

QUADRATIC HADAMARD MEMORIES I

Hendricus G. Loos

Laguna Research Laboratory
3015 Rainbow Glen, Fallbrook CA 92028-9765

December 1989

DTIC
ELECTE
JAN 26 1990
S D CS D

AD-A217 224

Sponsored by
Defense Advanced Research Projects Agency (DoD)
Aerospace Technology Office
"Adaptive Stochastic Content-Addressable Memory"
ARPA Order No. 6429
Issued by U.S. Army Missile Command Under
Contact # DAAH01-88-C-0887

Technical Report #1

Approved for public release; distribution unlimited

Name of Contractor: Laguna Research Laboratory
Business Address: 3015 Rainbow Glen, Fallbrook CA 92028-9765
Effective Date of Contract: 26 Sep 1988
Contract Expiration Date: 30 Aug 1990
Principal Investigator: Hendricus G. Loos, Tel. 619-728-8767
Work Title: "Adaptive Stochastic Content-Addressable Memory"

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research projects Agency or the U.S. Government.

90 01 25 0 27

ABSTRACT

-1-

)

9

$$(KR) \leftarrow$$

quadratic
us discre

FIELD
COPY
INSPECTED
1

Associative memories are to perform associative recall of stored vectors. The simplest such memory that has fault tolerance is the Hopfield memory [1], in which the connection matrix is the sum of the outer products of the stored vectors, with the diagonal removed. The problem with such memories is early saturation. This problem can be circumvented by using vectors with dilute information [2], or by employing a more sophisticated connection matrix [3,4]. The former of these approaches makes inefficient use of dimension, and the latter requires non-local learning rules, which complicate hardware implementations. Bidirectional Associative Memories [5-7] by

A-1

DISCLAIMER NOTICE

THIS DOCUMENT IS THE BEST
QUALITY AVAILABLE.

COPY FURNISHED CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

themselves do not solve the early saturation problem either [8].

The present DARPA SBIR project is an attempt to find an approach to associative memories that overcomes the early saturation problem in such a manner that efficient use is made of dimension, local learning can be used, and hardware implementations are simple. In Phase I of the project an approach was outlined that showed promise for reaching these objectives. It uses a modification of the Bidirectional Associative Memory (BAM) in which the rear threshold operation is replaced by something we have called a Dominant Label Selector (DLS). The idea is to choose backstates of the BAM as orthonormal vectors, which we consider to be labels for the stored frontstates. For any input state, the backstate of the BAM is then a linear combination of these labels, and the task of the DLS is to find the dominant label in this linear combination, and return it to the BAM, to be used in its backstroke. It is easy to see that the device performs perfect associative recall if the DLS works as required. To find such a DLS is the main challenge of the project.

A solution to this problem has been found in the present Phase II. The DLS conceived is an associative memory with an activation that is quadratic, rather than the customary linear (really, affine) function of the neuron signals. This approach falls under what is known as "higher-order neurons", which has been discussed in broad terms in the literature [9-12]. There is the immediate practical

question about the large number of connections (N^3) that appears to be required in such a case, but fortunately, the number of connections is found to have an upper bound of only N^2 , as for a Hopfield memory. The other practical question pertains to the hardware implementation of the signal product operation required in computing the quadratic activation. This question needs further work, but it should be mentioned already that we see a simple implementation by means of a field effect transistor operating in the unsaturated region.

The DLS acts on linear combinations of orthonormal labels, and returns the dominant label. Since the associative memory considered is bipolar, the orthonormal labels are bipolar. The matrix of a complete orthonormal set of bipolar vectors is known as a Hadamard matrix [13]. The DLS considered therefore may be called a "quadratic Hadamard memory". This report is about such memories. The questions discussed are stability and stable states, and the number of physical connections. These questions are investigated in the asynchronous discrete model.

In regard to conventions and notations, following Kosko [14], we call the input and output of the threshold function respectively the "activation" and the "signal" of the neuron.

If a neuron does not perform thresholding, then the activation is defined as the result of the summing operation done by the neuron.

The vectors $\hat{x}$ and $\hat{y}$ are bipolar, i.e., all their components are either 1 or -1.

In order to keep formulas simple, we have used the summation convention of tensor calculus where convenient. In order to distinguish from unsummed expressions, the summation convention has been used in its strict form [15]: in a product, summation over a repeated index is implied only if the index appears once as a subscript, and once as a super-

script. For instance, $u^a v_a$ is summed, but $u_a v_a$ is not.

Indices are used as follows. i, j , and k denote components in input space. a, b, c and p denote components in the space between the front end and the DLS, and also components of DLS state vectors. α, β , and γ are used to name stored vectors and Hadamard vectors. All indices range from 1 to N .

Indices are raised and lowered with the Kronecker delta as metric tensor, hence v^a and v_a have the same numerical value.

As a further simplification of appearance, 1 is often written as + and -1 as -, when no confusion with composition symbols can arise.

In the theorems, $\forall$ means "for all", $\exists$ means "such that", $\Rightarrow$ means "implies", and $\Leftarrow$ means "is implied by". $\exists$ means "there exists".

SELECTIVE REFLEXIVE MEMORY (SRM)

The associative memory under discussion consists of a front end followed by a Dominant Label Selector (DLS). The front end is much like a Bidirectional Associative Memory (BAM) [5]. However, there is no rear thresholding, and therefore the device is given a separate name, "Bidirectional Linear Transformer" (BLT). The BLT has thresholding of frontstates, as in the BAM. The whole machine, BLT plus DLS, is called "Selective Reflexive Memory" (SRM), in which "selective" indicates the selection of the dominant label by the DLS, and "reflexive" alludes to the signals passing back and forth through the BLT. The SRM is shown in Fig. 1.

In the discrete model, the BLT performs a linear transformation in both forward and backstrokes, the forward stroke being

$$u_b = B_b^i x_i, \quad (1)$$

where the bipolar vector x_i , $i=1$ to N , is the input vector, and u_b , $b=1$ to N , is the output of the BLT. The backstroke through the BLT gives

$$w_i = y_b B_{bi}^{\cdot}, \quad (2)$$

in which y_b is the bipolar vector returned by the DLS to the back of the BLT, and w_i is the result of the backstroke, appearing in front of the BLT. Subsequently, w_i is thresholded, and the result is a bipolar vector $x_i^!$. There is the option of using $x_i^!$ to upgrade the input x_i , as depicted in Fig. 2a, or to use it as output of the whole machine, as shown in Fig. 2b. Note that the linear transformations in front and backstrokes are the transpose of each other, just as in the BAM [5]. In hardware, the BLT is conveniently implemented by a bidirectional device capable of linear transformations of vectors, such as the Goodman optical matrix-vector multiplier [17], or an equivalent electronic device.

The connection matrix of the BLT is chosen as

$$B_{bi}^{\cdot} = h_b^{\alpha} q_{\alpha i}^{\cdot}, \quad (3)$$

where $\vec{q}_{\alpha}$, $\alpha=1$ to N , are the stored vectors, and $\vec{h}_{\alpha}$ are

Hadamard vectors. In (3), i and b are component indices. We have chosen the same dimension N for the BLT front and rear vector spaces, and have taken the number of stored states equal to N as well. It will be clear from the theory how to modify these choices if desired. The structure (3) of the BLT connection matrix B is Hebbian, i.e., it can be built up adaptively by Hebb learning. All adaption occurs in the BLT; the DLS is a fixed device.

The Hadamard vectors $\vec{h}_{\alpha}$ are rows of a Hadamard matrix, i.e., an orthogonal matrix (up to a scalar factor) with entries $+$ or $-$. Material on Hadamard vectors used in this report is shown in the Appendix. The Hadamard vector $\vec{h}_{\alpha}$ serves as a label for the stored state $\vec{q}_{\alpha}$. In its front stroke the BLT takes the N stored states $\vec{q}_{\alpha}$ into certain linear combinations of the orthogonal vectors $\vec{h}_{\alpha}$. The orthogonality of the $\vec{h}_{\alpha}$ allows for a clean backstroke through the BLT, in which a single label $\vec{h}_{\alpha}$ produces a single

stored state $\vec{q}_\alpha$. The action of the SRM is as follows. Let the input $\vec{x}$ be such that it is closest to stored vector $\vec{q}_\beta$. The BLT forward stroke gives, by (1) and (3),

$$u_b = h_b^\alpha q_\alpha^i x_i = c_\alpha h_b^\alpha, \quad (4)$$

where $c_\alpha = \vec{x} \cdot \vec{q}_\alpha$. c_β is the largest among the coefficients c_α , $\alpha=1$ to N . The vector $\vec{u}$ is presented to the input of the DLS, which selects from the linear combination (4) the Hadamard vector $\vec{h}_\beta$ which appears with the largest coefficient. The DLS returns $\vec{h}_\beta$ to the BLT, which performs the backstroke (2) with the result

$$w_\beta^i = h_\beta^b u_b = h_\beta^b h_b^\alpha q_\alpha^i = N \delta_{\beta\alpha} q_\alpha^i = N q_\beta^i, \quad (5)$$

where use has been made of (3) and (A2). Thresholding with the signum function s then gives

$$x'^i = s(w_\beta^i) = q_\beta^i, \quad (6)$$

which is, indeed, the stored vector that is closest to the DLS input $\vec{x}$.

Hence, if the DLS works as required, the SRM performs perfect associative recall of any one of N stored vectors. There are then no spurious stable states, and the memory can be loaded up fully to the dimension N . The question is, of course, how to construct a DLS with the required performance.

DOMINANT LABEL SELECTOR (DLS)

The DLS must select, from an input

$$u_b = c_\alpha h_b^\alpha, \quad (7)$$

the dominant Hadamard vector $\vec{h}_\beta$, i.e., the index β is such that c_β is maximum among the coefficients c_α in (7). Of course, $\vec{h}_\beta$ is then also the Hadamard vector with the largest scalar product $\vec{u} \cdot \vec{h}_\alpha$. Therefore, the DLS itself may be considered an associative memory with stored states $\vec{h}_\alpha$. In those terms, the DLS is to produce, from the input $\vec{u}$ of (7), the closest stored state, $\vec{h}_\beta$.

One may think that, since the stored vectors are orthogonal, a Hopfield memory will do fine in this case. That

is not so. The Hopfield matrix [1] would be

$$S_a^b = h_a^\alpha h_\alpha^b - N \delta_a^b, \quad (8)$$

but with (A1) this is zero. Leaving off the term $N \delta_a^b$ does

not help, since $h_a^\alpha h_\alpha^b$ is N times the identity operator, for which any vector is an eigenvector. The same results are found for any other complete orthonormal set of stored states, from the Spectral Theorem [16].

We have found that a DLS with the required properties may be obtained by using a quadratic activation,

$$v_a = S_{abc} y^b y^c + c r_a, \quad (9)$$

where the connection tensor S_{abc} is fully symmetric. The

last term in (9) may be seen either as external coupling to a vector r_a with coupling strength c , or as defining the

thresholds.

Neurons with activation (9) are a special case of the higher order neurons considered in the literature [9-12]. In the present discrete model, the upgrading of the output signal of neuron b from y_b to y'_b is done asynchronously

according to the assignment

$$\begin{aligned} y'_b &= + \text{ if } v_b > 0, \\ y'_b &= y_b \text{ if } v_b = 0, \\ y'_b &= - \text{ if } v_b < 0. \end{aligned} \quad (10)$$

The "energy" may be taken as

$$E = -(1/3) S_{abc} y^a y^b y^c - c r_a y^a, \quad (11)$$

which is a simple extension of the Hopfield energy [1]. Hopfield's argument showing that the energy cannot increase in the discrete model is easily modified [10] for the case of quadratic activation, as follows.

For a single Hamming step δy_p in the direction of the p th coordinate the energy (11) changes by

$$\delta_p E = -S_{abc} y^a y^b \delta y_p^c - S_{abc} y^a \delta y_p^b \delta y_p^c - (1/3) S_{abc} \delta y_p^a \delta y_p^b \delta y_p^c - c r_a \delta y_p^a. \quad (12)$$

The Hamming step $\delta_p \vec{y}$ in the direction of coordinate p may be written

$$\delta_p \vec{y} = 2e_p \delta_p^a \quad (13)$$

where the sign factor $e_p = \pm 1$ must be chosen such that the step $\delta_p \vec{y}$ keeps $\vec{y}$ in the hypercube. This requires that

$$e_p = -y_p. \quad (14)$$

In (13) δ_p^a is the Kronecker delta, so that $\delta_p^a = 1$ if $p=a$, and otherwise zero. With (13) and (14), the energy increment (12) due to a Hamming step in the p direction may be written

$$\delta_p E = -S_{abc} y^a y^b \delta_p^c - 4S_{app} y^a + (8/3) y_p S_{ppp} - c r_a \delta_p^a y^a. \quad (15)$$

Substituting the activation (9) gives

$$\delta_p E = -v^a \delta_p^a y^a - 4S_{app} y^a + (8/3) S_{ppp} y_p. \quad (16)$$

According to the upgrade scheme (10), $v_a \delta_p^a y^a$ is either

positive or zero, so that the first term in (16) cannot be positive in the asynchronous upgrading. Hence, if the connection tensor S_{abc} is restricted such that its components with at least two indices the same are zero, i.e.,

$$S_{app} = 0, \quad \forall a, \quad \forall p, \quad (17)$$

then the last two terms in (16) vanish, and we have the result

Theorem 1: The energy (11), with S subject to (17), cannot increase in asynchronous discrete upgrading.

This result is known [10], and it shows that the memory with quadratic activation with the connection tensor S discussed is stable in the asynchronous discrete model.

In addition to stability we require the absence of spurious stable states, something that is much harder to prove. We have such a proof for the case that the connection tensor is given by

$$S_{abc} = \sum_{\alpha} h_{\alpha a} h_{\alpha b} h_{\alpha c} - N \delta_{ab} \delta_c - N \delta_{bc} \delta_a - N \delta_{ca} \delta_b + 2N \delta_a \delta_b \delta_c, \quad (18)$$

where $\vec{h}_{\alpha}$ are Hadamard vectors. The first term on the r.h.s.

of (18) has a strict Hebbian form; the last four terms have been added in order to satisfy the condition (17), while retaining full symmetry. With (18), the activation (9) becomes

$$v_a = \sum_{\alpha} y_{\alpha}^2 h_{\alpha a} - 2N y_a y_1 - (N^2 - 2N) \delta_{a1} + c r_a, \quad (19)$$

where

$$y_{\alpha} = h_{\alpha}^a y_a. \quad (20)$$

From (20) we have

$$\sum_{\alpha} y_{\alpha}^2 = N^2, \quad (21)$$

where the orthonormality (A1) of Hadamard vectors has been used. (21) and the fact that h_1 has all + components has an important consequence for $a=1$ in (19):

$$v_1 = c r_1. \quad (22)$$

Choosing

$$c r_1 > 0, \quad (23)$$

v_1 of (22) is positive, and hence, by (10),

$$y_1 = +. \quad (24)$$

(24) means that only half of the hypercube $I_N = \{-1, 1\}^N$ is dynamically accessible to the DLS. In hardware, (24) is implemented by omitting neuron $a=1$, and by connecting the y_1 signal line to the logical "on" voltage.

The energy increment (16) due to a single Hamming step in the p direction may be written

$$\delta_p E = -v^a \delta_p y_a, \quad (25)$$

where (17) has been used. Because of (24), steps in the 1 direction are not considered:

$$p \neq 1. \quad (26)$$

The form (25) for the energy increment is reminiscent of the energy increment in thermodynamics, y_a and v_a being respec-

tively intensive and extensive variables. This typification of variables is consistent with the observation that as the neural net is scaled up in size, the signals y_a keep their

magnitude range, while the activations generally scale upward, due to the increase in number of signals to be summed.

For future use we write the Hadamard transform of the step $\delta_p y_a$,

$$\delta_p y_{\alpha} = h_{\alpha}^a \delta_p y_a; \quad (27)$$

with (13), (27) may be written

$$\delta_p y_{\alpha} = 2e h_{\alpha}^p. \quad (28)$$

STABLE POINTS OF THE DLS

We proceed to find the stable points. Calling

$$Q_a = \sum_{\alpha} y_{\alpha}^2 h_{\alpha a}, \quad (29)$$

and using (19), (13), and (24), Eq.(25) may be written

$$a \neq 1, \quad \oint_a E = -2e_a (Q_a + cr_a - 2Ny_a). \quad (30)$$

An alternate form for (30) is

$$a \neq 1, \quad \oint_a E = -2e_a (Q_a + cr_a) - 4N, \quad (31)$$

which is derived by using (14).

(21) and $h_{\alpha 1} = +, \forall \alpha$, have a consequence for $a=1$ in (29):

$$Q_1 = N^2. \quad (32)$$

In preparation for a determination of the energy minima we calculate bounds on Q_a . Define the index set

$$A = \{\alpha | y_{\alpha} \neq 0\}, \quad (33)$$

and the disjoint pieces

$$A^+ = \{\alpha | \alpha \in A \text{ \& } h_{\alpha a} = +\}, \quad (34)$$

and

$$A^- = \{\alpha | \alpha \in A \text{ \& } h_{\alpha a} = -\}, \quad (35)$$

where it is understood that the sets A^+ and A^- depend on the index a . (29) may then be written as

$$Q_a = \sum_{\alpha \in A^+} y_{\alpha}^2 - \sum_{\alpha \in A^-} y_{\alpha}^2. \quad (36)$$

One also has $\sum_{\alpha} y_{\alpha}^2 = N^2,$ (37)

as can be shown from (20), the bipolarity of $\vec{y}$, and the orthonormality of the Hadamard vectors. (37) may be written

$$N^2 = \sum_{\alpha \in A^+} y_{\alpha}^2 + \sum_{\alpha \in A^-} y_{\alpha}^2. \quad (38)$$

(36) and (38) imply

$$2 \sum_{\alpha \in A^+} y_{\alpha}^2 = N^2 + Q_a, \quad (39)$$

$$2 \sum_{\alpha \in A^-} y_{\alpha}^2 = N^2 - Q_a, \quad (40)$$

and since the left hand sides are nonnegative, this gives

$$-N^2 \leq Q_a \leq N^2 \quad (41)$$

We need to clarify the attainment of the bounds $-N^2$ and N^2 in (41). Anyone of these bounds is attained by Q_a , for some

index a , at the Hadamard points, with the exception of $\vec{h}_1$.

This may be seen as follows. For $\vec{y} = \vec{h}_\beta$ one has $y_\alpha = 0, \forall \alpha \neq \beta$, and $y_\beta = N$. (29) then gives

$$Q_a = N^2 h_{\beta a}. \quad (42)$$

For $\beta=1$ we have $h_{1a} = +, \forall a$, and it follows that $Q_a = N^2, \forall a$:

$$\vec{y} = \vec{h}_1 \implies Q_a = N^2, \forall a. \quad (43)$$

For $\beta \neq 1$, $\vec{h}_\beta$ has + as well as - components, and (42) shows

that $\beta \neq 1 \implies \exists a \ni Q_a = N^2,$

and $\beta \neq 1 \implies \exists a \ni Q_a = -N^2.$

(42) further shows that

$$Q_1 = N^2, \quad \forall \beta. \quad (43)$$

Hence, when discussing the case $Q_a = -N^2$, $a=1$ needs to be excluded.

For $Q_a = -N^2$ the left hand side of (39) is zero, and

hence the set A^+ is empty. This implies that $h_{\alpha a} = -$ for all α in A . Therefore we have

$$a \neq 1, \quad Q_a = -N^2 \implies h_{\alpha a} = -, \quad \forall \alpha \in A, \quad (45)$$

$$\text{and} \quad Q_a = N^2 \implies h_{\alpha a} = +, \quad \forall \alpha \in A, \quad (46)$$

as follows from a similar argument. The converse of (45) is also true, since from $h_{\alpha a} = -, \forall \alpha \in A$, it follows that $a \neq 1$ and

$$Q_a = \sum_{\alpha \in A} y_\alpha^2 h_{\alpha a} = - \sum_{\alpha \in A} y_\alpha^2 = -N^2,$$

where (37) has been used. Hence we have

Lemma 1: $a \neq 1, \quad Q_a = -N^2 \iff h_{\alpha a} = -, \quad \forall \alpha \in A.$

Next, we calculate the Hadamard components y_α given by (20).

Define the index set

$$B = \{b | y_b = -\} \quad (47)$$

Then we may write (20) as $y_\alpha = h_\alpha^b y_b = - \sum_{b \in B} h_\alpha^b + \sum_{b \notin B} h_\alpha^b$. (48)

The Hadamard vectors chosen have the property (A8):

$$\sum_b h_\alpha^b = N \delta_{\alpha 1} \quad (49)$$

It follows that $\sum_{b \notin B} h_\alpha^b = N \delta_{\alpha 1} - \sum_{b \in B} h_\alpha^b$, (50)

so that (48) yields

$$y_\alpha = N \delta_{\alpha 1} - 2 \sum_{b \in B} h_\alpha^b \quad (51)$$

Suppose $Q_a = -N^2$, $\forall a \in B$. Then, Lemma 1 and (51) give

$$\alpha \in A: y_\alpha = N \delta_{\alpha 1} + 2(N-W)/2 = N \delta_{\alpha 1} + N - W \quad (52)$$

where

$$W = \sum_b y_b \quad (53)$$

is the weight of the bipolar vector $\vec{y}$, and $(N-W)/2$ is the cardinality of the set B.

Hence we have

Lemma 2: $Q_a = -N^2$, $\forall a \in B \Rightarrow y_\alpha = N \delta_{\alpha 1} + N - W, \forall \alpha \in A$.

There appear to be two cases:

Case 1: $\alpha=1 \in A$, and Case 2: $\alpha=1 \notin A$.

Case 1: $\alpha=1 \in A$. Calculating y_α of Lemma 2, we have

$y_1 = 2N - W = W$, from (20), (53), and the fact that $\vec{h}_1$ has all + components. It follows that $W=N$, so that $\vec{y} = \vec{h}_1$, the all positive vector. However, for $\vec{y} = \vec{h}_1$, one has $Q_a = N^2$ by (43), in contradiction with $Q_a = -N^2$ of Lemma 2. Hence, Case 1 is not possible within the premises of Lemma 2.

Case 2: $\alpha=1 \notin A$. From (20) for $\alpha=1$ and (53) it follows that $W=0$. Hence, for $\alpha \neq 1$, the y_α of Lemma 2 is just N. With (37) it follows that

$$N^2 = \sum_{\alpha \in A} y_\alpha^2 = r N^2 \quad (54)$$

where r is the cardinality of the set A. Since (54) implies

$r=1$, the set A contains only a single element, say β . It follows that $\vec{y} = \vec{h}_\beta$, and with Lemma 2 we may conclude

$$Q_a = -N^2, \forall a \in B \Rightarrow \vec{y} \text{ is Hadamard } \vec{h}_1.$$

The converse is also true, since for $\vec{y} = \vec{h}_\beta, \beta \neq 1$, we have $\vec{y}_\alpha = 0$,

$\alpha \neq \beta$, and $y_\beta = N$, so that

$$Q_a = N^2 h_{\beta a}. \quad (55)$$

For index a such that $y_a = h_{\beta a} = -$ it follows that $Q_a = -N^2$.

Hence we have

Theorem 2: $Q_a = -N^2, \forall a \ni y_a = - \iff \vec{y} \text{ is Hadamard } \vec{h}_1.$

With Lemma 1, this theorem also can be put in the form

Theorem 3: $h_{\alpha a} = -, \forall \alpha \in A \text{ and } \forall a \ni y_a = -$
 $\iff \vec{y} \text{ is Hadamard } \vec{h}_1.$

The condition $y_a = -$ in Theorems 2 and 3 implies that $e_a = +$, by (14). Using $Q_a = -N^2$ and $e_a = +$ in (31) gives

$$\oint_a E = 2N^2 - 2cr_a - 4N. \quad (56)$$

We want to choose the threshold arrangement, comprised by c and $\vec{r}$, such that $\oint_a E$ of (56) is positive for all a involved, i.e., a such that $e_a = +$. This is the case if

$$cr_a < N^2 - 2N, \quad (57)$$

and $N > 2; \quad (58)$

the dimension N is henceforth restricted in this manner.

We wish the left hand side of (57) to be independent of a ; this is the case if $\vec{r}$ is chosen as the all positive bipolar

vector, $\vec{r} = \vec{h}_1. \quad (59)$

With this choice, (57) becomes

$$c < N^2 - 2N \quad (60)$$

For the Hamming steps considered above we have $e_a = +$. It remains to investigate steps with $e_a = -$, $a \neq 1$, made from a Hadamard point. For $\vec{y} = \vec{h}_\beta$, $Q_a = N^2 h_{\beta a}$. By (14), for a step with $e_a = -$ we need $y_a = h_{\beta a}$, and hence $Q_a = -N^2$. Substitution in (30) gives

$$a \neq 1, \quad \Delta_a E = 2(N^2 + c - 2N) \quad (61)$$

We want the energy increment (61) to be positive as well, for all steps considered in this case, i.e., a such that $e_a = -$. This requires that

$$c > -(N^2 - 2N) \quad (62)$$

(60) and (62) show that stability of the Hadamard points requires the common threshold c to satisfy the condition

$$-(N^2 - 2N) < c < N^2 - 2N \quad (63)$$

The upper bound is positive if (58) is satisfied. With (23), (59), and the fact that $\vec{h}_i$ has all positive components, we can sharpen (63) to

$$0 < c < N^2 - 2N \quad (64)$$

Hence we have

Theorem 4: For a quadratic Hadamard memory with asynchronous discrete dynamics given by (9) and (10), with connection tensor (18), and a threshold c subject to (64), the Hadamard points are stable equilibria.

It remains to be shown that a threshold c subject to (64) exists such that there are no other stable points besides Hadamard points. This is done by showing that the possible values of Q_a have a gap of magnitude $4N$ at the lower bound

$-N^2$. With (41) this implies that there is a gap of magnitude $8N$ at the upper bound of the energy increments $\Delta_a E$. This means

that the threshold c can be chosen such that the energy increment gap straddles the origin. We then have the desirable situation that for Hadamard points, and only for those, $\Delta_a E > 0 \quad \forall a \neq 1$; for any other points $\vec{y}$ there always is an index a such that $\Delta_a E < 0$. This, of course, shows that the Hadamard

points are the only stable points of the DLS. We proceed with

the details of the outlined steps.

By Theorem 2, for $\vec{y}$ Hadamard $\vec{h}_1$ and for Hamming steps in the direction a such that $e_a = +$ (and therefore $y_a = -$, by (14)) we have $Q_a = -N^2$. Writing

$$c = N^2 - 2N - K, \quad 0 < K < 2N^2 - 4N, \quad (65)$$

(56) gives

$$\delta_a E = 2K, \quad (66)$$

where (59) has been used. Vice versa,

$$\delta_a E = 2K, \quad \forall a \ni e_a = + \Rightarrow Q_a = -N^2 \Rightarrow \vec{y} \text{ is Hadamard } \vec{h}_1, \quad (67)$$

where use has been made of (14), (31), (59), and Theorem 2. Hence we have

$$\text{Theorem 5: } \vec{y} \text{ is Hadamard } \vec{h}_1 \iff \delta_a E = 2K, \quad \forall a \ni y_a = -. \quad (68)$$

Next, we calculate the increment $\delta_p Q_a$ of Q_a , $a \neq 1$, due to a single Hamming step $\delta_p \vec{y}$ in the direction $p \neq 1$. From (29) one has

$$\delta_p Q_a = 2 \sum_{\alpha} y_{\alpha} \delta_p y_{\alpha} h_{\alpha a} + \sum_{\alpha} \delta_p y_{\alpha}^2 h_{\alpha a}. \quad (69)$$

With (28), the first term in (69) may be written

$$2 \sum_{\alpha} y_{\alpha} 2e_p h_{\alpha p} h_{\alpha a} = 4e_p \sum_{\alpha} y_{\alpha} h_{\alpha p} h_{\alpha a}. \quad (70)$$

The group property of the Hadamard vectors, discussed in the Appendix, makes the (component-wise) product of two Hadamard vectors a Hadamard vector, i.e.,

$$h_{\alpha p} h_{\alpha a} = h_{\alpha c}, \quad \forall \alpha, \quad (71)$$

where $c = f(p, a)$; see (A16). For the present purpose we do not need to know what this function is. Since our proof of the group property given in the Appendix requires that the dimension is a power of 2, N will henceforth be restricted in this manner. Furthermore, in order that (58) be satisfied, the power is restricted to be at least 2. Using (71) in (70) gives

$$4e_p \sum_{\alpha} y_{\alpha} h_{\alpha c} = 4e_p N y_c, \quad (72)$$

by applying the inverse of the Hadamard transform (20).

With (28) and (A7), the second term in (69) may be written

$$\sum_{\alpha} \delta y_{p\alpha}^2 h_{\alpha a} = \sum_{\alpha} (2e_{p\alpha} h_{\alpha p})^2 h_{\alpha a} = 4 \sum_{\alpha} h_{\alpha a} = 0, \quad (73)$$

since $a \neq 1$. (72) and (73) give for the increment δQ_{pa} of (69)

$$a \neq 1, p \neq 1, \quad \delta Q_{pa} = 4e_p N y_c. \quad (74)$$

which shows that the only possible values for δQ_{pa} are $\pm 4N$.

With Theorem 2, (41), and the fact that any of the two bounds N^2 and $-N^2$ in (41) are attained by Q_a for some $a \neq 1$, it follows that for $a \neq 1$ the range of Q_a is included in the set

$$\{-N^2, -N^2 + 4N, -N^2 + 8N, \dots, N^2\}. \quad (75)$$

With (31) and (66) this implies that the range of energy increments δE_a is contained in

$$\{2K, 2K-8N, 2K-16N, \dots, 2K-8N^2\}. \quad (76)$$

(76) shows that the energy increment gap is $8N$. The gap may be placed symmetrically over the origin by choosing K such that

$$2K-8N = -2K; \quad (77)$$

this is the case if $K=2N$. (78)

With (78) and (65) we have for the threshold

$$c = N^2 - 4N. \quad (79)$$

The symmetric placing of the energy increment gap maximizes fault tolerance of the associative memory.

With (78), the energy increment range is contained in the set

$$\{4N, -4N, -12N, \dots, 4N-8N^2\}. \quad (80)$$

Theorem 5 together with (78) show that only at Hadamard points the increments δE_a are positive for all steps with

$e_a = +$. For other points $\vec{y}$, there always is a Hamming step

which gives a negative energy increment. With Theorem 1 it follows that such points are not stable. Hence, we have proved

Theorem 6: For the asynchronous discrete quadratic Hadamard memory with connection tensor (18) and threshold (79) the stable points are just the Hadamard points.

Evidently, there are no spurious stable states in the quadratic Hadamard memory considered. In contrast, for an N -dimensional Hopfield memory with N stored states, all taken as Hadamard, every state is a spurious stable state,

excepting of course the Hadamard states themselves.

It follows that an N-dimensional SRM with the DLS taken as a quadratic Hadamard memory has perfect associative recall of up to N bipolar stored vectors. The latter can be chosen arbitrarily. There are no spurious states.

ALTERNATE SYSTEM

There is a variation on the system considered above, which gives the same main results. In the ongoing research we stumbled first upon the alternate system, and much computer work was done in its exploration.

In the alternate system the connection tensor is taken to have the strictly Hebbian form

$$S_{abc} = \sum_{\alpha} h_{\alpha a} h_{\alpha b} h_{\alpha c}, \quad (81)$$

without any of the subtractions of (18). The activation is then simply

$$v_a = Q_a + c, \quad (82)$$

where (59) has been used. With (43) we have

$$v_i = N^2 + c; \quad (83)$$

if c is restricted by (64), then v_i is positive, and thus

$$y_i = +, \quad (24)$$

as for the subtracted tensor.

The energy increment (16) involves the terms with S_{app} and S_{ppp} , which must be calculated, since the connection tensor is "unsubtracted". For these terms we find for $p \neq 1$

$$-4S_{app} y^a = -4 \sum_{\alpha} h_{\alpha a} h_{\alpha p} h_{\alpha p} y^a = -4 \sum_{\alpha} h_{\alpha a} y^a = -4N \delta_{a1} y^a = -4Ny_1 = -4N$$

by (A7) and (24), and

$$(8/3)y_p S_{ppp} = (8/3)y_p \sum_{\alpha} (h_{\alpha p})^3 = (8/3)y_p \sum_{\alpha} h_{\alpha p} = (8/3)y_p N \delta_{p1} = 0,$$

since $p \neq 1$. Hence, for the unsubtracted connection tensor (81) and for $p \neq 1$, (16) becomes

$$p \neq 1, \quad \delta_p E = -v^a \delta_p y_a - 4N. \quad (84)$$

Since $v^a \delta_p y_a \geq 0$ by (10), and the extra term is negative, we have

Theorem 1': For the quadratic Hadamard memory with unsubtracted connection tensor (81), the energy (11) cannot increase in asynchronous discrete upgrading.

The increment of the energy (11) due to a single Hamming step in the direction $a \neq 1$ may also be written

$$\Delta_a E = -2e_a (Q_a + c) - 4N \quad , \quad (85)$$

where (82) has been used. Since (85) is identical to (31) with (59), the arguments leading to Theorem 6 go through for the alternate system, and we have

Theorem 6': For the asynchronous discrete model of the quadratic Hadamard memory with connection tensor (81) and threshold (79), the stable points are just the Hadamard points.

COMPUTER SEARCH FOR STABLE POINTS

As a check on Theorems 6 and 6' we have performed a search for stable points $\bar{y}$ for a discrete quadratic Hadamard memory of dimension 16. The program involves a search for equilibrium points $\bar{y}$ of the hypercube $I_{16} = \{-1, 1\}^{16}$. For each such point $\bar{y}$ the activations (9) of the 15 neurons $a \neq 1$ are calculated, and (10) is used to see if upgrading causes any change in the signals y . If there is no change, then $\bar{y}$ is an equilibrium point of the quadratic Hadamard memory considered. If $\bar{y}$ is found to be an equilibrium point, a subroutine is called which computes the energy (11) at $\bar{y}$ and at all the neighboring points one Hamming step away from $\bar{y}$. If $E(\bar{y})$ is smaller than the energy at each of these neighboring points, the equilibrium point $\bar{y}$ is stable, in the discrete sense, and the case is reported. The search extends over all 64K points of the hypercube I_{16} .

Programs have been run using the subtracted connection tensor (18), as well as the unsubtracted tensor (81). For both cases, the computer runs show as only stable points the Hadamard points, in agreement with Theorems 6 and 6'.

CONNECTIONS IN THE QUADRATIC HADAMARD MEMORY

There is a practical concern about the number of physical connections required in the quadratic Hadamard memory. Physical connections are required between neurons a , b , and c , if $S_{abc} \neq 0$. Fortunately, for S_{abc} given by (18)

or (81), most components are zero. This is due to the group property of the Hadamard vectors.

For the connection tensor given by (81):

$$S_{abc} = \sum_{\alpha} h_{\alpha a} h_{\alpha b} h_{\alpha c} \quad (81)$$

the product of two Hadamard vectors may be replaced by a single Hadamard vector, by virtue of the group property discussed in the Appendix:

$$h_{\alpha b} h_{\alpha c} = h_{\alpha d} \quad , \quad (86)$$

where $d=f(b,c)$ is the structure function of the group. Substitution in (81) and using the group property once more,

$$h_{\alpha a} h_{\alpha d} = h_{\alpha e} \quad , \quad (87)$$

$e=f(a,d)$, gives

$$S_{abc} = \sum_{\alpha} h_{\alpha e} = N \delta_{e1} \quad , \quad (88)$$

where (A7) has been used in the last step. (88) shows that S_{abc} is nonzero only for $e=1$. By (87) this requires that

$$a=d, \quad (89)$$

as follows from the material on the Hadamard group shown in the Appendix. It follows that S_{abc} is nonvanishing only if

$$a=f(b,c) \quad . \quad (90)$$

Hence we have

Theorem 7: For connection tensor (81) the quadratic

Hadamard memory requires N^2 physical connections; each of these has the weight N .

For the subtracted connection tensor (18) the number of nonvanishing connections is further reduced by N , the number of pairs (b,c) for which $b=c$. Hence we have

Theorem 8: The quadratic Hadamard memory with subtracted

connection tensor (18) requires $N^2 - N$ physical connections; all of these have weight N .

The quadratic Hadamard memory, serving as a DLS in our SRM, achieves the desirable absence of spurious states by using quadratic activation. It is fortunate that this can be done without the proliferation of physical connections which usually accompanies higher-order neurons [9]. The fact that all the nonzero weights have the same value N very much simplifies the DLS hardware: the interconnects may be taken just as wires, since the factor N is not essential.

The main complication facing us is the implementation of the signal product operation required for computing the quadratic activation. Such signal multiplication can be done with a field effect transistor which is made to operate in the unsaturated region. The details of this operation need to be studied. It is expected that considerable deviations from a strict product can be permitted, because of the fault tolerance of the quadratic Hadamard memory.

COUPLING BETWEEN BLT AND DLS

The forward-stroke output $\vec{u}$ of the BLT must be coupled to the DLS in such a manner that the DLS can extract the dominant Hadamard vector in $\vec{u}$. We have been deliberately vague about this coupling, for the following reason.

The BLT forward-stroke output $\vec{u}$ is generally not bipolar. However, the BLT model used in this report is discrete, so that the BLT signal state $\vec{y}$ is bipolar. Hence, $\vec{u}$ cannot be used as initial value for $\vec{y}$. Of course, one could threshold $\vec{u}$ and apply the result as initial $\vec{y}$. That would be a pity though, because in the thresholding the fidelity of the Hadamard expansion comprising $\vec{u}$ would be lost. $\vec{u}$ must be coupled to the DLS in some other way. We have found how to do this, but the description requires the continuum model which allows $\vec{y}$ to have as components real numbers in the range between -1 and 1 , and accounts for input capacitance and resistance of amplifiers. In this model, the DLS dynamics is described, in normalized form, by

$$\dot{v}_a = -v_a + S_{abc} y^b y^c, \quad (91)$$

$$y_a = s(v_a), \quad (92)$$

where the dot denotes differentiation with respect to time, and s is a sigmoid function. The forward coupling between BLT and DLS can be arranged such that $\mu \vec{u}$ is applied as initial value for $\vec{v}$, μ being a coupling constant. Choosing μ is a delicate matter which will be discussed in a future report entitled "Quadratic Hadamard Memories II". This report will deal with the continuum model of the SRM, the DLS, and of the entire system, SRM.

APPENDIX

Hadamard vectors

In the report we use bipolar vectors which form an orthonormal set. A matrix, the rows of which are such vectors, is called a Hadamard matrix [13]. Such a matrix is orthogonal, up to a scalar factor N, the dimension. Hadamard matrices have an important application in optimal multiplexing [13]; they are used in spectrometers and imagers in order to improve the signal-to-noise ratio. Optics employing this method is called "Hadamard transform optics" [13]. Hadamard matrices are also used in error-correcting codes [13]. It is not surprising that they can be used to advantage in neural nets as well.

The rows of a Hadamard matrix are here called "Hadamard vectors". They are special elements of the hypercube that form an orthogonal set. The N-dimensional Hadamard vectors used here are denoted by $\vec{h}_\alpha$, $\alpha = 1$ to N. Labeling the vector components by the index $b=1$ to N, the Hadamard matrix with rows $\vec{h}_\alpha$ has the elements $h_{\alpha b}$. The orthonormality of the Hadamard vectors is expressed by

$$h_{\alpha a} h_{\alpha b} = N \delta_{ab}, \quad (A1)$$

where δ_{ab} is the Kronecker. From (A1) and the linear independence of the Hadamard vectors a second set of orthonormality conditions can be derived:

$$h_{\alpha a} h_{\beta a} = N \delta_{\alpha\beta}. \quad (A2)$$

For use in this report we restrict the Hadamard matrix to be symmetric, and to have their first components equal to +. For dimensions N that are powers of 2 there are the so-called "Sylvester-type Hadamard matrices" [13], which are defined recursively by the scheme

$$H_{2N} = \begin{pmatrix} H_N & H_N \\ H_N & -H_N \end{pmatrix}, \quad (A3)$$

$$H = \begin{pmatrix} + & + \\ + & - \end{pmatrix}. \quad (A4)$$

For example, the Sylvester-type Hadamard matrix of dimension 16 is shown below.

$$H_{16} = \begin{matrix} + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\ + & - & + & - & + & - & + & - & + & - & + & - & + & - & + \\ + & + & - & - & + & + & - & - & + & + & - & - & + & + & - \\ + & - & - & + & + & - & - & + & + & - & - & + & + & - & + \\ + & + & + & + & - & - & - & - & + & + & + & + & - & - & - \\ + & - & + & - & - & + & - & + & + & - & + & - & - & + & - \\ + & + & - & - & - & - & + & + & + & + & - & - & - & + & + \\ + & - & - & + & - & + & + & - & + & - & - & + & - & + & + & - \\ + & + & + & + & + & + & + & + & - & - & - & - & - & - & - & - \\ + & - & + & - & + & - & + & - & - & + & - & + & - & + & - & + \\ + & + & - & - & + & + & - & - & - & - & + & + & - & - & + & + \\ + & - & - & + & + & - & - & + & - & + & + & - & - & + & + & - \\ + & + & + & + & - & - & - & - & - & - & - & - & + & + & + & + \\ + & - & + & - & - & + & - & + & - & + & - & + & + & - & + & - \\ + & + & - & - & - & - & + & + & - & - & + & + & + & + & - & - \\ + & - & - & + & - & + & + & - & - & + & + & - & - & + & - & + \end{matrix} \quad (A5)$$

There is a second type of Hadamard matrices. These matrices are constructed from so-called cyclic S-matrices [13], and they exist only for certain dimensions N, including all powers of 2. The construction of the matrix starts with choosing the first Hadamard vector to have all components +. The remaining N-1 Hadamard vectors are found by taking their first component as +, and by taking the remaining N-1 components as left shifts of a special N-1 dimensional bipolar vector z_{N-1} , the construction of which is discussed by Harwit and Sloane [13]. They also show a list of such vectors for several small values of N. Some examples, taken from [13] are:

N-1	z_{N-1}
3	--+
7	---+---
11	--+---+---+
15	+++---+---+---

The Hadamard matrix constructed from z for N=16 is

$$H_{16} = \begin{matrix} + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\ + & + & + & + & - & + & + & - & - & + & - & + & - & - & - \\ + & + & + & - & + & + & - & - & + & - & + & - & - & - & + \\ + & + & - & + & + & - & - & + & - & + & - & - & - & - & + \\ + & - & + & + & - & - & + & - & + & - & - & - & - & + & + \\ + & + & + & - & - & + & - & + & - & - & - & - & + & + & + \\ + & + & - & - & + & - & + & - & - & - & - & + & + & + & - \\ + & - & - & + & - & + & - & - & - & - & + & + & + & - & + \\ + & - & + & - & + & - & - & - & - & + & + & + & - & + & + \\ + & + & - & + & - & - & - & - & + & + & + & - & + & + & - \\ + & - & + & - & - & - & - & + & + & + & - & + & + & - & + \\ + & + & - & - & - & - & + & + & + & - & + & + & - & - & + \\ + & - & - & - & - & + & + & + & - & + & + & - & - & + & - \\ + & - & - & - & + & + & + & - & + & + & - & - & + & - & + \\ + & - & - & + & + & + & - & + & + & - & - & + & - & + & - \\ + & - & + & + & + & - & + & + & - & - & + & - & + & - & - \\ + & - & + & + & + & - & + & + & - & - & + & - & + & - & - \end{matrix} \quad (A6)$$

One has (see [13]),

Theorem A1: If N is a power of 2, the Hadamard matrix constructed from a cyclic S matrix can be transformed into a Sylvester-type Hadamard matrix by permutation of rows and of columns.

This theorem is mentioned here because it will be needed below. For the Hadamard vectors $\underline{h}_\alpha$ used in this report, we have

$$\sum_{\alpha} h_{\alpha a} = N \delta_{a1} ; \quad (A7)$$

by the symmetry of the Hadamard vectors used here, this may also be written as

$$\sum_a h_{\alpha a} = N \delta_{\alpha 1} . \quad (A8)$$

Group property of Hadamard vectors

For dimension $N=2^k$, k natural, we have the Sylvester construction of Hadamard vectors, as discussed above. The component-wise product of any two such Hadamard vectors is a Hadamard vector:

$$h_{\alpha}^b h_{\beta}^b = h_{\gamma}^b , \quad \forall b , \quad (A9)$$

where $\gamma = f(\alpha, \beta)$. This may be shown by induction, as follows.

The property holds for $N=2$, since by (A9) we have then $\underline{h}_1 = ++$, and $\underline{h}_2 = +-$, and (A9) is true, as can be seen by

inspection.

Assume that the group property holds for Sylvester-type Hadamard vectors of dimension N , a power of 2. Construct a $2N$ -dimensional Hadamard matrix by means of the Sylvester construction:

$$H_{2N} = \begin{array}{|c|c|} \hline \begin{array}{c} H \\ N \end{array} & \begin{array}{c} H \\ N \end{array} \\ \hline \begin{array}{c} H \\ N \end{array} & \begin{array}{c} -H \\ N \end{array} \\ \hline \end{array} \quad (A10)$$

where H_N is the Sylvester-type Hadamard matrix of dimension N . Consider an index pair α, β to be used in the product $h_\alpha^b h_\beta^b$ on the left side of (A9). We have the following cases (see Fig. 3):

1) $\alpha \leq N, \beta \leq N$. This means that the Hadamard vectors $\vec{h}_\alpha$ and $\vec{h}_\beta$ both lie in the upper half of the matrix (A10). Since the group property holds for $H_N, \exists \gamma \leq N \ni (A9)$ holds for $b=1$ to N . But for the upper half of H_{2N} , the matrix H_N of the left half is repeated on the right, $h_\alpha^b = h_\alpha^{N+b}, b=1$ to N . It follows that (A9) is true for all $b=1$ to $2N$.

2) $\alpha > N, \beta > N$. Hadamard vectors $\vec{h}_\alpha$ and $\vec{h}_\beta$ then lie both in the lower half of H_{2N} . Since the group property holds for dimension $N, \exists \gamma > N \ni (A9)$ is true for $b=1$ to N . But, in the lower half of H_{2N} we have

$$h_\alpha^b = -h_\alpha^{N+b}.$$

Hence, for $c=N+1$ to $2N, h_\alpha^c h_\beta^c = -h_\gamma^c$.

Subtract N from $\gamma: \gamma' = \gamma - N$. Then,

$$\text{for } b=1 \text{ to } N, h_{\gamma'}^b = h_{\gamma}^b,$$

and for $b=N+1$ to $2N, h_{\gamma'}^b = -h_{\gamma}^b$, because of the

block structure (A10). It follows that for all $b=1$ to $2N$ we have

$h_\alpha^b h_\beta^b = h_{\gamma'}^b$; hence, (A9) is true for dimension $2N$ in this case.

3) $\alpha \leq N, \beta > N$. Let $\alpha' = N + \alpha$. Then, $\exists \gamma > N$ such that,

$$\text{for } b=1 \text{ to } N, h_{\alpha'}^b h_\beta^b = h_\gamma^b. \quad (A11)$$

Let $\gamma' = \gamma - N$. Then,

$$\begin{aligned} b \leq N, \quad h_{\alpha'}^b &= h_\alpha^b, \\ h_{\gamma'}^b &= h_\gamma^b, \\ h_\alpha^b h_\beta^b &= h_{\gamma'}^b. \end{aligned}$$

and (A11) gives

$$h_{\alpha'}^b h_\beta^b = h_{\gamma'}^b. \quad (A12)$$

For $c > N$, we have

$$\begin{aligned} h_{\alpha'}^c &= -h_\alpha^c, \\ h_{\gamma'}^c &= -h_\gamma^c, \end{aligned}$$

and (A11) gives

$$h_{\alpha'}^c h_\beta^c = h_{\gamma'}^c. \quad (A13)$$

(A12) and (A13) give

$$a=1 \text{ to } 2N, \quad h_{\alpha}^a h_{\beta}^a = h_{\gamma}^a, \quad (A14)$$

Since the product in (A9) is commutative, the remaining case, $\alpha > N, \beta \leq N$ reduces to case 3).

Hence, the group property (A9) holds for dimension $2N$. By induction it follows that the property holds for all Sylvester-type Hadamard matrices.

In the DLS, we have used the symmetric Hadamard matrices constructed from cyclic S matrices. By Theorem A1, if N is a power of 2, these matrices can be transformed to Sylvester-type matrices by index permutations. We just have shown that for the latter type Hadamard matrices the group property holds. Index permutations do not affect the group property. Hence we have

Theorem A2: For dimensions that are a power of 2, the Hadamard vectors constructed from cyclic S matrices form a group under component-wise multiplication.

We call the group under discussion the "Hadamard group". In this group, the first Hadamard vector, which has all components +, acts as the identity. Every Hadamard vector is its own inverse. The function $\gamma = f(\alpha, \beta)$ which gives the "product index" γ for any pair of "factor indices" α, β may be called the "structure function" of the group. Since the component multiplication in (A9) is commutative, the Hadamard group is abelian.

Since the Hadamard matrices used here are symmetric, the group property may also be expressed as

$$h_{\alpha}^a h_{\alpha}^b = h_{\alpha}^c, \quad \forall \alpha, \quad (A16)$$

with $c=f(a,b)$.

REFERENCES

- [1] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", Proc. Natl. Acad. Sci. USA 79, 2554 (1982).
- [2] D. J. Amit, H. Gutfreund, and H. Sompolinski, "Information storage in neural networks with low levels of activity", Phys. Rev. A35, 2293 (1987).

- [3] L. Personnaz, I. Guyon, and G. Dreyfus, "Collective computational properties of neural networks: New learning mechanisms", Phys. Rev. A34, 4217 (1986).
- [4] I. Kanter and H. Sompolinsky, "Associative recall of memory without errors", Phys. Rev. A35, 380 (1987).
- [5] B. Kosko, "Competitive Adaptive Bidirectional Associative Memories", Proc. IEEE Int'l Conf. on Neural Networks (ICNN-87) vol. II, 261 (1987).
- [6] B. Kosko, "Bidirectional Associative Memories", IEEE Trans. Systems, Man, Cybernetics, vol. SMC-18, 49 (1988).
- [7] B. Kosko, "Adaptive Bidirectional Associative Memories", Applied Optics vol. 26, No. 23, 4947 (1987).
- [8] H. G. Loos, "Reflexive Associative Memories", "Neural Information Systems", Dana. Z. Anderson, Ed., p. 495, American Institute of Physics, New York, 1988.
- [9] M. L. Minsky and S. Papert, PERCEPTRONS, MIT Press, Cambridge, MA, 1969.
- [10] H. H. Chen, Y. C. Lee, G. Z. Sun, and H. Y. Lee, "High Order Correlation Model for Associative Memory", "Neural Networks for Computing", AIP Conference Proceedings # 151, p. 86, (1986).
- [11] T. Maxwell, C. L. Giles, Y. C. Lee, and H. H. Chen, "Nonlinear Dynamics of Artificial Neural Systems", in "Neural Networks for Computing", AIP Conference Proceedings # 151, p. 299, (1986).
- [12] M. B. Reid, L. Spirkovska, and E. Ochoa, "Rapid Training of Higher-Order Neural Networks for Invariant Pattern Recognition", First Int'l Joint Conf. on Neural Networks (IJCNN-89), (1989).
- [13] M. Harwit and N. J. A. Sloane, HADAMARD TRANSFORM OPTICS, ACADEMIC PRESS, NEW YORK, 1979.
- [14] B. Kosko, "Feedback Stability and Unsupervised Learning" IEEE Int'l Conf. on Neural Networks, p. I-141, (1988).
- [15] J. A. Schouten, RICCI CALCULUS, Springer, New York, 1954.
- [16] P. R. Halmos, FINITE-DIMENSIONAL VECTOR SPACES, D. Van Nostrand, New York, 1958.
- [17] J. W. Goodman, A. R. Dias, and L. M. Woody, "Fully Parallel, High-Speed Incoherent Optical Method for Performing Discrete Fourier Transforms", Opt. Lett. 2, 1 (1987).

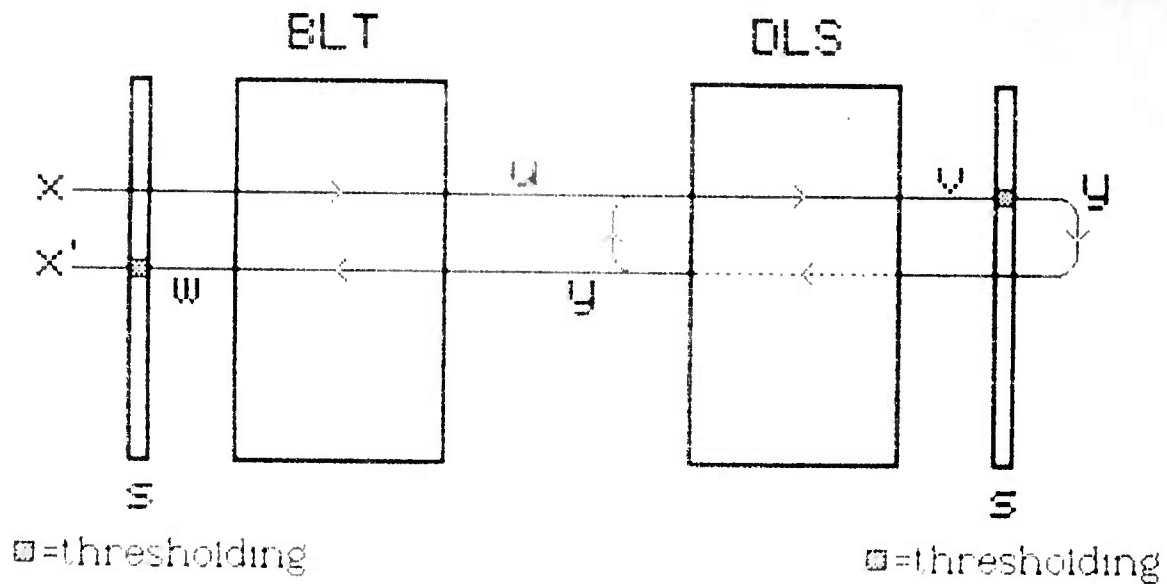
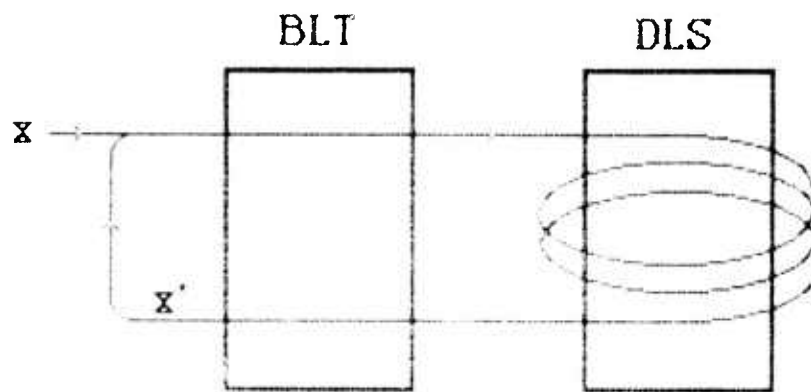
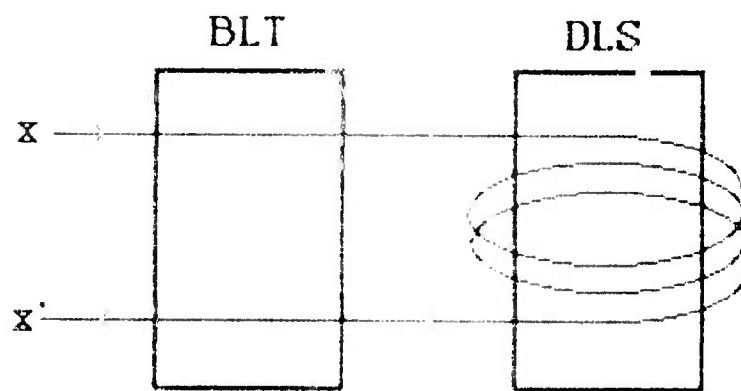


Fig.1 . Selective Reflexive Memory (SRM), consisting of a Bidirectional Linear Transformer (BLT) and a Dominant Label Selector (DLS). The SRM takes a bipolar input $\mathbf{x}$ to a bipolar output $\mathbf{x}'$. The BLT processes data in a forward stroke: $\mathbf{x} \rightarrow \mathbf{u}$, and a backstroke: $\mathbf{y} \rightarrow \mathbf{w}$. The DLS computes from $\mathbf{u}$ a quadratic activation $\mathbf{v}$, which is thresholded to the bipolar vector $\mathbf{y}$. $\mathbf{y}$ is fed back to the DLS input. Also, $\mathbf{y}$ is processed by the BLT in the backstroke: $\mathbf{y} \rightarrow \mathbf{w}$, and $\mathbf{w}$ is thresholded to give the bipolar output $\mathbf{x}'$.



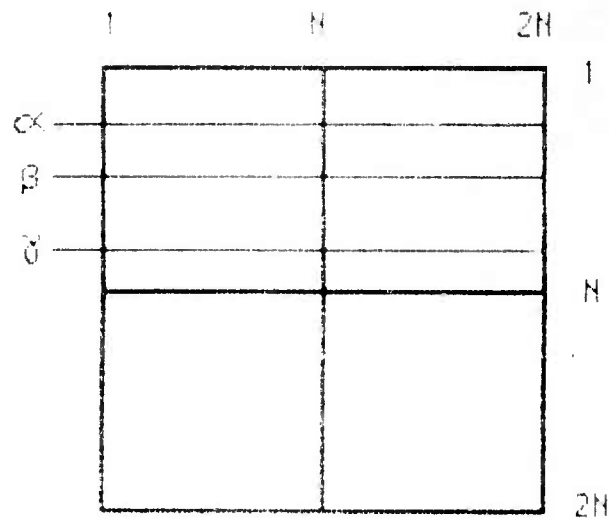
(a)



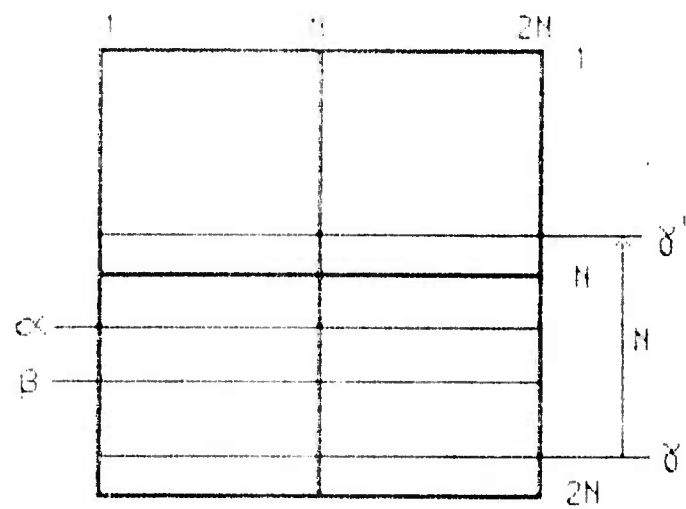
(b)

Fig. 2 . Two modes of operation of the SRM. In (a) x' is used to upgrade the input x . In (b) x' is used as the output of the machine.

Case 1)



Case 2)



Case 3)

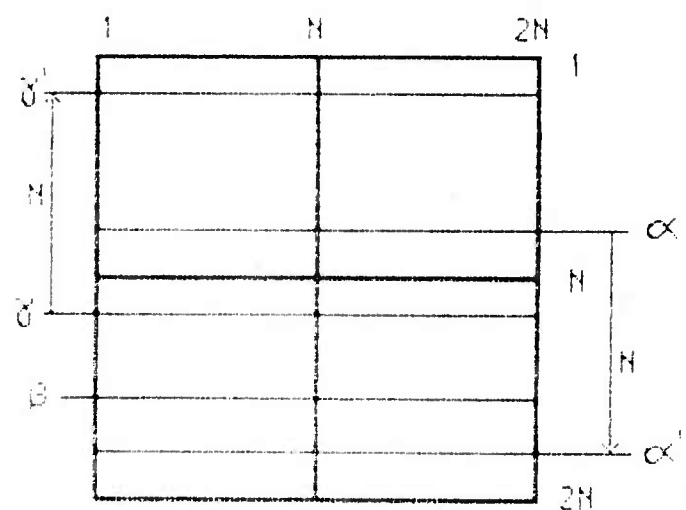


Fig. 3. Cases 1), 2), and 3) used in the proof of the group property of Hadamard vectors in the Appendix.