mu_i h_i(\xi^\mu)$. By using Eq. (4) we have that the mean value satisfies
\[ \langle h_1 \xi^\mu_i \rangle = 1 + \frac{1}{N} \sum_{j=1}^{N} \sum_{\xi^\mu_j} \langle \xi^\mu_i \xi^\mu_j \xi^\mu_i \xi^\mu_j \rangle = 1 + \frac{1}{N} \sum_{j=1}^{N} \langle \xi^\mu_i \xi^\mu_j \rangle \langle \xi^\mu_i \xi^\mu_j \rangle, \]
where we used the fact that different patterns are uncorrelated. Exploiting the expression of the two point correlations Eq. (18) we get
\[ \langle h_1 \xi^\mu_i \rangle = 1 + \frac{K}{N} \sum_{j=1}^{N} \exp \left[-\frac{2j-1}{L} \right] \]
\[ = 1 + \frac{K-1}{N} \left\{ 1 + \sum_{x=0}^{N-1} \exp \left[-\frac{2x}{L} \right] \right\}, \]
where we defined $x = j - 1$. In the limit of large $N$, this yields
\[ \langle h_1 \xi^\mu_i \rangle = 1 + \frac{K-1}{N} \frac{1}{1 - e^{-2/L}} = 1 + \frac{K-1}{N} \frac{1}{e^{2/L} - 1}. \]
\[ \text{(20)} \]

\footnote{Note that we already made the approximation $\frac{N-1}{N} = 1$ since we are interested in the large $N$ limit.}
The quantity \( \langle (h_1 \xi_1^2) \rangle \) can obtained by very similar passages, but since the computation is quite lengthy we only report the final result; the interested reader can find the detailed computation in Appendix A. It holds

\[
\begin{align*}
\langle (h_1 \xi_1^2) \rangle &= \frac{K - 1}{N} + 4 \frac{K - 1}{N} \frac{1}{e^{2L/1 - 1}} + \\
&\quad + \frac{(K - 1)(K - 2)}{N^2} \frac{1}{(e^{2L/1 - 1})^2} + 1. \\
\end{align*}
\]

By putting this expression and Eq. (20) into the condition for dynamical stability given by Eq. (7) we conclude that patterns can be retrieved up to a critical correlation length \( L_c \) satisfying

\[
C - \alpha + \frac{1}{e^{2L_c/1 - 1}} = 0.
\]

We recall that here \( \alpha = K/N \) is the load of the Hopfield neural network. Solving for \( L_c \) this yields

\[
L_c = \frac{2}{\log \left[ \frac{\alpha^2 + \alpha \sqrt{\alpha^2 - 4C + 2C - 2\alpha C}}{2(C - \alpha)} \right]}
\]

and so the inverse temperature \( \beta_c \) of the one dimensional Ising model above which retrieval breaks due to spatial correlations is

\[
\beta_c = \tanh^{-1} \left[ \frac{2(C - \alpha)}{\sqrt{\alpha^2 + \alpha \sqrt{\alpha^2 - 4C + 2C - 2\alpha C}}} \right].
\]

Now, if we set \( \beta_c = 0 \) we return to uncorrelated Rademacher patterns and so to the original HNN for which the maximal load is \( \alpha_c \approx 0.14 \). As a consequence we can fix \( C \) by imposing \( T_c(\alpha_c) = \infty \), which gives \( C = \alpha_c \approx 0.14 \). In conclusion the \( \beta_c \) above which Hopfield neural network can no more handle the spatial correlation of patterns is

\[
\beta_c(\alpha) = \tanh^{-1} \left[ \frac{2(\alpha_c - \alpha)}{\sqrt{\alpha^2 + \alpha \sqrt{\alpha^2 - 4\alpha_c + 2\alpha_c - 2\alpha \alpha_c}}} \right].
\]

This last expression allows us to draw a phase diagram in the variables \( \beta, \alpha \) dividing a fuzzy phase from a retrieval phase. This is shown in Fig. 1, where we also plotted the results of numerical simulations confirming the validity of Eq. (22). As we mentioned in the introduction, when data show internal correlation they carry a lower amount of information. One would thus expect them to be easier to store and so the capacity of the neural network to increase, while the results we obtained go in the opposite direction. This apparently paradoxical behavior can be explained noticing that also the coupling matrix, which is defined starting from the patterns themselves, is less complex than in the uncorrelated case and thus if on one side data are easier to memorize, on the other the neural network has de facto less resources to handle them.

### IV. DENSE ASSOCIATIVE MEMORIES WITH CORRELATED PATTERNS

We now move to the case of Dense Associative Memories storing correlated patterns generated by the 1D Ising model at temperature \( \beta \). The procedure is analogous to that followed in Sec. II, but as done in Sec. III we now have to consider correlations among the different components of the patterns. We only consider an Hamiltonian with interactions involving an odd number of neurons since in this way both computations and numerical simulations are easier to perform. As shown in the Appendix B in the case under consideration it can be proven that

\[
\langle h_1(p) \xi_1 \rangle = 1
\]

and

\[
\begin{align*}
\langle (h_1(p) \xi_1^2) \rangle &= 1 + \sum_{n=0}^{p-1} \frac{K - 1}{N^{2n-2-n}} \frac{[p - 1]!^2}{n![(p - n - 1)!]^2} \\
&\quad \left[ \frac{2}{e^{2L_c/1 - 1}} \right]^{p-n-1}. \\
\end{align*}
\]

From Eq. (7) we thus obtain that retrieval is possible up to a maximal correlation length \( L_c \) satisfying

\[
K - \frac{1}{N^{p-1}} \sum_{n=0}^{p-1} \frac{[p - 1]!^2}{n![(p - n - 1)!]^2} \left[ \frac{2}{e^{2L_c/1 - 1}} \right]^{p-n-1} = C.
\]

In the high load case \( K = \alpha N^{p-1} \) this yields

\[
\sum_{n=0}^{p-1} \frac{[p - 1]!^2}{n![(p - n - 1)!]^2} \left[ \frac{2}{e^{2L_c/1 - 1}} \right]^{p-n-1} = C.
\]

Analogously to the case \( p = 2 \) the constant \( C \) can be determined requiring \( L_c \) to be zero when \( \alpha = \alpha_c^{(p)} \). Here by \( \alpha_c^{(p)} \) we denote the critical load of the \( p \)-spin DAM with uncorrelated patterns above which retrieval breaks. Note that sometimes the load of DAM is defined as \( \alpha = K/N \) as in the Hopfield case, but we prefer to set \( \alpha = K/N^{p-1} \) in order to make it possible to make a comparison for different values of \( p \). For \( L_c = 0 \), the patterns becomes spatially uncorrelated and so only the term with \( n = p - 1 \) should be considered, as a consequence we get

\[
(p - 1)! = \frac{C}{\alpha_c^{(p)}} \rightarrow C = (p - 1)!\alpha_c^{(p)}.
\]

In conclusion our implicit expression for \( L_c \) is

\[
\sum_{n=0}^{p-1} \frac{[p - 1]!}{n![(p - n - 1)!]^2} \left[ \frac{2}{e^{2L_c/1 - 1}} \right]^{p-n-1} = \alpha_c^{(p)}/\alpha.
\]

For \( p = 3 \) this last equation can be solved analytically, indeed we have

\[
\frac{2}{(e^{2L_c/1 + 1})^2} + \frac{4}{e^{2L_c/1 + 1}} + 1 = \frac{\alpha_c^{(3)}}{\alpha}.
\]
Figure 2. Numerical simulations for a $p = 3$ DAM endowed with spatially correlated patterns. We simulated a DAM with $p = 3$-body interaction, $N = 100$ neurons and varying load $\alpha$ under the dynamics Eq. (3) and using as starting state one of the patterns $\sigma = \xi^\mu$ generated using a linear Ising chain at temperature $\beta$. We then collected the Mattis magnetization $m_n$ for each possible starting pattern, we repeated the process $M = 3$ times and we computed the average magnetization ($m^\mu$). For any $\alpha$ we computed the critical $\beta_c^{(3)}$ above which retrieval breaks as the value for which the average magnetization is equal to Abbott’s magnetization $m = 0.807$ [28], which is the equivalent of Sompolinsky magnetization for the case $p = 3$. Since for $p = 3$ the critical load $\alpha_c^{(3)}$ is unknown, we left it as a free parameter and we fitted the theoretical prediction Eq. (23) to simulated data (dashed black line). Uncertainties have been computed as the semi-difference between the last point above Abbott’s threshold and the one just below.

whose solution is

$$L_c^{(3)} = \frac{2}{\log \left( \frac{\alpha_c^{(3)} + \sqrt{(2\alpha_c^{(3)} + \alpha^2 + \alpha)}}{\alpha_c^{(3)} - \alpha} \right)}.$$

The critical inverse temperature $\beta_c^{(3)}$ is therefore

$$\beta_c^{(3)} = \tanh^{-1} \left[ \frac{\alpha_c^{(3)} - \alpha}{\sqrt{(2\alpha_c^{(3)} + \alpha^2 + \alpha)}} \right].$$

Again we have an expression allowing us to sketch a phase diagram in the variables $\beta, \alpha$ dividing, as for the Hopfield case, a fuzzy phase from a retrieval phase. This is shown in Fig. 2, where we also plotted the results of numerical simulations. Note that while for $p = 2$ an analytical estimate of $\alpha_c$ as been computed, for the case $p = 3$ only approximated values are available. As a consequence we decided to leave $\alpha_c^{(3)}$ as a free parameter, which we then estimated by fitting Eq. (23) to the simulated data. Results are shown in Fig. 2a and from this procedure we obtain $\alpha_c^{(3)} \approx 0.18$. This value is larger than analytical results obtained under the replica symmetry assumption and confirms the expected increase of $\alpha_c^{(3)}$ as when also replica symmetry breaking is taken into account [28].

V. CONCLUSIONS

Artificial neural networks are nowadays used in uncountable applications, but despite the great interest they gathered we are still lacking a general theory describing how and why they work. Understanding the mechanism behind their ability in handling complex problems is thus a central issue and many scientists are working in this direction. Hopfield neural network is one of the few examples where analytical results can be obtained and great effort has been devoted to draw phase diagrams for this particular system. Indeed, despite it being very simple if compared to modern deep neural networks, the Hopfield model still show a very rich phenomenology characterized by different phase transitions. Moreover, this neural network has also been shown to be equivalent to Restricted Boltzmann Machines, which are used in a number of real life applications. However, in order to make the model feasible to be analyzed with the tools of statistical mechanics, one often has to make strong assumption on the nature of the information to be stored. Indeed, one generally considers the case of i.i.d. random data, which is quite an unrealistic scenario since in most real life situations patterns are correlated one with the other or they are structured, thus showing spatial correlations.

In the present paper we considered how the HNN behaves when fed with spatially correlated data, that we generated from a linear Ising chain with inverse temperature $\beta$. These synthetic patterns can be seen as prototypes of textual data and their spatial correlation can be tuned by changing the temperature of the Ising chain. By using a signal to noise approach, we obtain an analytical expression for the critical inverse temperature $\beta_c^{(3)}$ above which correlation becomes too strong and retrieval breaks. This allows to draw a phase diagram in the variables $\alpha, \beta$ where a fuzzy phase and a retrieval region can be identified. Remarkably, we observe both numerically and analytically that the introduction of spatial correlations diminishes the critical load of the Hopfield network, while one would naively expect the opposite to occur. These results are then extended to Dense Associative Memories, namely generalized Hopfield networks with p-body interactions. In particular we consider the case of odd $p$ and we derive again an expression for the critical inverse temperature $\beta_c^{(3)}$ as. As in the $p = 2$ case corresponding to the HNN, we observe the critical load to diminish when spatial correlations are increased, a result which is also confirmed by numerical simulations performed for a DAM with $p = 3$.

Spatial mono-dimensional correlations, typical of textual data, are only an example of the possible structure real patterns can show: as also mentioned in the introduction, another straightforward example is given by two dimensional correlations in images. As the 1d Ising chain con-
figurations can be used as prototypes of textual data, in the very same way synthetic spatially correlated images could be generated by means of a two dimensional Ising model. Extending the analysis here performed to such a case would be a very interesting problem to investigate in a future publication, also considering the non-trivial n-body correlation emerging in the 2d Ising model.

ACKNOWLEDGMENTS

We are grateful to Elena Agliari, Adriano Barra and Francesco Alemanno for their useful comments and discussion. We acknowledge the project “Social and Economic Complexity” of Enrico Fermi Research Center.

Appendix A: Detailed computations for HNN

Here we show how to compute the quantity $\langle (h_1 \xi_1)^2 \rangle$ feed with one dimensional Ising configurations. By virtue of Eq. (9) we have

$$\langle (h_1 \xi_1)^2 \rangle = \frac{2h_1^2 \xi_1^2}{A} + \frac{1}{N^2} \sum_{j,k=1}^{N} \sum_{\mu, \rho = 2}^{K} \xi_1^\mu \xi_1^\rho \xi_1^\mu \xi_1^\rho \xi_1^\mu \xi_1^\rho \xi_1^\mu \xi_1^\rho$$

The expectation of the first term can be easily computed using Eq. (20)

$$\langle A \rangle = 2 \frac{K - 1}{N} \frac{1}{e^{2L} - 1} + 1.$$  \quad (A1)

For what concerns term $B$ we first have to get rid of self-interactions. In particular we rewrite it as

$$B = \frac{1}{N^2} \sum_{\mu, \rho = 2}^{ \mu, \rho = 2} \left[ \sum_{j, \mu = 2}^{N} \xi_1^\mu \xi_1^\rho + \sum_{j, \mu = 2}^{N} \xi_1^\rho \xi_1^\mu \right]$$

and going on we get

$$B = \frac{1}{N^2} \left\{ \sum_{\mu, \rho = 2}^{K} \left[ \sum_{j, \mu = 2}^{N} \xi_1^\mu \xi_1^\rho + \sum_{j, \mu = 2}^{N} \xi_1^\rho \xi_1^\mu \right] + \sum_{\mu, \rho = 2}^{K} \sum_{j, \mu = 2}^{N} \xi_1^\mu \xi_1^\rho \xi_1^\mu \xi_1^\rho \xi_1^\mu \xi_1^\rho \right\}.$$  \quad (A2)
We thus have four terms

\[ B = \left( \frac{K - 1}{N} \right) (N - 1) + \frac{1}{N^2} \sum_{\mu \neq 1, \mu \neq 1} \sum_{j > 1} \sum_{k \neq 1, j} \xi^\mu \xi^j \xi^1 \xi^k + \frac{1}{N^2} \sum_{\mu \neq 1, \mu \neq 1} \sum_{j > 1} \sum_{k \neq 1, j} \xi^\mu \xi^j \xi^1 \xi^k \]

\[ + \frac{1}{N^2} \sum_{\mu \neq 1, \mu \neq 1} \sum_{j > 1} \sum_{k \neq 1, j} \xi^\mu \xi^j \xi^1 \xi^k + \frac{1}{N^2} \sum_{\mu \neq 1, \mu \neq 1} \sum_{j > 1} \sum_{k \neq 1, j} \xi^\mu \xi^j \xi^1 \xi^k \]

The average of term I is trivial, for what concerns term II we have

\[ \langle II \rangle = \frac{1}{N^2} \sum_{\mu \neq 1, \mu \neq 1} \sum_{j > 1} \sum_{k \neq 1, j} \langle \xi^\mu \xi^j \rangle \langle \xi^1 \xi^k \rangle = \frac{K - 1}{N^2} - \left( -\frac{2}{L} \right)^j = 2^{K - 1} \cdot \frac{1}{N^2} \sum_{j > 1} \sum_{k > j} \exp \left( -\frac{2}{L} (j - k) \right). \]

Defining \( x = k - j \) we get

\[ \langle II \rangle = \frac{2}{N^2} \sum_{j = 1}^{N - 1} \left[ -1 + \sum_{x = 0}^{N - j} \left( e^{-2/L} \right)^x \right] = \frac{2}{N^2} \sum_{j = 2}^{N - 1} \frac{e^{2(j - N)/L} - 1}{e^{2/L} - 1} = \frac{2}{N^2} \cdot \frac{1}{1 - e^{2/L}} \left[ -N^2 - 2 \sum_{j = 0}^{N - 2} \left( e^{-2/L} \right)^j \right], \]

where we defined \( y = N - j \). For \( N \to \infty \) the sum, being convergent, is a negligible correction and therefore we can write

\[ \langle II \rangle = \frac{2}{N^2} \cdot \frac{1}{1 - e^{2/L}} \frac{e^{-2/N} - 1}{1 - e^{-2/L}}. \]

Going on to term III we have

\[ \langle III \rangle = \frac{1}{N^2} \sum_{\mu \neq 1, \mu \neq 1} \sum_{j > 1} \sum_{k \neq 1, j} \langle \xi^\mu \xi^j \rangle \langle \xi^1 \xi^k \rangle = \frac{(K - 1) (K - 2)}{N^2} \sum_{j = 2}^{N - 1} \sum_{k > j} \exp \left( -\frac{2}{L} (j - k) \right). \]

that is, for large \( N \)

\[ \langle III \rangle = \frac{(K - 1) (K - 2)}{N^2} \cdot \frac{1}{1 - e^{2/L}}. \]

Finally we have term IV

\[ \langle IV \rangle = \frac{1}{N^2} \sum_{\mu \neq 1, \mu \neq 1} \sum_{j > 1} \sum_{k \neq 1, j} \langle \xi^\mu \xi^j \rangle \langle \xi^1 \xi^k \rangle \]

\[ = \frac{2}{N^2} \sum_{j = 2}^{N - 1} \sum_{k > j} \exp \left( -\frac{2}{L} (j - k) \right) e^{-(k - 2)/L} e^{-(j - 2)/L}. \]

Proceeding as done before and considering \( N \gg 1 \) one finds

\[ \langle IV \rangle = \frac{2}{N^2} (K - 1) (K - 2) \frac{1}{(e^2 - 2) (L - 1)^2}. \]

Exploiting Eqs. Eq. (A2), Eq. (A3), Eq. (A4) and Eq. (A5), the expectation of \( B \) is then

\[ \langle B \rangle = \frac{K - 1}{N} + \frac{2}{N^2} \sum_{j = 2}^{N - 1} \left( \frac{1}{e^{2/L} - 1} \right) \left( \frac{1}{(e^{2/L} - 1)^2} \right) + \frac{(K - 1) (K - 2)}{N^2} \frac{1}{(e^{2/L} - 1)^2} + \frac{2}{N^2} \sum_{j = 2}^{N - 1} \left( \frac{1}{(e^{2/L} - 1)^2} \right) \right. \]

and combining this expression with Eq. (A1) we obtain for the expectation of \( \langle h_1 \xi^1 \rangle^2 \)

\[ \left( \langle h_1 \xi^1 \rangle^2 \right) = \frac{K - 1}{N} + \frac{2}{N^2} \sum_{j = 2}^{N - 1} \left( \frac{1}{e^{2/L} - 1} \right) \left( \frac{1}{(e^{2/L} - 1)^2} \right) + \frac{(K - 1) (K - 2)}{N^2} \frac{1}{(e^{2/L} - 1)^2} + 1. \]

Appendix B: Detailed computations for the Dense Associative Memories

Here we show how to compute the quantities \( \langle h_1 \xi^1 \rangle \) and \( \langle (h_1 \xi^1)^2 \rangle \) for a DAM with odd bodies interaction.

Using Eq. (13) we can write the former as

\[ \langle h_1 \xi^1 \rangle = \sum_{\mu = 1}^{K} \sum_{i_2, \ldots, i_p} \left( \xi^\mu \xi^1 \prod_{k=i_2, \ldots, i_p} \xi^\mu \xi^1 \right) \]

\[ = 1 + \sum_{\mu = 1}^{K} \sum_{i_2, \ldots, i_p} \left( \xi^\mu \xi^1 \prod_{k=i_2, \ldots, i_p} \xi^\mu \xi^1 \right) \]

Since different patterns are uncorrelated, we can rewrite the expectation as

\[ \langle h_1 \xi^1 \rangle = \sum_{\mu = 1}^{K} \sum_{i_2, \ldots, i_p} \left( \xi^\mu \xi^1 \prod_{k=i_2, \ldots, i_p} \xi^\mu \xi^1 \right) = 0, \]

where the last equality derives from the fact that the expectation of an odd number of spins is null for the 1D Ising model. Concluding we have

\[ \langle h_1 \xi^1 \rangle = 1. \]
We can now turn to the expectation of the local field squared so to compute the variance. Analogously to Eq. (9) we have
\[
\langle (h_1 \xi_1^2) \rangle = \frac{2h_1 \xi_1^2}{\langle A \rangle} + \left[ \frac{1}{N^{p-1}} \sum_{\mu=2}^{K} \sum_{i_2 \cdots i_p} \xi_1^\mu \xi_1^\phi \prod_{k=i_2 \cdots i_p} \xi_k^\phi \right]^2
\]
The expectation of the first term is trivial and satisfies \(\langle A \rangle = 1\), so let us focus on the last term, its expectation is
\[
\langle B \rangle = \frac{1}{N^{2p-2}} \sum_{\mu, \phi = 2}^{K} \sum_{j_2 \cdots j_p} \sum_{i_2 \cdots i_p} \left\langle \xi_1^\mu \xi_1^\phi \prod_{k=i_2 \cdots i_p} \xi_k^\phi \prod_{l=j_2 \cdots j_p} \xi_l^\phi \right\rangle
\]
that is
\[
\langle B \rangle = \frac{1}{N^{2p-2}} \sum_{\mu, \phi = 2}^{K} \sum_{j_2 \cdots j_p} \sum_{i_2 \cdots i_p} \left\langle \prod_{k=i_2 \cdots i_p} \xi_k^\phi \prod_{l=j_2 \cdots j_p} \xi_l^\phi \right\rangle.
\]
If \(\mu \neq \phi\), the second expectation can be factorized in two expectations each containing an odd number of terms which thus give a null contribution. We can then set \(\mu = \phi\)

obtaining
\[
\langle B \rangle = \frac{1}{N^{2p-2}} \sum_{\mu, \phi = 2}^{K} \sum_{j_2 \cdots j_p} \sum_{i_2 \cdots i_p} \left\langle \prod_{k=i_2 \cdots i_p} \xi_k^\phi \prod_{l=j_2 \cdots j_p} \xi_l^\phi \right\rangle
\]
which yields
\[
\langle B \rangle = \frac{K-1}{N^{2p-2}} \sum_{i_2 \cdots i_p} \sum_{j_2 \cdots j_p} \left\langle \prod_{k=i_2 \cdots i_p} \xi_k^\phi \prod_{l=j_2 \cdots j_p} \xi_l^\phi \right\rangle^2.
\]
Let us focus on the expectation, by writing it in full we have
\[
\left\langle \prod_{k=i_2 \cdots i_p} \xi_k^\phi \prod_{l=j_2 \cdots j_p} \xi_l^\phi \right\rangle = \left\langle \xi_{i_2}^\phi \cdots \xi_{i_p}^\phi \xi_{j_2}^\phi \cdots \xi_{j_p}^\phi \right\rangle.
\]
Here we have to pay attention to all terms in which two or more of the \(\xi\)s are equal. Since no self-interaction is present in the HNN, this can occur only if \(i_a = j_b\) for some \(a, b\). As a consequence we have to count in how many ways we can form \(n\) pairs between the \(\xi_i\) and the \(\xi_j\). It is easy to see that such a number \(M(n, p)\) is given by
\[
M(n, p) = \frac{[(p-1)!]^2}{n!(p-n-1)!^2},
\]
because the first pair can be chosen in \([(p-1)!]^2\) ways, the second in \([(p-2)!]^2\) and so on, but we have also to divide by \(n!\) in order to not to give any relevance to the order by which the different pairs are chosen. Exploiting this result we can rewrite Eq. (B3) as
\[
\langle B \rangle = \frac{K-1}{N^{2p-2}} \sum_{i_2 \cdots i_p} \sum_{j_2 \cdots j_p} \sum_{n=0}^{p-1} \left[ \frac{[(p-1)!]^2}{n!(p-n-1)!^2} \right]^2 \left\langle \prod_{k=i_2 \cdots i_p} \xi_k^\phi \prod_{l=j_2 \cdots j_p} \xi_l^\phi \right\rangle^2 \delta_{i_m, j_m} = \sum_{n=0}^{p-1} \left[ \frac{K-1}{N^{2p-2-n}} \frac{[(p-1)!]^2}{n!(p-n-1)!^2} \right]^2 \sum_{i_2 \cdots i_p} \sum_{j_2 \cdots j_p} \left\langle \prod_{k=i_2 \cdots i_p} \xi_k^\phi \prod_{l=j_2 \cdots j_p} \xi_l^\phi \right\rangle^2 = \sum_{n=0}^{p-1} \left[ \frac{K-1}{N^{2p-2-n}} \frac{[(p-1)!]^2}{n!(p-n-1)!^2} \right]^2 \sum_{k=2}^{N} \sum_{l=k+1}^{N} \left\langle \xi_k^\phi \xi_l^\phi \right\rangle \right]^{p-n-1}.
\]
Note that in writing this expression we are using the convention that when there is not an explicit kronecker delta making two indices equal they are different. Moreover we exploited the fact that in the one dimensional Ising model the correlation between an even number of spins can be factorized in two-bodies correlations. We can finally rewrite the expectation of \(B\) recurring to Eq. (A3), this gives
\[
\langle B \rangle = \sum_{n=0}^{p-1} \left[ \frac{K-1}{N^{2p-2-n}} \frac{[(p-1)!]^2}{n!(p-n-1)!^2} \right]^2 \left[ \frac{2N}{e^{2L/1}} \right]^{p-n-1}.
\]
From this result, going back to Eq. (B2), we obtain

\[
\langle (h_\xi_1^2) \rangle = 1 + \sum_{n=0}^{p-1} \left\{ \frac{K - 1}{N^{2p-2-n}} \frac{[(p - 1)!]^2}{n![(p - n - 1)!]^2} \left[ \frac{2N}{e^{2L} - 1} \right]^{p-n-1} \right\}
\]