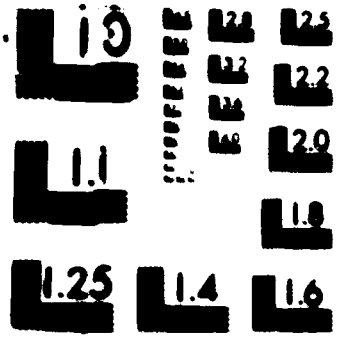
AD-A181 860 A RELAXATION MODEL FOR MEMORY WITH HIGH STORAGE DENSITY 1/1
(U) BROWN UNIV PROVIDENCE RI CENTER FOR NEURAL SCIENCE
C M BACHMANN ET AL 26 JUN 87 TR-41 N00014-86-K-0041
UNCLASSIFIED F/G 6/4 NL


1.0
1.1
1.25
1.4
1.6
2.0
2.2
2.5

(12)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER #41 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) A RELAXATION MODEL FOR MEMORY WITH HIGH STORAGE DENSITY | | 5. TYPE OF REPORT & PERIOD COVERED TECHNICAL |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) Charles M. Bachmann, Leon N Cooper, Amir Dembo, and Ofer Zeitouni | | 8. CONTRACT OR GRANT NUMBER(s) N000-14-86-K-0041 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Brown University Center for Neural Science Providence, Rhode Island 02912 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS N-201-484 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Program Office of Naval Research, Code 442 PT | | 12. REPORT DATE June 26, 1987 |
| | | 13. NUMBER OF PAGES 10 pages |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report) UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

AD-A181 860

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.



17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This is to be published in the Proceedings of the National Academy of Sciences.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

High Density Storage Model
Hopfield Model
Relaxation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We present a relaxation model based on an N-dimensional Coulomb potential. The model has arbitrarily large storage capacity and, in addition, well-defined basins of attraction about stored memory states. The model is compared with the Hopfield relaxation model.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Neural Networks

# A Relaxation Model for Memory with High Storage Density

Charles M. Bachmann, Leon N Cooper, Amir Dembo,
and Ofer Zeitouni

Department of Physics, Applied Mathematics
and Center for Neural Science
Brown University
Providence, Rhode Island 02912

Key Terms : High Density Storage Model, Hopfield Model, relaxation

## Abstract

We present a relaxation model based on an N-dimensional Coulomb potential. The model has arbitrarily large storage capacity and, in addition, well-defined basins of attraction about stored memory states. The model is compared with the Hopfield relaxation model.

# 1 Introduction

Equilibrium associative and distributed memories that are content addressable and have the ability to recall stored memories more or less imperfectly have been known and studied for years [1], [2], [3], [4], [5]. At the same time, relaxation models have been the subject of much exploration [6]. In 1982, Hopfield [7] introduced a relaxation model of memory storage and retrieval, that incorporates simultaneously a distributed memory correlation matrix and a relaxation process from a given input to an equilibrium state. Although learning procedures can be included, the model has not emphasized these. Among its problems are poor recall of stored memories when the number of items stored exceeds some percentage of the number of neurons involved.

The correlation matrix originally employed by Hopfield has relatively weak recall properties when employed as an equilibrium distributed memory. It gives perfect recall only when the inputs are orthogonal. When the inputs are not orthogonal, one can still achieve perfect recall by some orthogonal modification procedure such as Widrow-Hoff [8], or what Kohonen calls an *optimal associative mapping*. [9] Such procedures work if the number of stored memories is equal to or smaller than the dimension of the system (the number of synapses on each neuron). A procedure for storing as many memories as desired for a given dimension has also been discussed[10]. In this procedure items can be stored at arbitrary points on a hypersphere with variable regions of influence.

In this paper we present a general method for the construction of a relaxation memory in which an arbitrary number of items can be stored. The essence of the problem is to define a function whose minima lie at designated points, corresponding to the items to be stored, and to show that these are the only minima of the function. Then an appropriate relaxation procedure is defined, so that any entering pattern relaxes to one of the stored items.

# 2 Hopfield's Model and Some Improvements

In the Hopfield model[7], neurons are binary-valued threshold units and are completely interconnected, with the strength of the connections given by a correlation matrix formed from the memory states to be stored in the system:

$$w_{ij} = \sum_{s=1}^{m} \mu_i^s \mu_j^s , \tag{1}$$

where $\mu_i = \pm 1$. Input states are relaxed to local minima of a Liapunov function,

$$\xi = -\frac{1}{2} \sum_{i,j} w_{ij} \mu_i \mu_j \tag{2}$$

by random, asynchronous updating of the neurons in the layer according to:

$$\mu_i \rightarrow 2\theta(\sum_{j=1}^{N} w_{ij} \mu_j) - 1. \tag{3}$$

In its original form, the Hopfield model functions poorly as a categorizer when $(m/N) \gtrsim 0.1$, where m = the number of stored states and N = the number of neurons. Given the limitations of the original model, improvements have been sought. "Unlearning", an approach first tried by Hopfield [11], employs the relaxation of random states to a stable state (often spurious attractors); a correlation matrix is formed from the relaxed state, and then an amount proportional to this is subtracted from the original matrix:

$$w_{ij} \rightarrow w_{ij} - \alpha \mu_i^{relaxed} \mu_j^{relaxed} \tag{4}$$

With "unlearning" the number of stored states that can be correctly recalled approaches the dimensionality, N, and error correction is improved but falls to zero as $m \rightarrow N$.[12]

Recently, an interesting variation of Hopfield's "unlearning" has been studied by Potter.[12] The algorithm is a hybrid combining elements of Hopfield's "unlearning" with a modification reminiscent of the Widrow-Hoff algorithm[8]:

$$w_{ij} \rightarrow w_{ij} - \alpha(\mu_i^{target} - \mu_i^{relaxed}) \mu_j^{input} (\mu_j^{input} + 1) . \tag{5}$$

The symmetry of the synaptic matrix is preserved by making the same modification to $w_{ji}$ each time a modification is performed on the element $w_{ij}$. In simulations for which all of the input states at a radius of one Hamming unit from each stored state were used for the modification procedure, a radius of attraction of one Hamming unit was observed for m just below the dimensionality, N. Above the dimensionality, the radius of attraction and the percentage of stable stored states decays. In [13], it has been shown that Potter's algorithm may be viewed as an *"effective orthogonalization"* of the input with respect to the nonlinear relaxation process; a more complete discussion of Potter's algorithm is given there.

## 3 High Density Storage Model

In what follows we present a general method for the construction of a high storage density neural memory. We define a function with an arbitrary number of minima that lie at preassigned points and define an appropriate relaxation procedure.

Let $\vec{x}_1, ..., \vec{x}_m$ be a set of m arbitrary distinct memories in $R^N$. The "energy" function we will use is:

$$\xi = -\frac{1}{L}\sum_{i=1}^{m} Q_i \mid \vec{\mu} - \vec{x}_i \mid^{-L} \tag{6}$$

where we assume throughout that $N \geq 3$, $L \geq (N-2)$, and $Q_i > 0$ and use $\mid \cdots \mid$ to denote the Euclidean distance. Note that for L=1,N=3, $\xi$ is the electrostatic potential induced by negative fixed particles with charges $-Q_i$. This "energy" function possesses global minima at $\vec{x}_1, ..., \vec{x}_m$ (where $\xi(\vec{x}_i) = -\infty$) and has no local minima except at these points. A rigorous proof is presented in Dembo and Zeitouni[14] together with the complete characterization of functions having this property.

As a relaxation procedure, we can choose any dynamical system for which $\xi$ is strictly decreasing. In this instance, the theory of dynamical systems guarantees that for almost any initial data, the trajectory of the system converges to one of the desired points $\vec{x}^1, ..., \vec{x}^m$. However, to give concrete results and to further exploit the resemblance to electrostatics, consider the relaxation:

$$\dot{\vec{\mu}} = \vec{E}_{\vec{\mu}} \triangleq -\sum_{i=1}^{m} Q_i \mid \vec{\mu} - \vec{x}_i \mid^{-(L+2)} (\vec{\mu} - \vec{x}_i) \tag{7}$$

where for N=3,L=1, equation ( 7) describes the motion of a positive test particle in the electrostatic field $\vec{E}_{\vec{\mu}}$ generated by the negative fixed charges $-Q_1, ..., -Q_m$ at $\vec{x}_1, ..., \vec{x}_m$.

Since the field $\vec{E}_{\vec{\mu}}$ is just minus the gradient of $\xi$, it is clear that along trajectories of ( 7), $\frac{d\xi}{dt} \leq 0$, with equality only at the fixed points of ( 7), which are exactly the stationary points of $\xi$.

Therefore, using ( 7) as the relaxation procedure, we can conclude that entering at any $\vec{\mu}(0)$, the system converges to a stationary point of $\xi$. The space of inputs is partitioned into m domains of attraction, each one corresponding to a different memory, and the boundaries (a set of measure zero), on which $\vec{\mu}(0)$ will converge to a saddle point of $\xi$.

We can now explain why $\xi_{\vec{\mu}}$ has no spurious local minima, at least for L=1,N=3, using elementary physical arguments. Suppose $\xi$ has a spurious

local minima at $\vec{y} \neq \vec{x}_1, ..., \vec{x}_m$, then in a small neighborhood of $\vec{y}$ which does not include any of the $\vec{x}_i$, the field $\vec{E}_{\vec{\mu}}$ points towards $\vec{y}$. Thus, on any closed surface in that neighborhood, the integral of the normal inward component of $\vec{E}_{\vec{\mu}}$ is positive. However, this integral is just the total charge included inside the surface, *which is zero*. Thus we arrive at a contradiction, so $\vec{y}$ can not be a local minimum.

We now have a relaxation procedure, such that almost any $\vec{\mu}(0)$ is attracted by one of the $\vec{x}_i$, but we have not yet specified the shapes of the basins of attraction. By varying the charges $Q_i$, we can enlarge one basin of attraction at the expense of the others (and vice versa).

Even when all of the $Q_i$ are equal, the position of the $\vec{x}_i$ might cause $\vec{\mu}(0)$ not to converge to the closest memory, as emphasized in the example in fig 1. However, let $r = min_{1 \leq i \neq j \leq m} \mid \vec{x}_i - \vec{x}_j \mid$ be the minimal distance

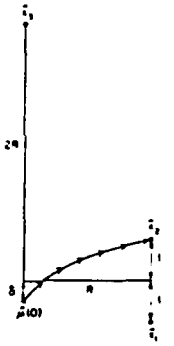


Figure 1: $\vec{\mu}(0)$ closer to $\vec{x}_1$ but converges to $\vec{x}_2$, due to the existence of $\vec{x}_3$ (assuming $R \gg 1$ and $\delta \ll 1$).

between any two memories; then, if $\mid \vec{\mu}(0) - \vec{x}_i \mid \leq \frac{r}{(1+3^{\frac{1}{k}})}$, it can be shown that $\vec{\mu}(0)$ will converge to $\vec{x}_i$, provided that $(k \triangleq \frac{L+1}{N+1} \geq 1)$. Thus, if the memories are densely packed in a hypersphere, by choosing k large enough (i.e. enlarging the parameter L), convergence to the closest memory for any "interesting" input, that is an input $\mu(\vec{0})$ with a distinctive closest memory, is guaranteed.

The detailed proof of the above property is given in [14]. It is based on bounding the number of $\vec{x}_j$, $j \neq i$, in a hypersphere of radius R ($R \geq r$)

around $\vec{x}_i$, by $[2\frac{R}{r}+1]^N$, then bounding the magnitude of the field induced by any $\vec{x}_j$, $j \neq i$, on the boundary of such a hypersphere by $(R- \mid \vec{\mu}(0)-\vec{x}_i \mid)^{-(L+1)}$, and finally integrating to show that for $\mid \vec{\mu}(0)-\vec{x}_i \mid \leq \frac{\theta r}{1+3^{\frac{1}{L}}}$, with $\theta < 1$, the convergence of $\vec{\mu}(0)$ to $\vec{x}_i$ is within finite time T, which behaves like $\theta^{L+2}$ for $L \gg 1$ and $\theta < 1$ and fixed. Intuitively the reason for this behaviour is the short-range nature of the fields used in equation ( 7). Because of this, we also expect extremely low convergence rate for inputs $\vec{\mu}(0)$ far away from *all* of the $\vec{x}_i$.

The radial nature of these fields suggests a way to overcome this difficulty, that is to increase the convergence rate from points very far away, without disturbing all of the aforementioned desirable properties of the model. Assume that we know in advance that all of the $\vec{x}_i$ lie inside some large hypersphere S around the origin. Then, at any point $\vec{\mu}$ outside S, the field $\vec{E}_{\vec{\mu}}$ has a positive projection radially into S. By adding a long-range force to $\vec{E}_{\vec{\mu}}$, effective only outside of S, we can hasten the movement towards S, from points far away, without creating additional minima inside of S. As an example the force ($-\vec{\mu}$ for $\vec{\mu} \notin S$; 0 for $\vec{\mu} \in S$) will pull any test input $\vec{\mu}(0)$ to the boundary of S within the small finite time $T \approx \frac{1}{|S|}$, and from then on the system will behave inside S according to the original field $\vec{E}_{\vec{\mu}}$.

Up to this point, our derivations have been for a continuous system, but from it, we can deduce a discrete system. We shall do this mainly for a clearer comparison between our high density memory model and the discrete version of Hopfield's model. Before continuing in that direction, note that our continuous system has ***unlimited storage capacity*** unlike Hopfield's continuous system [15], which like his discrete model, has limited capacity.

For the discrete system, assume that the $\vec{x}_i$ are composed of elements $\pm 1$ and replace the Euclidean distance in ( 6) with the normalized Hamming distance $\mid \vec{\mu_1} - \vec{\mu_2} \mid \stackrel{\Delta}{=} \frac{1}{N}\sum_{j=1}^{N} \mid \mu_j^1 - \mu_j^2 \mid$. This places the vectors $\vec{x}_i$ on the unit hypersphere.

The relaxation process for the discrete system will be of the type defined in Hopfield's model in equation( 3). Choose at random a component to be updated (that is, a neighbor $\vec{\mu}'$ of $\vec{\mu}$ such that $\mid \vec{\mu}' - \vec{\mu} \mid = \frac{2}{N}$), calculate the "energy" difference, $\delta\xi = \xi(\vec{\mu}') - \xi(\vec{\mu})$, and only if $\delta\xi < 0$, change this component, that is:

$$\mu_i \rightarrow \mu_i \cdot \theta(\xi(\vec{\mu}') - \xi(\vec{\mu})), \tag{8}$$

where $\xi(\vec{\mu})$ is the potential energy in (6). Since there is a finite number of possible $\vec{\mu}$ vectors ($2^N$), convergence in finite time is guaranteed.

This relaxation procedure is rigid since the movement is limited to points with components $\pm 1$. Therefore, although the local minima of $\xi(\vec{\mu})$ defined in ( 6) are only at the desired points $\vec{x}_i$, the relaxation may get stuck at some $\vec{\mu}$ which is not a stationary point of $\xi(\vec{\mu})$. However, the short-range behaviour of the potential $\xi(\vec{\mu})$, unlike the long-range behaviour of the quadratic potential used by Hopfield (equation ( 2)), gives rise to results similar to those we have quoted for the continuous model (equation ( 7)).

Specifically, let the stored memories $\vec{x}_1, ..., \vec{x}_m$ be separated from one another by having at least $\rho N$ different components ($0 < \rho \leq \frac{1}{2}$ and $\rho$ fixed), and let $\vec{\mu}(0)$ agree up to at least one $\vec{x}_i$ with at most $\theta \rho N$ errors between them ($0 \leq \theta < 1/2$, with $\theta$ fixed), then $\vec{\mu}(0)$ converges monotonically to $\vec{x}_i$ by the relaxation procedure given in equation ( 8).

This result holds independently of m, provided that N is large enough (typically, $N\rho \ln(\frac{1-\theta}{\theta}) \geq 1$) and L is chosen so that $\frac{N}{L} \leq \ln(\frac{1-\theta}{\theta})$. The proof is constructed by bounding the cummulative effect of terms $\mid \vec{\mu} - \vec{x}_j \mid^{-L}$, $j \neq i$, to the energy difference $\delta\xi$ and showing that it is dominated by $\mid \vec{\mu} - \vec{x}_i \mid^{-L}$. For details, we refer the reader again to [14].

Note the importance of this property: unlike the Hopfield model which is limited to $m \leq N$, the suggested system is optimal in the sense of Information Theory, since for every set of memories $\vec{x}_1, ..., \vec{x}_m$ separated from each other by a Hamming distance $\rho N$, up to $\frac{1}{2}\rho N$ errors in the input can be corrected, provided that N is large and L properly chosen.

As for the complexity of the system, we note that the nonlinear operation $a^{-L}$, for $a > 0$ and L integer (which is at the heart of our system computationally) is equivalent to $e^{-L\ln(a)}$ and can be implemented, therefore, by a simple electrical circuit composed of diodes, which have exponential input-output characteristics, and resistors, which can carry out the necessary multiplications.

Further, since both $\mid \vec{x}_i \mid$ and $\mid \vec{\mu} \mid$ are held fixed in the discrete system, where all states are on the unit hypersphere, $\mid \vec{\mu} - \vec{x}_i \mid^2$ is equivalent to the inner product of $\vec{\mu}$ and $\vec{x}_i$, up to a constant.

To conclude, the suggested model involves about $m \cdot N$ multiplications, followed by m nonlinear operations, and then $m \cdot N$ additions. The original model of Hopfield involves $N^2$ multiplications and additions, and then N nonlinear operations, *but is limited to* $m \leq N$. Therefore, whenever the Hopfield model is applicable the complexity of both models is comparable.

**Acknowledgements**

This work was supported by ONR contracts N00014-86-K0041 and N00014-85-K-0607, ARO contract DAAG-29-84-K-0262 and DAAG29-84-K-0082,

and the Weizmann postdoctoral fellowship.

## References

[1] Anderson, J. A. (1970)*Math. Biosciences* **8**, 137-160.

[2] Anderson, J. A. (1972) *Math. Biosciences* **4**, 197-220.

[3] Cooper, L. N (1974) in *Proc. Nobel Symposium on Collective Properties of Physical Systems*, ed. B. Lindquist and S. Lindquist, (Academic Press, New York), pp. 252-264.

[4] Kohonen, T. (1972) *IEEE Trans. on Computers, C.* **21**, 353-359.

[5] Kohonen, T. (1977) *Associative Memory: A System Theoretic Approach*, (Springer-Verlag, Berlin).

[6] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953) *Journ. Chem. Phys.*, **21:6**, 1087-1092.

[7] Hopfield, J. J. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2554-2558.

[8] Widrow, G. & Hoff, M. E. (1960) *Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record, Part 4*, 96-104.

[9] Kohonen, T. (1984) *Self-Organization and Associative Memory* (Springer-Verlag, Berlin), pp. 172-174.

[10] Reilly, D. E., Cooper, L. N & Elbaum, C. (1982) *Biol. Cybern.* **45**, 35-41.

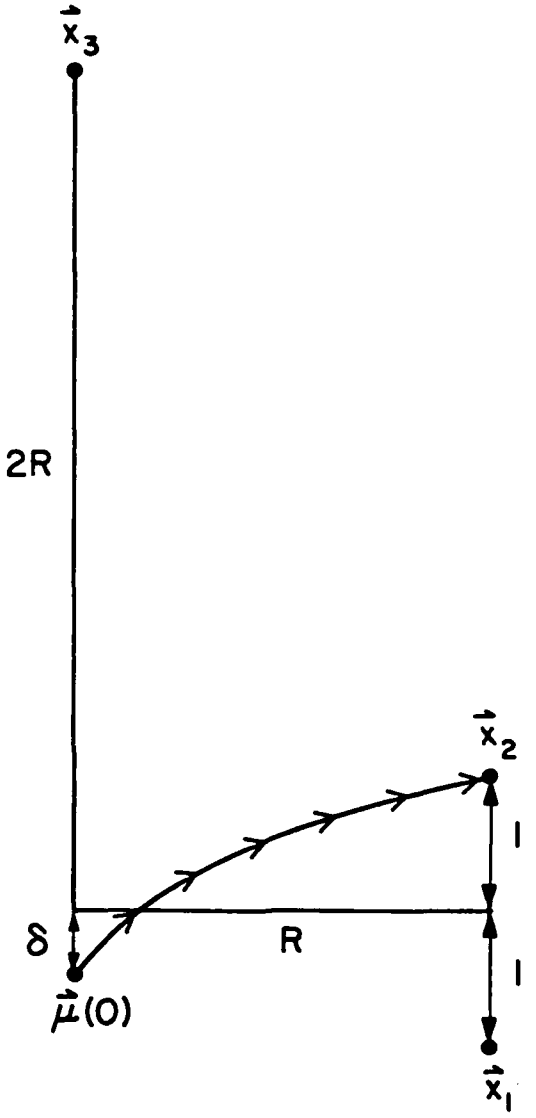
[11] Hopfield, J. J., Feinstein, D. I. & Palmer, R. G. (1983) *Nature* **304**, 158-159.

[12] Potter, T. W. (1987) Ph.d. Dissertation in Advanced Technology, S.U.N.Y. Binghampton, (unpublished).

[13] Bachmann, C.M. (1986) ARO Technical Report, Brown University, Center for Neural Science, Providence, R. I..

[14] Dembo, A., Zeitouni, O. (1987) ARO Technical Report, Brown University, Center for Neural Science, Providence, R.I., also submitted to *Phys. Rev. A.*

[15] Hopfield, J. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3088-3092.


$\vec{x}_3$
2R
$\vec{x}_2$
l
δ
R
l
$\vec{\mu}(0)$
$\vec{x}_1$

END

7-87

DTIC