Hierarchical model of memory and memory loss

1988 J. Phys. A: Math. Gen. 21 4443

(http://iopscience.iop.org/0305-4470/21/23/024)

J. Phys. A: Math. Gen. 21 (1988) 4443-4454. Printed in the UK

Hierarchical model of memory and memory loss

J P Sutton, J S Beis and L E H Trainor
Department of Physics, University of Toronto, Toronto, Canada, M5S 1A7

Received 30 November 1987, in final form 29 April 1988

Abstract. A neural circuit model is presented that consists of a hierarchy of nested clusters of neuron-like elements. Each hierarchical level of cluster organisation encodes a different level of distributed memory. Associative memory properties within and between levels are investigated numerically in simple versions of the model. The results may provide insight into mechanisms of memory storage, recall and loss in real neural circuits, such as those found in the cerebral cortex.

1. Introduction

Several recent neural circuit models of associative memory rely on the proposition that every neural element is effectively connected with every other element (Little 1974, Hopfield 1982, Amit *et al* 1985). When the connections are disordered and symmetric, as in the case of spin-glass systems, the models are characterised by complicated energy surfaces having dominant valleys, whose local minima can be identified with memory states. In the limit of infinitely large networks, the distribution of overlaps among these valleys has an ultrametric structure (Mézard *et al* 1984, Rammal *et al* 1986). This interesting feature has been utilised to construct hierarchical models of memory capable of categorising information (Dotsenko 1985, Ioffe and Feĭgel'man 1986, Parga and Virasoro 1986, Toulouse *et al* 1986, Cortes *et al* 1987, Gutfreund 1988).

In contrast to models that possess fully interconnected geometries among similar neural elements, biological circuits, containing of the order of 1000 or more neurons, demonstrate considerable heterogeneity in the distribution of synaptic connections formed by individual neurons (Szentágothai 1977). Here it is not the case that every neuron is connected to every other neuron. The actual synaptic distributions depend, in part, upon the size and shape of the neuron types involved. While some types of neurons are confined to distances of less than 100 μm, other types project over distances of several centimetres or more. Furthermore, the number of connections formed by different neuron types can be quite variable.

Despite enormous diversity in the connection patterns associated with individual neurons, many neural circuits can be subdivided into essentially similar subcircuits, where each subcircuit contains many types of neurons. Even the cerebral cortex, which has long been considered the most complex and elusive of neural circuits, has basic design principles which give regularity to its subcircuits (Mountcastle 1978). Cortical subcircuits are arranged in a nested fashion, with clusters of subcircuits at the first level coalescing together to form subcircuits at the second level, which cluster to form third-level subcircuits, and so on. This nesting arrangement serves to link different and often widely separated regions of the cortex in a precise but distributed manner.

Subcircuits at the first level are organised into parallel columns, each containing of the order of 100 to 1000 neurons. Several physiological responses, such as those occurring in columns in the visual cortex elicited by various optical stimuli, may be associated with each first-level subcircuit. The boundary relations, formed by the clustering of first- and higher-level subcircuits, are associated with successive levels of complex cortical behaviour that integrate and modify responses at lower levels. While this viewpoint oversimplifies the neurobiology of the cerebral cortex, it captures an important design strategy that has been largely overlooked in mathematical studies of neural circuits.

In this paper, a hierarchical model of neural circuits is proposed based on the principle that regions of the cortex, and possibly other neural circuits as well, are topologically organised into nested distributed subcircuits. At the first level of organisation, model clusters are composed of neural elements that are completely interconnected in an apparently disordered manner. Each of these first-level clusters, however, has content-addressable memory capabilities similar to those described in the model by Hopfield (1982). A memory state in a first-level cluster is taken to represent a single physiological response. The number of memory states in each cluster need not be large since only a limited number of responses is assumed to be associated with each first-level subcircuit in real systems. Emphasis in the model is therefore placed on the hierarchical nesting of the clusters and not on the optimal enhancement of memory storage capacity.

First-level clusters in the model are linked together by subsets of elements, termed 'projection elements', to form second-level clusters. Other subsets of projection elements join second-level clusters to produce third-level clusters. Continuing in this manner, an r-level hierarchy of nested clusters of neural elements is constructed (figure 1). Although the subsets of projection elements in each first-level cluster could, in

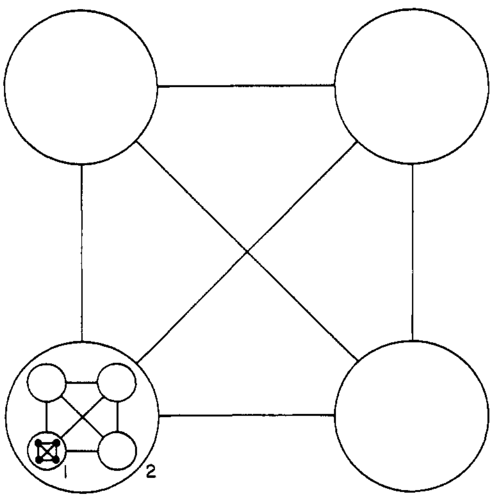

Figure 1. Proposed hierarchy of neural clusters demonstrating three levels of organisation. Clusters at the first level contain neural elements (small dots) which are fully interconnected with each other, as shown in cluster 1. Certain linkages between first-level clusters form second-level clusters, such as cluster 2. Second-level clusters coalesce to create a cluster at the third level. Lines joining different first- and second-level clusters represent connections formed by many, but not all, elements in different first-level clusters.

principle, be overlapping, it is assumed for modelling purposes that the subsets are mutually exclusive. In this way, $r-1$ populations of projection elements are identified within each first-level cluster, such that each population establishes connections with elements in its own cluster and with elements in other clusters nested together at a unique level. This will be made clear in the mathematical formulation of the model. The first level at which two clusters are linked is a measure of what will be termed the 'functional distance' separating them. This measure reflects the anatomical distance between two first-level subcircuits in real circuits, as well as any differences in the physiological responses associated with the subcircuits.

As mentioned above, each first-level cluster has several first-level memory states. Second- and higher-level memory states consist of sets of correlated first-level memory states, the correlations being determined by the details of the intercluster connections. The memory is hierarchical in the sense that lower-level memory states combine to create emergent and possibly degenerate higher-level memory states; higher-level states, in turn, categorise lower-level states. This hierarchical organisation of memory states corresponds to a multi-level ordering of physiological responses in real circuits.

The hierarchical memory associated with the nesting of neural clusters arises naturally from design principles found in finite but complex neural circuits. The memory is to be distinguished from that based on the ultrametric properties of some infinite and fully interconnected models because the proposed layering of memory states arises from heterogeneity in the patterns of connections formed by different populations of elements. This feature breaks the symmetry associated with the complete connectivity of elements found in the ultrametric models. Moreover, there is no need in the present model to formulate abstract measures of memory state correlations that are necessary in the ultrametric models. These measures include the use of graded phase space overlaps among memory states (Ioffe and Feĭgel'man 1986, Parga and Virasoro 1986, Toulouse *et al* 1986, Cortes *et al* 1987, Gutfreund 1988) and the introduction of magnetisation-like quantities (Dotsenko 1985), whose justification and relevance to real neural circuits remains unclear.

A hierarchical spin model of memory that does not explicitly rely on fully interconnected circuitry has been proposed by Dotsenko (1986). The model consists of a layered sequence of spin clusters, where the elements in each cluster are symmetrically connected with one another. Connections between the clusters converge multiple elements at one level to a single element at the next level. Memory states are hierarchically arranged, based partially on the degree of coarse graining performed by individual elements at successive levels. This contrasts with the proposed model, wherein the activity of an individual element cannot depict the collective activity of many other elements. Instead, each element is taken to represent a single neuron-like entity and the graded projections of otherwise similar elements are utilised to construct a hierarchical circuit model. The result is that nested clusters and their associated memory properties are organised in a complex and distributed manner at each level of the hierarchy. This notion includes but transcends the traditional sequential arrangement of clusters utilised in neural circuit models of memory, including the Dotsenko model, to describe successive levels of memory storage and recall.

The following outline is adopted with respect to the proposed model. In § 2, a hierarchical notation is introduced and the connections within and between levels are described. The model is formulated in a manner similar to that of the Hopfield model, but the emphasis is placed on the construction of nested neural clusters. This is achieved by means of the spatial dilution of connections from a fully interconnected circuit and by the utilisation of graded phase lags. To allow the model generality, the hierarchical indices are maintained throughout the formulation. Following a discussion of the general model, the results of computer simulations performed on simple versions of the model are given in § 3. The dependencies of hierarchical memory recall on structural variables, such as the fractions of various projection elements and their relative connection strengths, are delineated. In § 4, the model is examined numerically using percolation techniques to develop a simple multilevel model of neurodegenerative processes. The results of the various studies are placed into perspective in § 5.

2. Hierarchical cluster model of memory

In the proposed scheme, nested clusters of neural elements are labelled at the kth level by sequences of indices $i_k \ldots i_r$, $k=1,\ldots,r$, where i_k denotes a particular cluster at the kth level, i_{k+1} at the next higher level, and so on, to the highest level, where $i_r \equiv 1$ denotes the unique cluster at that level (figure 2). For convenience, the elements themselves are considered to be zeroth-level clusters, with one element per cluster. A cluster $i_1 \ldots i_r$ at the first level contains $N_{i_1 \ldots i_r}$ elements, while a kth-level cluster $i_k \ldots i_r$ contains $N_{i_k \ldots i_r}$ clusters at the $(k-1)$th level, $2 \leqslant k \leqslant r$. The full circuit is composed of N_{i_r} clusters at the $(r-1)$th level, where the i_r label is explicitly maintained even though $i_r \equiv 1$.

The essentially binary state firing property of a neuron is represented by an Ising spin $S_{i_0 \ldots i_r}$, where $S_{i_0 \ldots i_r} = +1(-1)$ characterises the firing (resting) state of the $(i_0 \ldots i_r)$th element. Patterns of memory embedded in each of the first-level clusters are $p_{i_1 \ldots i_r}$ states of the system $\{\xi^{\mu}_{i_0 \ldots i_r}\}^{N_{i_1 \ldots i_r}}_{i_0=1}$, $\mu = 1, \ldots, p_{i_1 \ldots i_r}$. These patterns are random in so far as the $\xi^{\mu}_{i_0 \ldots i_r}$ assume the values $+1$ and -1 with equal probabilities. For convenience below, a memory state $\{\xi^{\mu}_{i_0 \ldots i_r}\}^{N_{i_1 \ldots i_r}}_{i_0=1}$ will also be represented below by the symbol $\chi^{\mu}_{i_1 \ldots i_r}$.

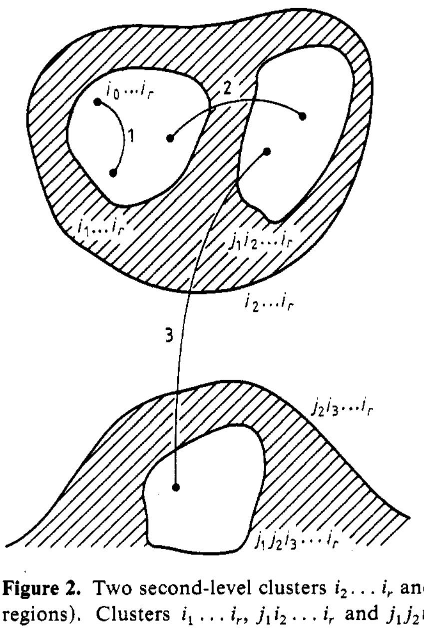

Figure 2. Two second-level clusters $i_2 \ldots i_r$ and $j_2 i_3 \ldots i_r$ are shown schematically (shaded regions). Clusters $i_1 \ldots i_r$, $j_1 i_2 \ldots i_r$ and $j_1 j_2 i_3 \ldots i_r$ are first-level clusters nested within these second-level clusters, and $i_0 \ldots i_r$ labels a single neural element. The number labelling a representative connection between two elements denotes the level of that connection.

The pairwise connections between all the elements within a first-level cluster are, as in the case of the Hopfield model (Hopfield 1982), given by the symmetric Hebb rule:

$$J_{i_0 i_1 \ldots i_r, j_0 i_1 \ldots i_r} = \frac{1}{N_{i_1 \ldots i_r}} \sum_{\mu=1}^{p_{i_1 \ldots i_r}} \xi^{\mu}_{i_0 i_1 \ldots i_r} \xi^{\mu}_{j_0 i_1 \ldots i_r} \qquad i_0 \neq j_0. \tag{1}$$

The connections in equation (1) are referred to as first-level connections. Higher-level connections link elements in different first-level clusters and are, in general, asymmetric. To specify the nature of these connections, consider a single first-level cluster $i_1 \ldots i_r$. Its elements can be divided into two types, namely those elements that form only first-level connections and those elements that, in addition to having first-level connections with every other element in the cluster, also form connections with elements in other clusters. Elements of the latter type are referred to as projection elements. They can be subdivided into $r-1$ disjoint subsets, where elements in each subset form connections with all the elements in other clusters separated from the $(i_1 \ldots i_r)$th cluster by a unique functional distance. The number of projection elements, expressed as a fraction of $N_{i_1 \ldots i_r}$, in the subset associated with the functional distance at the kth level is given by $x^{(k)}_{i_1 \ldots i_r}$, $k = 2, \ldots, r$.

To illustrate the proposed patterns of higher-level connections, consider two first-level clusters $i_1 \ldots i_r$ and $j_1 \ldots j_{k-1} i_k \ldots i_r$ separated by a functional distance at the kth level, $k > 1$. Further suppose that elements $i_0 \ldots i_r$ and $j_0 \ldots j_{k-1} i_k \ldots i_r$ are the only projection elements in their respective first-level cluster subsets to partake in kth-level interactions between the two clusters. That is, $x^{(k)}_{i_1 \ldots i_r} = 1/N_{i_1 \ldots i_r}$ and $x^{(k)}_{j_1 \ldots j_{k-1} i_k \ldots i_r} = 1/N_{j_1 \ldots j_{k-1} i_k \ldots i_r}$. Between the clusters, element $i_0 \ldots i_r$ forms connections with all the $\{j_0 \ldots j_{k-1} i_k \ldots i_r\}_{j_0=1}^{N_{j_1 \ldots j_{k-1} i_k \ldots i_r}}$, but only element $j_0 \ldots j_{k-1} i_k \ldots i_r$ has a possibly non-zero connection with $i_0 \ldots i_r$ at the kth level. Similarly, element $j_0 \ldots j_{k-1} i_k \ldots i_r$ forms connections with elements $\{i_0 i_1 \ldots i_r\}_{i_0=1}^{N_{i_1 \ldots i_r}}$ but not *vice versa*, with the exception of $i_0 \ldots i_r$.

Connections formed by projection elements serve to correlate memory patterns occurring in different first-level clusters. At the kth level, $1 < k \leq r$, the connection from $j_0 \ldots j_{k-1} i_k \ldots i_r$ to $i_0 \ldots i_r$ is written as

$$J_{i_0 \ldots i_r, j_0 \ldots j_{k-1} i_k \ldots i_r} = \frac{\gamma_k}{(N_{i_1 \ldots i_r} N_{j_1 \ldots j_{k-1} i_k \ldots i_r})^{1/2}} \sum_{\langle \mu \mu' \rangle} \xi^{\mu}_{i_0 \ldots i_r} \xi^{\mu'}_{j_0 \ldots j_{k-1} i_k \ldots i_r} \qquad i_{k-1} \neq j_{k-1} \tag{2}$$

where the sum is taken over those pairs of first-level memory states, labelled by μ and μ', which are correlated by the connection; γ_k gives the average strength of the connections at the kth level relative to that of first-level connections, where $\gamma_1 \equiv 1$. Only those J values where $j_0 \ldots j_{k-1} i_k \ldots i_r$ is a projection element at the kth level will be potentially non-zero. Projection elements that form second-level connections, for example, link collections of first-level clusters together in such a manner as to correlate, via non-zero J values at the second level, certain combinations of first-level memory states. These combinations form second-level memories. While some first-level memories may contribute to several second-level memories, thereby introducing degeneracy into the higher levels, other first-level memories may not make any contribution to second-level memory formation. The actual strength of the correlations between the first-level memories at the second level depends upon the number of second-level projection elements, expressed by the fraction $x^{(2)}_{i_1 \ldots i_r}$, and on the scaling factor γ_2.

In general, kth-level connections correlate memory patterns that are present in different $(k-1)$th-level clusters, forming kth-level memories. Each kth-level memory $\chi^{\mu''}_{i_k \ldots i_r}$, $\mu''=1, \ldots, p_{i_k \ldots i_r}$, $k=2, \ldots, r$, can be expressed as an ordered array

$$\chi^{\mu''}_{i_k \ldots i_r} = (\chi^{\mu}_{1 i_k \ldots i_r}, \ldots, \chi^{\mu'}_{N_{i_k} i_k \ldots i_r}) \tag{3}$$

of $(k-1)$th-level memories. This formulation endows the circuit with a distributed memory that is hierarchically organised into r levels.

To investigate the effects that multiple levels of connections have on associative memory recall, the network is evolved sequentially by discrete time steps. At each step, an element chosen at random assesses and modifies its alignment as required by its local field according to

$$S_{i_0 \ldots i_r}(t+1) = \mathrm{sgn}[h_{i_0 \ldots i_r}(t)]. \tag{4}$$

Here, the local field is

$$h_{i_0 \ldots i_r}(t) = \sum_{k=1}^{r} h^{(k)}_{i_0 \ldots i_r}(t) + I_{i_0 \ldots i_r}(t) - \theta_{i_0 \ldots i_r} \tag{5}$$

where the first term in the sum gives the instantaneous local field contribution due to interactions within a first-level cluster, namely

$$h^{(1)}_{i_0 \ldots i_r}(t) = \sum_{j_0=1}^{N_{i_1 \ldots i_r}} J_{i_0 i_1 \ldots i_r, j_0 i_1 \ldots i_r} S_{j_0 i_1 \ldots i_r}(t) \qquad i_0 \neq j_0. \tag{6}$$

For $1 < k \leqslant r$

$$h^{(k)}_{i_0 \ldots i_r}(t) = \sum_{j_{k-1}=1}^{N_{i_k \ldots i_r}} \cdots \sum_{j_0=1}^{N_{j_1 \ldots j_{k-1} i_k \ldots i_r}} J_{i_0 \ldots i_r, j_0 \ldots j_{k-1} i_k \ldots i_r} S_{j_0 \ldots j_{k-1} i_k \ldots i_k}(t-\tau_k) \qquad i_{k-1} \neq j_{k-1} \tag{7}$$

where τ_k gives the integer-valued phase lag due to propagation delays between elements separated by the functional distance at the kth level relative to those separated by the functional distance at the first level (where $\tau_1 = 0$). It is expected that, in general, $\tau_{k+1} > \tau_k$, $1 \leqslant k < r$, in accord with the notion that functional distance incorporates a measure of neuro-anatomical distance in real circuits. Finally, $I_{i_0 \ldots i_r}(t)$ and $\theta_{i_0 \ldots i_r}$ in equation (5) represent the time-dependent external field and fixed threshold potential of the $(i_0 \ldots i_r)$th element, respectively, and are set equal to zero in the discussion that follows.

Despite the fact that elements are chosen sequentially in the evolution process given by equation (4), each step involves a field assessment of the entire network. State changes that result from this assessment tend to decrease a global function $E = \sum_{k=1}^{r} E^{(k)}$, where

$$\begin{aligned} E^{(k)}(t) = -\tfrac{1}{2} \sum_{i_{r-1}=1}^{N_{i_r}} \cdots \sum_{i_{k-1}=1}^{N_{i_k \ldots i_r}} \sum_{\substack{j_{k-1}=1 \\ j_{k-1} \neq i_{k-1}}}^{N_{j_k i_{k+1} \ldots i_r}} \cdots \\ \cdots \sum_{i_0=1}^{N_{i_1 \ldots i_r}} \sum_{j_0=1}^{N_{j_1 \ldots j_{r-1} i_r}} J_{i_0 \ldots i_r, j_0 \ldots j_{k-1} i_k \ldots i_r} S_{i_0 \ldots i_r}(t) S_{j_0 \ldots j_{k-1} i_k \ldots i_r}(t-\tau_k) \\ = -\tfrac{1}{2} \sum_{i_{r-1}=1}^{N_{i_r}} \cdots \sum_{i_0=1}^{N_{i_1 \ldots i_r}} h^{(k)}_{i_0 \ldots i_r}(t) S_{i_0 \ldots i_r}(t) \end{aligned} \tag{8}$$

corresponds to the kth-level contribution, $1 \leq k \leq r$. Beginning with some initial suggestive configuration of element states $S_{i_0 \ldots i_r}$, memory recall occurs by the mechanism given in equation (4) until a local E minimum associated with a complete set of corresponding first-level memory states (one from each cluster) is reached. This property is basically assured in the $r = 1$ case when p_1/N_1 is less than approximately 0.14 (Hopfield 1982, Amit *et al* 1985). The memory states in each disjoint cluster are then, on average, orthogonal, and behave as attractors within a partitioned configuration space. The addition of higher-level connections alters the topology of the E surface within this space. In contrast to first-level connections, the asymmetry of higher-level connections causes E to lose its energy-like properties that are present when $r = 1$. If it is assumed that $p_{i_1 \ldots i_r}/N_{i_1 \ldots i_r}$ is small, so that memory states at the first level are indeed potentially stable states, then $(k > 1)$-level connections can modify the attractor basins of the first-level memory states to produce intercluster memory correlations. In doing so, however, the stability of some first-level memory states may be lost.

From equation (7), it is evident that the summations performed in determining a local higher-level field of the $(i_0 \ldots i_r)$th element can, in principle, include many more contributions from elements located outside and successively further from the $(i_1 \ldots i_r)$th cluster than from those elements residing within the cluster. In real circuits, such as the cortex, there is roughly an inverse relationship between the number of connections involving small populations of neurons and the proposed functional distance measure over which the neurons have synaptic connections. It is therefore assumed that the fractions of projection elements are graded such that $x^{(k+1)}_{i_1 \ldots i_r} < x^{(k)}_{i_1 \ldots i_r}$, $k = 2, \ldots, r-1$. Moreover, $x^{(2)}_{i_1 \ldots i_r}$ is assumed to be less than the fraction of elements in the $(i_1 \ldots i_r)$th cluster that are not projection elements.

In addition to the constraints imposed on the patterns and the number of projection elements, it is assumed that the connection strengths at all the different levels are roughly of the same order. If this were not the case, an unusually small or large value of γ_k will distort the configuration space so as to effectively disconnect the circuit at the kth-level cluster or destroy memory correlations occurring at all but the kth level. To the authors' knowledge, there is no experimental evidence to suggest that some γ_k values are vastly different from other values or that the values are sequentially ordered. Furthermore, with respect to the graded phase lags, it is assumed that the τ_k are small, so that the influence of all levels during simulated memory recall is actually present. In this way, the entire circuit operates in a cooperative and parallel manner, and memory association at higher levels need not occur only after lower-level memory recall has been completed.

3. Numerical simulations

Several numerical simulations have been performed on two- and three-level versions of the model to quantitatively assess the effect of hierarchical organisation on memory recall. Only results pertaining to zero phase lag studies are presented here; non-zero phase lag effects are currently under investigation. For $r = 2$, $N_1 = 3$, $N_{i_1 1} = 30$, $x^{(2)}_{i_1 1} = x_2$, $p_1 = p_{i_1 1} = 3$, where $i_1 = 1, 2, 3$, and $\tau_2 = 0$, correlations among the first-level memory states in different clusters were studied as functions of the fraction of projection elements x_2 and the relative connection strength γ_2. As expected, with $p_{i_1 1}/N_{i_1 1} = 0.10$, all nine first-level memory states in this two-level model were found to be stable when second-level connections were ignored. Second-level connections, however, gave rise to second-level memories. Each of the three second-level memories consisted of a different triplet of first-level memory states, each such state being chosen at random from a different first-level cluster. Thus, three of 27 possible combinations of first-level memory states were pre-designated as second-level memories. Simulations were conducted by randomly cueing one of the clusters in a first-level memory state associated with one third of a second-level memory. Then, with the other two clusters in random initial states, the system was allowed to evolve according to equations (4)-(8) until a stable state was achieved. If the final state was one of the three second-level memory states, then a correct correlation of first-level memories was recorded; otherwise, it was not. An averaging was performed over initial states, cued memory states, and sets of first-level memory states.

The dependence of second-level memory recall, expressed as a percentage correlation of the pre-selected first-level memories, on x_2 and γ_2 under instantaneous field conditions is shown in figure 3. The range of correct correlations as a function of x_2 for selected values of $0.5 \leqslant \gamma_2 \leqslant 3.0$ is 7.3 to 94% (figure 3(a)). Most of the gain in second-level memory association occurs for $x_2 < 0.4$ when $\gamma_2 \geqslant 1.0$. This is not surprising in a two-level hierarchy containing similar neural clusters, since the number of terms in the field calculations scales as $N_{i_1 1}$ for $h^{(1)}_{i_0 i_1 1}$, whereas for $h^{(2)}_{i_0 i_1 1}$, the number of terms effectively increases by $x_2 N_1 N_{i_1 1}$. For values of $\gamma_2 < 0.5$, the recall ability of second-level memories deteriorates dramatically.

Note that as $x_2 \to 0$, the circuit becomes dissociated into disjoint clusters, and intercluster correlations of first-level memory states can no longer be stored. The *a priori* probability that relaxation to first-level memory states in each cluster would correspond to a second-level memory is $p_1/p_{i_1 1}^{N_1} = 0.11$. In practice, the correlation value was slightly less than this (0.07), since the evolution of the two non-cued clusters may result in spurious local E minima states that are not designated first-level memory states. This effect partially explains the less than perfect correlation of first-level memories seen as $x_2 \to 1$. In this limit, which also corresponds to the binary interaction case of an associative memory model proposed by Chen *et al* (1986), the relatively large number of second-level connections occasionally causes the cued memory state to become unstable. When random initial states of the non-cued clusters are associated with a deep valley on the E surface that excludes the cued memory state, the model tends to come to rest in either some combination of first-level memory states not representing a second-level memory or some other spurious state.

Second-level memory recall is optimised in a range of γ_2 values that varies with x_2 (figure 3(b)). When $\gamma_2 < 0.2$, the second-level effects of the model for $x_2 \leqslant 0.4$ are essentially negligible. On the other hand, for $\gamma_2 > 7$, the second-level correlations tend to enslave the first-level events, causing the model to lose the fine detail of the E surface and the efficiency of its memory storage and recall capabilities. Relatively large variability (roughly 15%) in second-level recall occurs in this range. A detailed theoretical account of why peak second-level associations occur for $2 < \gamma_2 < 5$ and $0.1 \leqslant x_2 \leqslant 0.4$ remains to be elicited.

The γ_2–γ_3 surface associated with third-level memory recall is shown in figure 4 in the case of $r = 3$, $N_1 = 2$, $N_{i_2 1} = 3$, $N_{i_1 i_2 1} = 30$, $x^{(2)}_{i_1 i_2 1} = x_2 = 0.27$, $x^{(3)}_{i_1 i_2 1} = x_3 = 0.13$, $p_1 = p_{i_2 1} = p_{i_1 i_2 1} = 3$, $i_1 = 1, 2, 3$, $i_2 = 1, 2$, and $\tau_2 = \tau_3 = 0$. In accordance with the development of the previous sections, the subsets of projection elements within a first-level cluster, that project over second- and third-level functional distances, were taken to be mutually exclusive. High third-level memory correlations in this instantaneous field simulation occur for values of $0.5 < \gamma_2 < 2$ and $2 < \gamma_3 < 5$. The actual shape of the γ_2–γ_3 surface

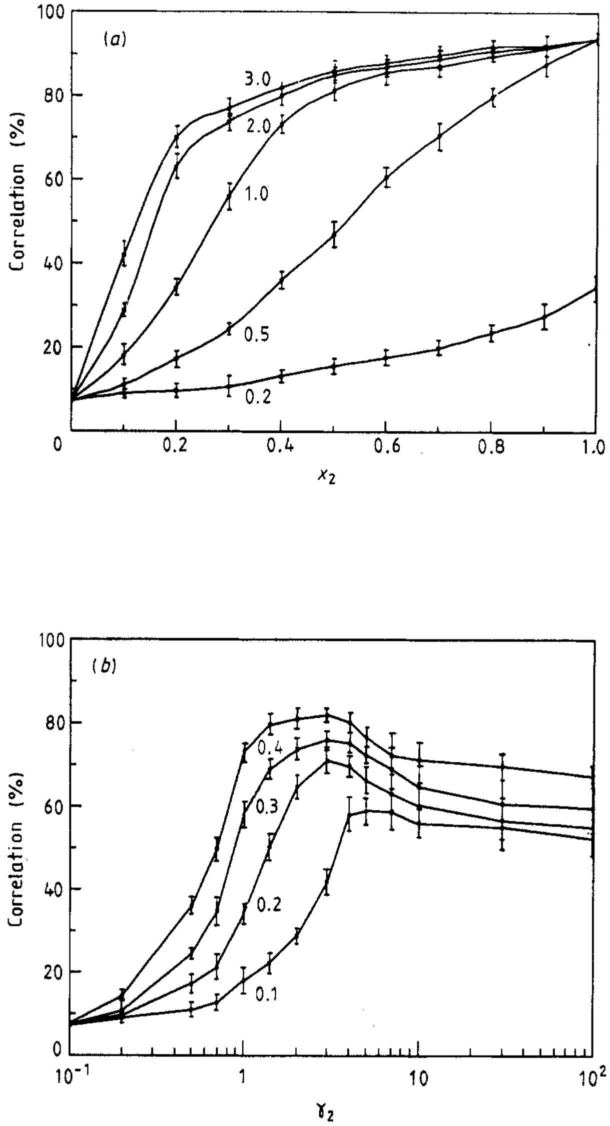

Figure 3. Results of second-level memory recall expressed as a percentage correct correlation among pre-designated first-level memories for $r=2$, $N_1=3$, $N_{i_1 1}=30$, $p_1=p_{i_1 1}=3$, $\tau_2=0$. (a) Percentage correlation dependence on x_2 for $\gamma_2=0.2$, 0.5, 1.0, 2.0, 3.0; (b) percentage correlation dependence on γ_2 for $x_2=0.1$, 0.2, 0.3, 0.4.

is largely a function of x_2 and x_3, as expected by the strong inter-dependence of x_2 and γ_2 observed in the two-level model.

4. A model of memory loss

To this point, the model has been discussed in the context of hierarchical storage and associative recall of information. It might also be useful in understanding mechanisms

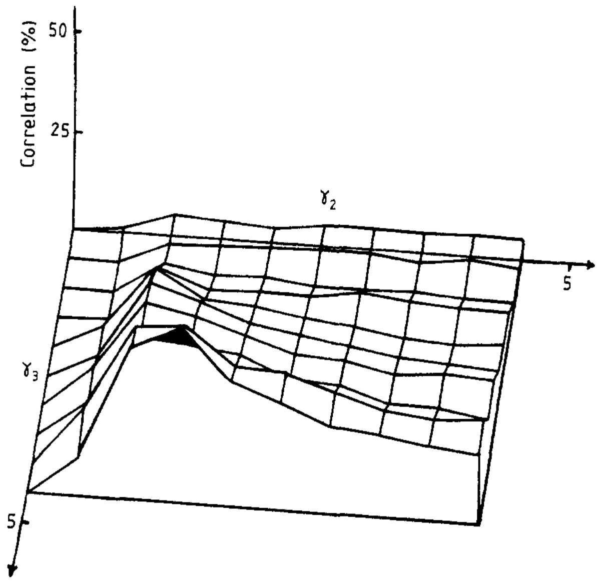

Figure 4. Third-level memory recall as a function of γ_2 and γ_3 with $r=3$, $N_1=2$, $N_{i_21}=3$, $N_{i_1i_21}=30$, $x_2=0.27$, $x_3=0.13$, $p_1=p_{i_21}=p_{i_1i_21}=3$, and $\tau_2=\tau_3=0$.

of memory loss in some common neurodegenerative disorders in real circuits. Alzheimer's disease, for example, primarily involves a selective destruction of large neurons, which comprise approximately 20% of the total cortical neuron population (Katzman 1986). These neurons project over considerable anatomical distances and are either partially or completely destroyed in the disease process. Small localised neurons remain relatively unaffected. It is estimated that in total, only 10% of all cortical neurons are actually lost in Alzheimer's disease (Katzman 1986). However, the loss to higher brain behaviours is progressive and often devastating.

The proposed model can be modified to simulate selective neuron loss and its associated effects on higher memory operations. The idea is that the selective loss of projection elements in the model destroys memory correlations between clusters, while preserving, in a relative sense, memory correlations within clusters. The loss of projection elements may be partial or complete. Partial destruction can be achieved by decreasing the numbers and relative strengths of the corresponding non-zero connections. Complete destruction can, if done randomly, correspond to a site percolation problem.

As a simple example of how the loss of projection elements affects memory correlations, a selective site percolation on the projection elements of the two-level model in § 3 was performed. A 50% selective loss of projection elements that reduced x_2 from 0.2 to 0.1 and an additional reduction in γ_2 on the remaining intercluster connections from 3.0 to 1.0 was found to decrease second-level memory recall from greater than 70% to approximately 10% (figure 5). The change in x_2 and γ_2 effectively dissociated the first-level clusters with respect to intercluster correlations; first-level memory properties, however, were essentially unaltered in the process. It is a straightforward matter to generalise these results to multiple-level versions of the model. It should be noted that the effects due to selective element loss are to be distinguished from percolation phenomena occurring in collective neural networks containing similar

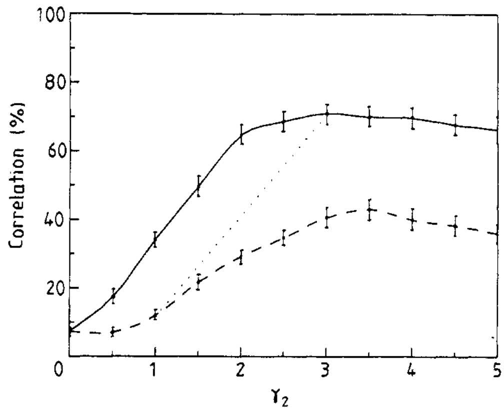

Figure 5. Percentage correlation as a function of γ_2 for $r=2$, $N_1=3$, $N_{i_11}=30$, $p_1=p_{i_11}=3$, $x_2=0.2$, $\tau_2=0$. The curves give the control (full curve) and 50% projection element loss (broken curve) simulation results. A possible loss in second-level memory recall is shown (dotted curve).

elements with homogeneous connection patterns. In such networks, a 10% loss of elements generally has minimal effects on memory loss.

5. Conclusions

While the proposed model possesses features that are suggestive of synaptic and memory organisation in some biological circuits, they are at best crude approximations to the complex architectures actually found in the mammalian brain. Nevertheless, it is argued that basic principles of hierarchical design associated with the heterogeneity of connection patterns within real and synthetic neural circuits are important to an understanding of the organisation of physiological responses, memory and memory loss. These principles can be formulated in terms of variables that have anatomical and physiological relevance and can be rigorously investigated in a manner similar to spin-glass studies. This line of investigation poses interesting theoretical challenges and may also have applications beyond those discussed in this paper.

Acknowledgments

The generous support of MRC and NSERC of Canada is gratefully acknowledged. Simulations were performed on the Cray X-MP/22 at the Ontario Centre for Large-Scale Computation.

References

Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 1530-3
Chen H H, Lee Y C, Sun G Z, Maxwell T and Giles C L 1986 *Neural Networks for Computing* (*AIP Conf. Ser.* **151**) ed J S Denker (New York: American Institute of Physics) pp 86-99
Cortes C, Krogh A and Hertz J A 1987 *J. Phys. A: Math. Gen.* **20** 4449-55

Dotsenko V S 1985 *J. Phys. C: Solid State Phys.* **18** L1017-22
—— 1986 *Physica* **140A** 410-5
Gutfreund H 1988 *Phys. Rev.* A **37** 570-7
Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554-8
Ioffe L B and Feĭgel'man M V 1986 *JETP Lett.* **44** 189-92
Katzman R 1986 *N. Engl. J. Med.* **314** 964-73
Little W A 1974 *Math. Biosci.* **19** 101-20
Mézard M, Parisi G, Sourlas N, Toulouse G and Virasorso M 1984 *Phys. Rev. Lett.* **52** 1156-9
Mountcastle V B 1978 *The Mindful Brain* ed G M Edelman and V B Mountcastle (Cambridge, MA: MIT Press) pp 7-50
Parga N and Virasoro M A 1986 *J. Physique* **47** 1857-64
Rammal R, Toulouse G and Virasoro M A 1986 *Rev. Mod. Phys.* **58** 765-88
Szentágothai J 1977 *Proc. R. Soc.* B **201** 219-48
Toulouse G, Dehaene S and Changeux J-P 1986 *Proc. Natl Acad. Sci. USA* **83** 1695-8