# Information retrieval in neural networks. - I. Eigenproblems in neural networks

A.D. Maruani, R.C. Chevallier, G.Y. Sirat

*Revue Phys. Appl.* **22** (1987) 1321-1325 OCTOBRE **1987**, PAGE 1321

Classification
*Physics Abstracts*
42.30 — 42.80 — 84.20P

# Information retrieval in neural networks. I. Eigenproblems in neural networks

A. D. Maruani, R. C. Chevallier and G. Y. Sirat

Ecole Nationale Supérieure des Télécommunications, Groupe Optique des Matériaux, 46, rue Barrault, 75634 Paris Cedex 13, France

*(Reçu le 13 novembre 1986, révisé le 8 juillet 1987, accepté le 9 juillet 1987)*

**Résumé.** — Nous étudions le problème des vecteurs propres et des valeurs propres de la matrice synaptique correspondant au modèle d'Hopfield. Nous montrons en particulier que les vecteurs propres se laissent ranger en deux sous-espaces orthogonaux : le premier espace est celui des mémoires prototypes ; la dispersion des valeurs propres correspondantes autour de leur valeur moyenne donne une mesure de l'orthogonalité de ces prototypes. Le second espace est quant à lui totalement dégénéré. En outre, nous montrons une similitude entre l'algorithme de mémoire associative et l'algorithme d'orthonormalisation de Gram-Schmidt. A partir de ces résultats, nous développons un modèle de matrice synaptique dont nous évaluons numériquement les performances en les comparant à celles d'un processeur d'Hopfield fonctionnant dans les mêmes conditions.

**Abstract.** — Consideration of the eigenproblem of the synaptic matrix in Hopfield's model of neural networks suggests to introduce a matrix built up from an orthogonal set, orthogonal to the original memories. With this new scheme, capacity storage is significantly enhanced and robustness at least conserved.

## Introduction.

A recent proposal by Hopfield of a neural network model has triggered a renewal of activity in this field [1]. Two particularities of Hopfield's model are, on the one hand, the emphasis on nonlinear feedback, and on the other hand, its remarkable ability for optical implementation. Farhat and Psaltis [2] have proposed and demonstrated such an ability. A problem in this type of associative memory is its capacity storage ; namely, what is the maximum content of the set of the memories for a given signal to noise ratio or, more generally, what are the attractors of the process, and their behaviour ? Those problems have been adressed by McEliece *et al.* [3], and Amit *et al.* [4].

More generally, the problem of optimal storage matrices has been adressed by T. Kohonen [5], Personnaz *et al.* [6], Venkatesh and Psaltis [7] and Kanter *et al.* [8]. In particular for linearly independant patterns the algorithm given in this paper leads to the same matrix as in these references.

A thorough comparative discussion of those approaches is in current progress (Venkatesh, Psaltis and Sirat), and a preliminary presentation has been done [9].

This paper is the first one of a series of papers, and is devoted to a brief discussion of the normal properties of synaptic matrices involved in the algorithms. The discussion will enable us to propose a new scheme, allowing both large capacity storage and robustness of the algorithm.

In a second paper we will propose an optical implementation of two-dimensional neural networks based on Frequency Multiplexed Raster ; in a third paper we will propose an optical implementation of shift invariant neural networks based also on Frequency Multiplexed Raster.

This paper is organized in the following way : in the first section we present briefly the Hopfield Model and discuss the cross correlation factor ; in the second section, we consider some eigenproblems and in section 3 we introduce and discuss schematically a new scheme.

## 1. Basics of Hopfield model and correlation factor.

In this section we recall the Hopfield model formalism, and discuss the behaviour of the cross-correlation factor.

Let a set of $M$ binary vectors $V^{(m)}$ be given as the memory set. Each component is $V_i^{(m)} = \pm 1$ for $1 \leqslant m \leqslant M$ and $1 \leqslant i \leqslant N$, so that the euclidian norm $\| V_m \|$ is $N$, the number of bits of each vector. We assume that the vectors are linearly independent.

The associated matrix, $T^{(m)}$ of vector $V^{(m)}$, is defined by :

$$T_{ij}^{(m)} = V_i^{(m)} V_j^{(m)} (1 - \delta_{ij}) . \qquad (1.1)$$

Where $\delta_{ij}$ is the Kronecker symbol.

The synaptic matrix is $T$ :

$$T = \sum_m T^{(m)} \quad \text{so that}$$
$$T_{ij} = \sum_{m=1}^{M} V_i^{(m)} V_j^{(m)} (1 - \delta_{ij}) . \qquad (1.2)$$

If the algorithm is addressed by a probe $X$ equal to one of the stored memories, say $V^{(\mu)}$, it will yield first the estimate for coordinate $i$ :

$$\begin{aligned} V_i &= \sum_{j=1}^{N} T_{ij} X_j = \sum_{j=1}^{N} T_{ij} V_j^{(\mu)} \\ &= \sum_{j=1}^{N} T_{ij}^{(\mu)} V_j^{(\mu)} + \\ &\quad + \sum_{m \neq \mu} \sum_{j=1}^{N} T_{ij}^{(m)} V_j^{(\mu)} \end{aligned} \qquad (1.3)$$

from (1.2) :

$$V_i = (N - 1) V_i^{(\mu)} + \sum_{m \neq \mu} \sum_{j \neq i} V_i^{(m)} V_j^{(m)} V_j^{(\mu)} . \qquad (1.4)$$

The first term will yield $V^{(\mu)}$ amplified by $(N - 1)$, while the second term is some kind of correlation of the probe with the remainder of the memories and represents an unwanted, cross talk, term. Its mean value is zero and its standard deviation is $[(N - 1)(M - 1)]^{1/2}$.

For $N$ sufficiently larger than $M$, the sign of the right hand side of (1.4) is safely that of $V_i$, i.e. positive if $V_i^{(\mu)} = 1$ and negative otherwise. Thresholding of $V_i$ will therefore yield $V_i^{(\mu)}$:

$$V_i^{(\mu)} = \text{sign } (V_i) . \qquad (1.5)$$

If the algorithm is addressed with a binary-valued vector that is not one of the stored memories but close to one of them in the Hamming sense, vector-matrix multiplication and thresholding operations yield an output binary vector which, in general, is an approximation of the stored memory that is at the shortest Hamming distance from the input vector. If in that case this output vector is fed back and used as the input to the system, the new output generally is a more accurate version of the stored memory and successive iterations converge to the correct vector.

A useful concept in this respect is the cross correlation factor $C$ :

$$C = \sum_i \sum_j T_{ij} X_i X_j \qquad (1.6)$$

which is just the negative of the energy defined in [1].

As was demonstrated [1], thresholding and feeding back make energy decrease and correlation increase. We shall use throughout the cross-correlation factor as an indicator of the presence of a nominal state vector equal, or nearly equal, to the probe ; utilizing (1.1) and (1.2), equation (1.2) can be rewritten as :

$$C = \sum_{m=1}^{M} C^{(m)} \qquad (1.7)$$

where $C^{(m)}$ the $m$-th partial cross-correlation factor is given by

$$\begin{aligned} C^{(m)} &= \sum_{i=1}^{N} \sum_{j=1}^{N} T_{ij}^{(m)} X_i X_j \\ &= \sum_{i=1}^{N} (V_i X_i) \sum_{j=1}^{N} (V_j X_j) - \\ &\quad - \sum_{i=1}^{N} \sum_{j=1}^{N} \delta_{ij} X_i X_j \\ &= (p_R^{(m)} - p_W^{(m)})^2 - N \\ &= (2 p_R^{(m)} - N)^2 - N \end{aligned} \qquad (1.8)$$

where $p_R^{(m)}$ is the number of right bits (bits for which $V_i^{(m)} . X_i = + 1$) and $p_W^{(m)}$ is the number of wrong bits (bits for which $V_i^{(m)} . X_i = - 1$) ; the sum of $p_R^{(m)}$ and $p_W^{(m)}$ is $N$.

If the algorithm is addressed by a probe $X$ equal to one of the stored memories, say $V^{(\mu)}$, the $\mu$-th partial cross-correlation factor is :

$$C^{(\mu)} = (2 p_R^{(\mu)} - N)^2 - N = N^2 - N = N (N - 1) . \qquad (1.9)$$

If the algorithm is addressed by a probe $X$, close to one of the stored memories, say $V^{(\mu)}$, the $\mu$-th partial cross-correlation factor is given by equation (1.8).

In the following we consider the thermodynamic limit which is operational when $N$ is large enough.

Let the probe and memory $V^{(m)}$ be a sequence of $N$ random binary bipolar digits ; then the probability of $p_R$ to be equal to a given integer $n$ is a binomial distribution whose thermodynamic limit is Gaussian :

$$p'(p_R^{(m)} = n) = \binom{N}{n} \approx \left( \frac{2}{\pi N} \right)^{1/2} \times \times \exp - \left[ \left( \frac{2}{N} \right) (n - N/2)^2 \right] \qquad (1.10)$$

so that the probability of $X = (2\,p_{\mathrm{R}}^{(m)} - N)$ is

$$p(X = X_0) = \frac{1}{2}\left(\frac{2}{\pi N}\right)^{1/2} \exp(-x_0^2/2\,N) \quad (1.11)$$

and the mean value of $C^{(m)}$ is :

$$C^{(m)} = \int_{-\infty}^{\infty} (X^2 - N)\,p(X)\,\mathrm{d}X = \frac{1}{2}\left(\frac{2}{\pi N}\right)^{1/2} \times \times \int_{-\infty}^{\infty} (x^2 - N)\exp(-x^2/2\,N)\,\mathrm{d}X = \left(\frac{2}{\pi N}\right)^{1/2}\int_0^{\infty}(x^2 - N) \times \times \exp(-x^2/2\,N)\,\mathrm{d}x = 0\,. \quad (1.12)$$

In the same way, we obtain that

$$\mathrm{var}\,(C^{(m)}) = \int_{-\infty}^{\infty} (x^2 - N)^2\,p(x)\,\mathrm{d}x = 2\,N^2\,. \quad (1.13)$$

The standard deviation of $C^{(m)}$ is $N\sqrt{2}$, so that the standard deviation of the cross-correlation factor for a probe uncorrelated to any one of the memories is $N(2\,M)^{1/2}$ and the noise part of a probe correlated to one of the memories, but uncorrelated to all the others is $N[2(M-1)]^{1/2}$.

In conclusion, the behaviour of the cross-correlation factor, for uncorrelated memories, is the following :

i) if the probe is equal to one of the stored memories the cross-correlation factor is equal to a signal term of value $N(N-1)$ and a noise term of standard deviation $N[2(M-1)]^{1/2}$ ; note that from (1.13) the cross-correlation factor is larger than $N(N-M)$ ;

ii) if the probe is close to one of the stored words the signal term is decreased to :

$$(2\,p_{\mathrm{R}}^{(\mu)} - N)^2 - N \quad (1.8)$$

and the noise term is (almost) unchanged.

iii) if the probe is uncorrelated with the memories, only a noise term is present with a standard deviation of $N(2\,M)^{1/2}$.

## 2. Eigenproblems for matrix $T$.

In this section we discuss some eigenproblems of the synaptic matrix.

The description in section 1 supports an elementary geometrical picture : from (1.1), one has, in the general case

$$T^{(m)}X = (V^{(m)}\,.\,X)\,V^{(m)} - X\,, \quad (2.1)$$

where the dot represents the ordinary scalar product.

All vectors are assumed to have the same length, $N$. Replacing $X$ in (2.1) by $V^{(m)}$ provides clearly the following result for the eigenproblem of partial matrix $T^{(m)}$ :

$V^{(m)}$ is eigenvector for $T^{(m)}$, with eigenvalue $(N-1)$.

The $(N-1)$ remaining eigenvectors span the subspace orthogonal to $V^{(m)}$ and the degenerate eigenvalue is $(-1)$.

From (1.2) and (2.1) one derives

$$TX = \sum_{m=1}^{M} [(V^{(m)}\,.\,X)\,V^{(m)} - X]\,. \quad (2.2)$$

A consequence of (2.2) is that, if the set of memories is a complete base, then $TX$ is null for any $X$, which is possible if and only if matrix $T$ is null itself. This extreme case suggests storage limitations of the synaptic matrix. Now, equations (2.1) and (2.2) show that, for memory $\mu$

$$TV^{(\mu)} = (N-1)\,V^{(\mu)} + \sum_{m \neq \mu} T^{(m)}V^{(\mu)}\,. \quad (2.3)$$

Then, noise in (1.4) is seen to be nothing but the action on $V^{(\mu)}$ of all the partial matrices $T^{(m)}$ except its own associate $T^{(\mu)}$.

The eigenproblem for the synaptic matrix reads as :

$$\sum_{m=1}^{M} (V^{(m)}\,.\,X)\,V^{(m)} = (c + M)\,X\,. \quad (2.4)$$

Hereabove, $X$ and $c$ are the sought after eigenvectors and eigenvalues. If $X$ belongs to the subspace orthogonal to that generated by the memory set, the left hand of (2.4) vanishes so that the solution is $c + M = 0$.

As a first result :

The synaptic matrix $T$ has eigenvalue $(-M)$ with degeneracy $(N-M)$ ; corresponding eigenvectors are orthogonal to the memory set.

If $(c+M)$ does not vanish, equation (2.4) can be rewritten as :

$$X = \sum_{m=1}^{M} \alpha_m\,V^{(m)} \quad (2.5)$$

where

$$\alpha_m = \frac{(X\,.\,V^{(m)})}{c + M} \quad (2.6)$$

so that $X$ belongs to the space generated by the memory set.

Assume for a while that the memory set is orthogonal ; then multiplying (2.4) by $V^{(\mu)}$ provides

$$\sum_{m=1}^{M} (V^{(m)}\,.\,X)(V^{(m)}\,.\,V^{(\mu)}) = (c + M)\,X\,.\,V^{(\mu)} \quad (2.7)$$

or using $V^{(m)} V^{(\mu)} = N \delta_{m\mu}$

$$N(V^{(\mu)} . X) = (c + M)(X . V^{(\mu)}) . \quad (2.8)$$

Equation (2.8) shows the seeked eigenvalue is $\langle c \rangle = N - M$, it is degenerate $M$ times ; clearly $V^{(\mu)}$ are the $M$ associate eigenvectors.

If, as in the general case, the memories are not orthogonal.

We define :

$$V^{(m)} . V^{(p)} = N \delta_{mp} + \varepsilon_{mp}(1 - \delta_{mp}) \quad (2.9)$$

$$c_i = N - M + e_i = \langle c \rangle + e_i \qquad i = 1, ..., M . \quad (2.10)$$

Equation (2.9) defines matrix $\varepsilon$. Its dimension is $M < N$. The eigenproblem transforms into :

$$e \alpha_m = \sum_{p \neq m} \varepsilon_{mp} \alpha_p \quad (e = (e_1, e_2, ..., e_\mu)) . \quad (2.11)$$

Dispersion of the eigenvalues is estimated by the variance

$$\sum_{i=1}^{M} e_i^2 .$$

This term appears to be the one with power $(M - 2)$ in the characteristic polynom of matrix $\varepsilon$, and then can be derived by direct inspection, with no need to actually compute the set of $e_i$'s.

Additional properties can be deduced from (2.4) by multiplying both side by $X$ ; the right hand side is a sum of squares and is positive, while the left hand size is $(c + M) N$. Note also that the right size is bounded by $MN^2$ — being $M$ sums of terms smaller or equal to $N^2$ ; so that

$$- M \leqslant c \leqslant M(N - 1) . \quad (2.12)$$

## 3. A modified synaptic matrix.

As a direct consequence of the above, we present a new scheme for calculating the synaptic matrix ; let the $V^{(m)}$ be the memory set and consider $L = N - M$ random binary vectors, assumed to be independant. On this new set, the Gram-Schmidt orthogonalization procedure is applied. Note first that from this point we shall work partially in the analog component scheme for the whole set of vectors, but that retrieval concerns only the original, binary vectors of the given memories, and not the orthogonalized vectors.

Another remark concerns the striking formal analogy between the Gram-Schmidt procedure and the Hopfield processor : actually, the transformation of an independant set $X_n$ into an orthonormal set $V_m$ makes use of the recursion formula

$$\left| \begin{array}{ll} \text{(a)} & V_1 = X_1 / \| X_1 \| \\ \text{(b)} & V_{\mu+1} = \sum_{m=1}^{\mu} [(V^{(m)} . X_{\mu+1}) V^{(m)}] - X_{\mu+1} \\ \text{(c)} & V_{\mu+1} = V_{\mu+1} / \| V_{\mu+1} \| \end{array} \right. \quad (3.1)$$

which has to be compared to equation (2.2).

The last $L$ vectors of the set are called the antimemories $W_i^{(1)}$ ; by construction the $V^{(m)}$ and the $W^{(1)}$ satisfy :

$$\left| \begin{array}{ll} \text{(a)} & \sum_{j=1}^{N} W_j^{(1)} V_j^{(m)} = 0 \quad \text{for any } l \text{ and } m \\ \text{(b)} & \sum_{j=1}^{N} W_j^{(1)} W_j^{(k)} = \delta_{1k} . \end{array} \right. \quad (3.2)$$

The new synaptic matrix, $S$, will be defined as

$$S_{ij} = \sum_{l=1}^{L} S_{ij}^{(1)} = \sum_{l=1}^{L} W_i^{(1)} W_j^{(1)} (\delta_{ij} - 1) . \quad (3.3)$$

Each one of the partial synaptic matrix $S_{ij}^{(1)}$ fulfills the relation :

$$\begin{aligned} \hat{U}_i^{(1)} &= \sum_{j=1}^{N} S_{ij}^{(1)} V_j^{(\mu)} \\ &= \sum_{j=1}^{N} W_i^{(1)} W_j^{(1)} V_j^{(\mu)} (\delta_{ij} - 1) \\ &= W_i^{(1)} \sum_{j=1}^{N} W_j^{(1)} V_j^{(\mu)} + [W_i^{(1)}]^2 V_i^{(\mu)} \\ &= a_i^{(1)} V_i^{(\mu)} \end{aligned} \quad (3.4)$$

where

$$a_i^{(1)} = [W_i^{(1)}]^2 \geqslant 0 .$$

Note that each one of the partial synaptic matrix fulfills a relation similar to (3.4) ; also

$$U_i = \sum_{l=1}^{L} U_i^{(1)} = \sum_{l=1}^{L} S_{ij} V_j^{(\mu)} = a_i^{(1)} V_i^{(\mu)} \quad (3.5)$$

where

$$a_i = \sum_{l=1}^{L} a_i^{(1)} = \sum_{l=1}^{L} [W_i^{(1)}]^2 . \quad (3.6)$$

It is important to remark that $a_i$ does not depend on the particular set of vectors $W^{(1)}$ and is ultimately determined by the memory set $V^{(m)}$.

With this scheme, the memories can be retrieved for almost any value of $M$ (of course smaller than $N$). For instance, computer simulations showed that retrieval was 100 % perfect up to $M = 254$ for $N = 256$ (see Fig. 1). Simultaneously a radius of attraction, characterized as a Hamming distance beyond which retrieval is lost, was shown to be comparable to or larger than that of the original synaptic matrix. The latter property was not systematically studied but numerically observed on a number of random sampling for the memories. Last, and as shown

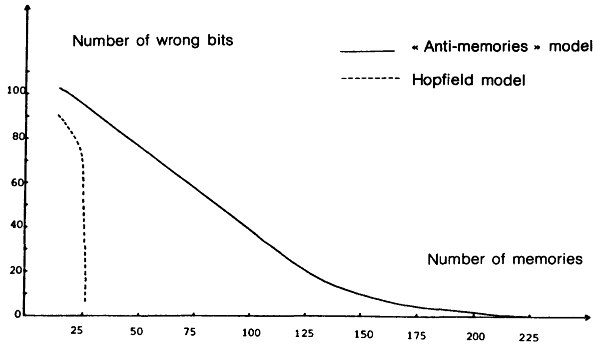


Fig. 1. — Computer simulations of the performances of the « anti-memories » model *versus* the Hopfield model. The graph shows the maximum number of wrong bits permitted as a function of the number of memories (90 % retrieval, $N = 256$).

before, the cross-correlation for any memory is exactly $(N - M)$ i.e. the lower bound of the cross-correlation factor but without any noise term.

**Conclusion.**

We have presented a brief discussion of the formal properties of the synaptic involved in neural networks and proposed a new scheme for building such matrices. An immediate property of the algorithm is its rather large capacity storage, together with its robustness. Extended discussion of this new model will be presented in a forthcoming article of this series.

**Acknowledgments.**

The authors thank Prof. Demetri Psaltis for numerous stimulating discussions. One of the authors (GYS) thank Demetri Psaltis, and his group, particularly Eung Gi Paek, John Hong, Santosh Venkatesh and Kelvin Wagner, for introducing him to this subject.

Programming and numerical calculations were performed by Didier Charlot and N. Boisgontier and B. Plessis. We thank Françoise Petiot for carefully typing the converging iterations of this paper. Last, we thank the referee for his most welcome comments which added lisibility and rigor to this paper.

**References**

[1] HOPFIELD, J. J., Neural networks and Physical systems with emergent Collective Computational Abilities. *Proc. Nat. Acad. Sci.* **79** (1982) 2254-2258.

[2] PSALTIS, D. and FAHRAT, N., *Optic Lett.* **10** (1985) 98-100.

[3] MCELIECE, R. J., POSNER, E. C., RODEMICH, E. R. and VENKATESH, S. S., *The capacity of the Hopfield associative memory*, submitted to the IEEE Transactions on Information Theory.

[4] AMIT, D., GUTFREUND, H. and SOMPOLINSKY, H., *Phys. Rev.* **A 32** (1985) 1007 and *Phys. Rev. Lett.* **55** (1985) 1530.

[5] KOHONEN, T., *Self-Organisation and associative memory* (Springer-Verlag) 1984.

[6] PERSONNAZ, L., GUYON, I. and DREYFUS, G., Information storage and retrieval in spin-like neural networks. *J. Physique Lett. (Paris)* **46** (1985) L359-L365 and

PERSONNAZ, L., GUYON, I., DREYFUS, G. and TOULOUSE, G., A biologically constrained learning mechanism in networks of formal neurons. *J. Stat. Phys.* **43** (1986) 411.

[7] VENKATESH, S. and PSALTIS, D., Information storage and retrieval in the associative nets, submitted to the *IEEE Trans. Info. Theory.*

[8] KANTER, I., SOMPOLINSKY, H. (Bar-Ilan preprint 1986).

[9] VENKATESH, S. and PSALTIS, Snowbird Meeting April 1987.