Statistics and Probability Letters 106 (2015) 256–261

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On the capacity of an associative memory model based on neural cliques

Judith Heusel[a], Matthias Löwe[a,*], Franck Vermet[b]

[a] Fachbereich Mathematik und Informatik, University of Münster, Einsteinstraße 62, 48149 Münster, Germany

[b] Laboratoire de Mathématiques, UMR CNRS 6205, Université de Bretagne Occidentale, 6, avenue Victor Le Gorgeu, F-29238 BREST Cedexm, France

ARTICLE INFO

Article history:
Received 13 March 2015
Accepted 22 July 2015
Available online 30 July 2015

MSC:
82C32
60K35
68T05
92B20

Keywords:
Neural networks
Associative memory
Exponential inequalities

ABSTRACT

Based on recent work by Gripon and Berrou (2011), we introduce a new model of an associative memory. We give upper and lower bounds on the memory capacity of the model.

1. Introduction

In Gripon and Berrou (2011) the authors introduced a new and biologically motivated model of an associative memory. This model is more effective than standard models of associative memories, in particular the Hopfield model. In this model, which we will call GB model for short, there are N neurons grouped into c clusters of l neurons. Typically, c is chosen to be $\log l$. One tries to store M messages $m^1, \dots, m^M$ in this network. These messages are sparse: Each message m^μ has c active neurons, one in each of the clusters. We write $m^\mu = (m_1^\mu, \dots, m_c^\mu)$ for $\mu = 1, \dots, M$. For each $i = 1, \dots, c$, m_i^μ denotes the active neuron of the message m^μ in the ith block. A message $m^0 = (m_1^0, \dots, m_c^0)$ is stored in the model, if all edges of the complete graph spanned by $(m_1^0, \dots, m_c^0)$ are present in the set of edges $\mathcal{E} := \{e : e \text{ is an edge of one of the } m^\mu\}$. Gripon et al. analyze the performance of this network for a random input (see e.g. Gripon and Berrou, 2011, Aliabadi et al., 2014). However, their analysis is either based on numerical simulations or on unjustified independence assumptions. Indeed, the precise form of the network makes it difficult to analyze it rigorously. In the present paper we strive for the rigorous analysis of another associative memory model, which is closely related to the GB model.

To motivate it, observe that the GB model can be described as follows: Let $\mathcal{A} = \{1, \dots, l\}$. A message m^μ is then a string $m^\mu = (m_1^\mu, \dots, m_c^\mu) \in \mathcal{A}^c$. With m^μ we associate a (column) vector $\psi(m^\mu) \in (\{0, 1\}^l)^c$ obtained by replacing the m_i^μ with the unit vector $e_{m_i^\mu}$. In a slight abuse of notation we will also use the notation $\mathcal{A}^c$ for the set $(\{0, 1\}^l)^c$. Now build the

* Corresponding author.
E-mail address: maloewe@math.uni-muenster.de (M. Löwe).

http://dx.doi.org/10.1016/j.spl.2015.07.026
0167-7152/

0–1-matrix $\tilde{W}$ given by $\tilde{W} = \max_{m \in \mathcal{M}} \psi(m)\psi(m)^t$ where $\mathcal{M} = \{m^1, \ldots, m^M\}$. Thus for $a \neq a'$ $\tilde{W}_{(a,k),(a',k')} = 1$ if and only if there is an edge between (a, k) and (a', k'). On the other hand, $\tilde{W}_{(a,k),(a,k')} = 1$ if and only if $k = k'$ and there exists μ such that the kth neuron in block a is 1. With $\tilde{W}$ one associates a dynamics D on $(\{0, 1\}^l)^c$: for $v \in (\{0, 1\}^l)^c$,

$$D(v)_{(a,k)} = \mathbf{1}_{\left\{\sum_{b=1}^{c} \mathbf{1}_{\left\{\sum_{r=1}^{l} \tilde{W}_{(a,k),(b,r)} v_{(b,r)} \geq 1\right\}} \geq c\right\}}.$$

Obviously, for all learned message $m \in \mathcal{M}$, we have $D(\psi(m)) = \psi(m)$. However, a more detailed analysis of the associative abilities of the network turns out to be difficult due to the double indicator structure of D and the max in $\tilde{W}$.

We therefore propose the following variant of the above model. Consider the random variables $\zeta^\mu_{(a,i)}$, that denote if neuron i of cluster a is part of message μ:

$$\zeta^\mu_{(a,i)} = \begin{cases} 1 & \text{if neuron } i \text{ of cluster } a \text{ is part of message } \mu, \text{ i.e. } m^\mu_a = e_i, \\ 0 & \text{otherwise.} \end{cases}$$

Under the above assumptions on the patterns the random variables $(\zeta^\mu_{(a,i)})^{1\leq\mu\leq M}_{1\leq a\leq c, 1\leq j\leq l}$ are Bernoulli variables, each with parameter $\frac{1}{l}$, and $\zeta^\mu_{(a,i)}$ is independent of $\zeta^\nu_{(b,j)}$ if $a \neq b$ or $\mu \neq \nu$. We define the matrix $W \in \mathbb{N}^{cl\times cl}$ by

$$W_{(a,i),(b,j)} = \sum_{\mu=1}^{M} \zeta^\mu_{(a,i)} \zeta^\mu_{(b,j)},$$

for $a, b \in \{1, \ldots, c\}$, $a \neq b$ and $i, j \in \{1, \ldots, l\}$. We set $W_{(a,i),(a,j)} = 0$ for all $i, j \in \{1, \ldots, l\}$ and $a \in \{1, \ldots, c\}$.

Given an input vector $v = (v_{(b,j)})_{1\leq b\leq c, 1\leq j\leq l} \in \{0, 1\}^{cl}$, we define the dynamics

$$\varphi_{(a,i)}(v) = \mathbb{1}_{\left\{\sum_{b=1}^{c}\sum_{j=1}^{l} W_{(a,i),(b,j)} v_{(b,j)} \geq \kappa c\right\}} \tag{1}$$

for some $\kappa > 0$. Obviously this model is closely related to the GB model. However, the structure of W and the dynamics that include sums of random variables rather maxima and minima, make it more accessible to probabilistic tools. Indeed from this point of view the model is somewhat similar to the Hopfield model, whose storage capacity has been analyzed in McEliece et al. (1987), Komlós and Paturi (1988), Newman (1988), Loukianova (1997), Talagrand (1998), Löwe (1998), Löwe and Vermet (2005b), Löwe and Vermet (2011), and many other papers.

However, it can be shown that the present model has a non-vanishing efficiency in the given range of parameters. In the present note in Section 2 we bound the storage capacity of our version of the GB model from below and above. Indeed, a decent upper bound can only be proven for very few models of associative memories. Sections 3 and 4 are devoted to the proofs of these results.

2. Bounds on the storage capacity

The purpose of the present section is to analyze whether one can find threshold κ such that an amount of $M = \alpha l^2$ patterns can be stored in the above described model. This is indeed the case: In this sense our model is considerably more effective than the Hopfield model, where only for $M \leq N/(2\log N)$ all patterns are fixed points of the retrieval dynamics (see McEliece et al., 1987, Burshtein, 1994, Bovier, 1999). In our variant of the GB model there are $N = l \log l$ neurons, while the number of messages to be stored is proportional to l^2. In terms of capacity and sparseness it is therefore related to the Willshaw model (Golomb et al., 1990) and the Blume–Emery–Griffiths model (Löwe and Vermet, 2005a). More precisely, we will prove the following theorem.

Theorem 2.1. *In the above model with coding matrix W, let $c = \log l$ and $M = \alpha l^2$. For $\kappa \leq 1 - 1/c$ we have*

1. *If $\alpha < \kappa$, for every fixed μ, every fixed block a and every fixed coordinate k, we have that*

$$\mathbb{P}\left(\varphi_{(a,k)}(m^\mu) \neq m^\mu_{(a,k)}\right) \to 0 \quad \text{as } l \text{ and therefore } N \text{ tends to infinity}.$$

2. *If $\alpha < \kappa \exp(-(3 + \kappa)/\kappa)$, for $N \to \infty$ we have that*

$$\mathbb{P}\left(\exists 1 \leq a \leq c, 1 \leq i \leq l, 1 \leq \mu \leq M : \varphi_{(a,i)}(m^\mu) \neq m^\mu_{(a,i)}\right) \to 0.$$

Moreover, we will also be interested in the error correcting abilities of the model. Such errors in a message occur, if some characters are false, or erased. In fact, both types of errors are equivalent, if we replace each missing character with a randomly chosen letter. If we not only require that the network is able to correct a certain percentage of errors, this may lower its capacity. However, the order of the capacity is maintained as can be read off from the following theorem. We will concentrate on one step of the parallel dynamics. To this end we define the discrete ball of radius r centered in m^μ as

$$\mathcal{B}(m^\mu, r) = \{m \in \mathcal{A}^c : d_H(m^\mu, m) = \operatorname{card}\{j : m^\mu_j \neq m_j\} \leq r\}.$$