Associative Memory in a Multi-modular Network

Nir Levy and David Horn
School of Physics and Astronomy
Tel Aviv University Tel Aviv 69978, Israel
nirlevy@post.tau.ac.il horn@neuron.tau.ac.il

Eytan Ruppin
Departments of Computer Science & Physiology
Tel Aviv University Tel Aviv 69978, Israel
ruppin@math.tau.ac.il

January 17, 1999

Abstract

Recent imaging studies suggest that object knowledge is stored in the brain as a distributed network of many cortical areas. Motivated by these observations, we study a multi-modular associative memory network, whose functional goal is to store patterns with different coding levels, i.e., patterns that vary in the number of modules in which they are encoded. We show that in order to accomplish this task, synaptic inputs should be segregated into intra-modular projections and inter-modular projections, with the latter undergoing additional nonlinear dendritic processing. This segregation makes sense anatomically if the inter-modular projections represent distal synaptic connections on apical dendrites. It is then straightforward to show that memories encoded in more modules are more resilient to focal afferent damage. Further hierarchical segregation of inter-modular connections on the dendritic tree improves this resilience, allowing memory retrieval from input to just one of the modules in which it is encoded.

1 Introduction

The knowledge of concrete objects, such as their physical features and functional properties, is stored in the brain in a distributed network of discrete cortical areas. Imaging studies [Martin *et al.*, 1995, Martin *et al.*, 1996] have shown that attributes that define an object are represented close to the cortical regions that mediate perception of these attributes. Moreover, the brain regions active during object identification are dependent on intrinsic properties of the object presented.

Cortical modules were observed in the somatosensory [Mountcastle, 1957], visual [Hubel and Wiesel, 1977] and association [Hevner, 1993] cortices. These modules differ in their structure and function but are likely to be elementary units of processing in the mammalian cortex. Within each module the neurons are interconnected. Input and output fibers to and from other cortical modules and subcortical areas connect to these neurons. In some parts of the cortex modules were shown to function as memory units (see [Amit, 1995] for a review).

It therefore behooves us to study neural networks of associative memory with modules representing different modalities or features. Previous studies of modular networks [O'Kane and Treves, 1992, O'Kane and Sherrington, 1993, Viana and Martínez, 1995] were limited to memory patterns which where coded in a fixed number of modules, equal for all memories. The computational challenge of storing memory patterns that are coded in a *variable* number of modules was first introduced by [Lauro-Grotto *et al.*, 1994]. The authors addressed the problem by using the category to which each memory belongs (which determines its level of coding) to modulate synapses and neuronal thresholds during retrieval. This process has led [Lauro-Grotto *et al.*, 1994] to conclude that storage and retrieval of memories that are coded in a variable number of modules must involve consciousness.

In this paper we introduce a low-level mechanism that solves the problem without requiring explicit knowledge of the coding level of each memory. Our solution is based on a functional distinction between the intra-modular and inter-modular couplings. While intra-modular connections are summed up in the conventional linear manner, we introduce a non-linear dendritic function that processes the activations arriving via the inter-modular connections. This non-linear processing operates as a *squashing* function on the inputs coming from other modules, thus eliminating the difference between contributions to the postsynaptic potential generated by memories with different levels of activity.

The biological motivation for intra/inter modular synaptic segregation hinges on the observation that neurons from distant modules synapse onto the distal part of the dendritic tree [Markram *et al.*, 1997, Yuste *et al.*, 1994, Hetherington and Shapiro, 1993]. Furthermore, [Yuste *et al.*, 1994] have suggested that neocortical pyramidal neurons have two functional zones: a basal compartment, comprising the basal dendrites, cell body, and the proximal $300\mu m$ of the apical dendrite; and an apical compartment, formed by the rest of the apical dendrite. Evidence for non-linear processing comes from the observation that neocortical apical dendrites are capable of producing regenerative events and thus may be involved in boosting distal synaptic inputs [Yuste *et al.*, 1994, Magee and Johnston, 1995, Johnston *et al.*, 1996, Yuste and Tank, 1996]. [Stuart *et al.*, 1997] have found that proximal synaptic stimulation (in layer 4) of layer 5 pyramidal neurons initiated action potentials first at the soma, whereas distal stimulation (upper layer 2/3) could initiate dendritic regenerative potential prior to somatic action potentials.

The need to store memories with different coding levels necessarily entails a significant loss of memory capacity in conventional associative memory networks. The concept of critical capacity needs to be modified to accommodate the fact that populations of memories with different coding levels have distinct critical capacities. The introduction of non-linear processing enables the storage of a broad spectrum of such populations and considerably increases the network's information capacity compared with conventional networks. Yet, the resulting critical capacity levels are still significantly smaller than those obtained in networks storing memories with a single, uniform coding level.

In the next Section, we present a model of a multi-modular associative memory, storing memory patterns encoded in a variable number of modules. In Section 3 we motivate the usage of non-linear dendritic processing, characterize the optimal dendritic processing function and study the network's memory capacity. Section 4 investigates the fault tolerance properties of a modular network and its consequences. Finally, the last Section summarizes our results. Details of analytic calculations are provided in Appendices.

2 Multi-Modular Organization

We study an excitatory-inhibitory associative memory network, storing M memory patterns in L modules of N binary neurons each. Each memory η^μ is defined on a subset of size Ω^μ of the L modules. We refer to Ω^μ as its *modular coding*. The sparse coding level inside a module is $p << 1$. The synaptic efficacy ${J_{ij}}^{lk}$ between the jth (presynaptic) neuron from the kth module and the ith (postsynaptic) neuron from the lth module is chosen in a Hebbian manner

$$ {J_{ij}}^{lk} = \frac{1}{Np} \sum_{\mu=1}^{M} {\eta^\mu}_{il} {\eta^\mu}_{jk} \ . \tag{1} $$

This synaptic matrix is a natural extension of [Tsodyks, 1989] formulation for the case of multi-modular network. Although the network is fully connected, in the sense that all synapses may be formed, note that most of the synapses vanish because of the sparse coding hence the effective connectivity is low. The updating rule for the activity state V_i^l of the ith binary neuron in the lth module is given by

$$ {V_i}^l(t+1) = \Theta \left[{h_i}^l(t) - \theta_n \right] \ , \tag{2} $$

where $\Theta[x]$ is the step function and θ_n is the neuronal threshold. The neuron's local field, or membrane potential, has two components

$$h_i{}^l(t) = h_i{}^l{}_{internal}(t) + h_i{}^l{}_{external}(t) \ . \tag{3}$$

The internal field is

$$h_i{}^l{}_{internal}(t) = \sum_{j \neq i}^{N} J_{ij}{}^{ll} V_j{}^l(t) - \gamma_n \mathcal{Q}^l(t) \ , \tag{4}$$

with inhibition proportional to the total activity inside the module

$$\mathcal{Q}^l(t) = \frac{1}{Np} \sum_{j}^{N} V_j{}^l(t) \ . \tag{5}$$

The external field is

$$h_i{}^l{}_{external}(t) = \mathcal{G}\left[\sum_{k \neq l}^{L} \sum_{j}^{N} J_{ij}{}^{lk} V_j{}^k(t) - \gamma_d \sum_{k \neq l}^{L} \mathcal{Q}^k(t) - \theta_d\right] \ , \tag{6}$$

where $\mathcal{G}[x]$ represents dendritic processing of post-synaptic currents from neurons situated in other cortical modules, and allows for the freedom of using more complicated behavior than the standard $\mathcal{G}[x] = x$ one.

Both local and global inhibition are used to keep the means of the internal and external crosstalk noise terms near zero, otherwise the capacity vanishes. Evidence for the existence of local cortical inhibition is abundant but that for global inhibition is weaker. Such global cortical inhibition may be mediated through diffuse thalamo cortical projections that terminate on certain classes of inhibitory interneurons (see [Mountcastle, 1997] for a review).

The retrieval quality at each recall trial is measured by the overlap function that describes the similarity between the final state V the network converges to and the memory pattern η^μ that is cued in each trial. It is defined by

$$m^\mu(t) = \frac{1}{p(1-p)N\Omega^\mu} \sum_{k=1}^{L} \sum_{i=1}^{N} (\eta^\mu{}_{ik} - p)\, V_i{}^k(t) \ , \tag{7}$$

where Ω^μ is the modular coding of η^μ.

3 Dendritic Processing

3.1 Motivation

The dendritic processing function $\mathcal{G}[x]$ separates between the proximal (intra-modular) and distal (inter-modular) synaptic connections. Hence, synaptic inputs from neurons that are situated in other modules undergo additional processing compared with inputs from neurons that belong to the same module. We first compare in simulations the linear (and, so far, standard) case $\mathcal{G}[x] = x$ of dendritic processing with the non-linear $\mathcal{G}[x] = \lambda\Theta[x]$, where Θ is the Heaviside step function.

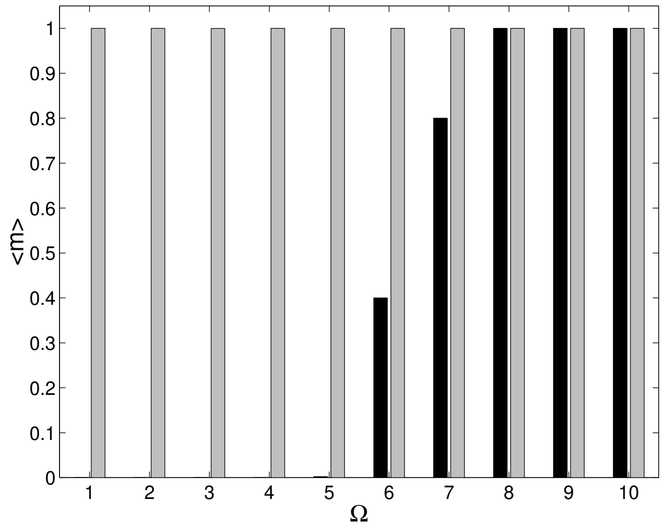

Figure 1: Quality of retrieval *vs.* memory modular coding. The dark shading represents the mean overlap of retrieved memories with different modular codings achieved by a network with linear synaptic couplings. The light shading represents the mean overlap of a network with non-linear processing of the inter-modular connections. The latter achieves perfect recall of all memory patterns. The simulation parameters are: $L = 10$, $N = 500$, $M = 50$, $p = 0.05$, $\lambda = 0.7$, $\theta_d = 2$ and $\theta_n = 0.6$. The encoded memories are distributed homogeneously over all possible modular codings.

Figure 1 shows the performance of the two networks when the stored memories have different levels of modular codings. We measure the network's performance by the mean overlaps of the retrieved memory patterns. We choose the modular coding Ω^μ to be homogeneously distributed with equal proportions for all memories e.g., we store 5 memories with $\Omega^\mu = 1$, another 5 memories with $\Omega^\mu = 2$, and so on. As can be seen the linear network can only sustain memories with high modular codings. In order to retrieve patterns with low Ω^μ, non-linear inter-connections are needed.

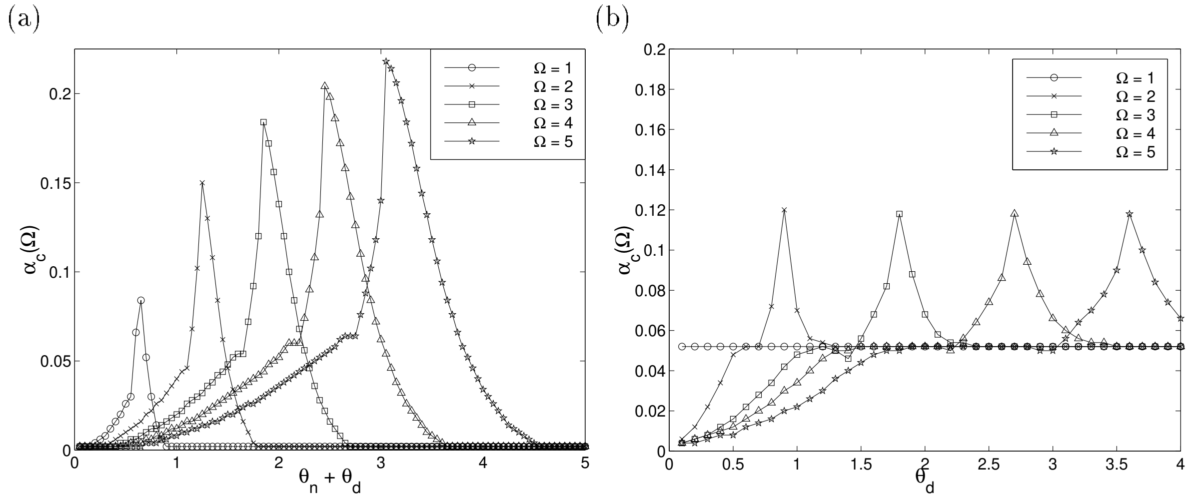

Figure 2: Critical capacities for different modular codings as function of threshold parameters. (a) A modular network with linear $\mathcal{G}[x] = x$. The network has $L = 5$ modules with $N = 1000$ neurons in a module and stores memory patterns with $p = 0.05$. (b) A modular network with non-linear $\mathcal{G}[x] = \lambda\Theta[x]$, using the same architecture with parameters $\theta_n = 0.7$ and $\lambda = 0.5$. Finite $\alpha_c \geq 0.05$ for all Ω is obtained for $\theta_d > 2$.

To understand this behavior we calculate the critical capacities $\alpha_c(\Omega) = \frac{M}{LN}$ of memories with modular coding Ω for the two networks as a function of threshold. We define $\alpha_c(\Omega)$ as the capacity above which memories of patterns with modular coding Ω are unstable. Note that $\alpha_c(\Omega)$ depends on the modular coding distribution of the whole memory population. In this paper, we focus on the case of equal number of memories for each modular coding (the pertaining capacity calculations are given in Appendix A). As can be seen in Figure 2(a), there exists no threshold for which linear modular networks can sustain their full memory repertoire with all modular codings. Linear dendritic processing allows for the coexistence of memories with only a relatively narrow span of different modular coding levels, as already observed by [Lauro-Grotto *et al.*, 1994]. However, in the case of the non-linear network with $\mathcal{G}[x] = \lambda\Theta[x]$ shown in Figure 2(b), a stable system is obtained, in which all possible modular codings can coexist (for values of $\theta_d > 2$).

3.2 Optimal Dendritic Functions

Next, we examine the performance of the modular network when a more general $\mathcal{G}[x]$ is taken as the dendritic processing function. We define the following piecewise linear family of functions

$$\mathcal{G}[x,\beta,\gamma] = \begin{cases} \tan\beta \cdot x & -\infty < x < 0 \\ \tan\gamma \cdot x & 0 < x < \cot\gamma \\ \tan\beta \cdot x + [1 - \cot\gamma \tan\beta] & \cot\gamma < x < \infty \end{cases} , \qquad (8)$$

plotted in Figure 3(a). Given $\mathcal{G}[x,\beta,\gamma]$ we search for the optimal dendritic processing function that will maximize the network's retrieval capabilities. We use the overlap Eq. 7 as a measure of performance and calculate $m^\nu(t=1)$ when pattern η^ν is introduced to the network as an input cue at $t = 0$. This one step calculation is derived in Appendix B for an arbitrary dendritic processing function.

Figure 3(b) shows $m^\nu(t=1)$ for $\mathcal{G}[x,\beta,\gamma]$ in two cases: memories with low modular coding ($\Omega = 10$) and with high modular coding ($\Omega = 40$), calculated for an optimal neuronal threshold θ_n. This analysis shows that choosing small β values (that is, a practically bounded dendritic processing function) is necessary for sustaining memories with low modular codings. These results are independent of γ. However large values of γ (i.e. non-linear dendritic functions such as the step function) are important for allowing strong interactions between different modules, leading to efficient retrieval from partial cues, as discussed in Section 4.

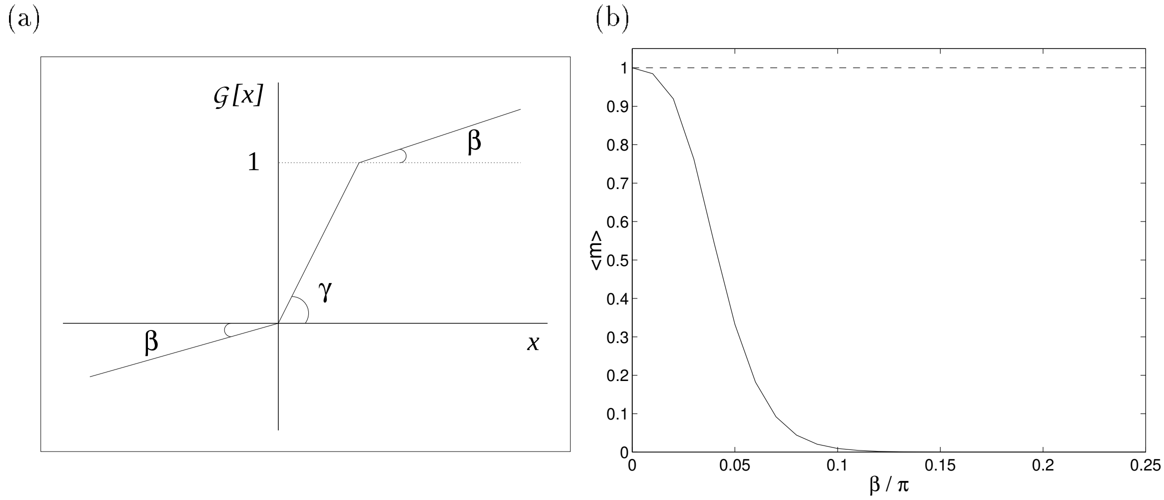

Figure 3: Performance of a modular network for different types of dendritic processing functions. (a) Piecewise linear $\mathcal{G}[x]$ used for the calculation. (b) The overlaps of patterns with $\Omega = 10$ (solid) and $\Omega = 40$ (dashed) for $\mathcal{G}[x]$ defined by Eq. 8 as a function of β. The results for $\frac{\pi}{2} \geq \beta \geq \frac{\pi}{4}$ remain similar to those reported for $\beta = \frac{\pi}{4}$. The modular network has $L = 50$ modules and $N = 1000$ neurons in a module, and stores memory patterns with $p = 0.025$. Other parameters are $\alpha = 0.02$, $\theta_d = 1$ and optimal θ_n. The results are practically independent of the value of γ.

3.3 Memory Capacity in Networks Storing Multi-coded Memories

Storing memories with variable levels of coding necessarily results in a significant reduction in storage capacity. This property is inherent to both modular and conventional, single module networks. In a conventional network the critical capacities of the different memory subpopulations strongly depend on the variance of memory coding levels. The analysis carried out in Appendix C leads to the results displayed in Figure 4, showing the critical capacities of two memory subpopulations η and ξ with coding levels p and f respectively (where $p < f \ll 1$) stored in a single module network.

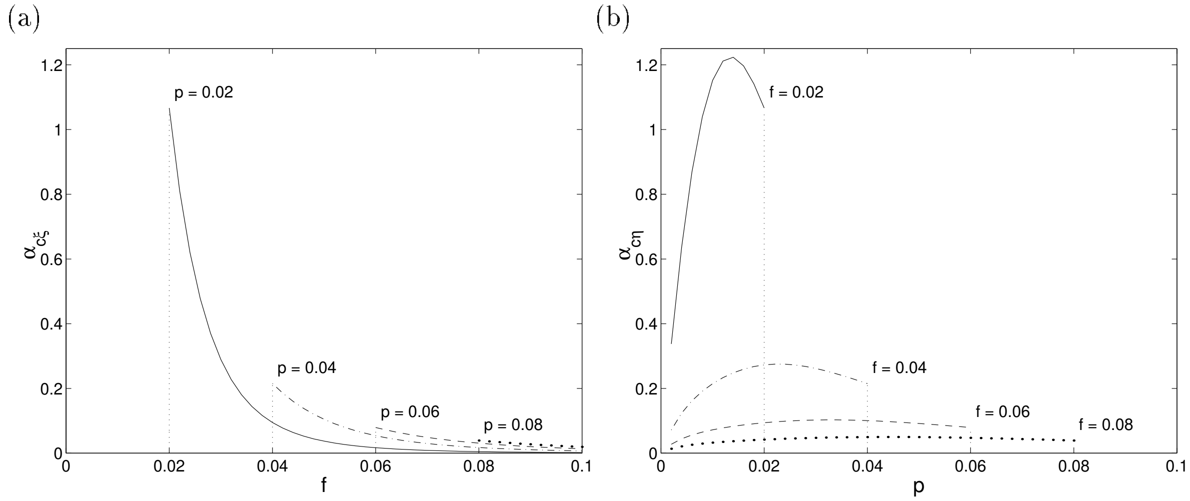

Figure 4: (a) The critical capacity $\alpha_{c\xi}$ *vs.* f, for specific p values, $\theta = 0.7$ and $N = 5000$. (b) $\alpha_{c\eta}$ *vs.* p for specific f values with the same parameters as in (a). The details of the calculation are given in Appendix C.

Figure 4(a) depicts the critical capacity of the high activity memories. As evident, their critical capacity is maximized when $f = p$. The larger the difference between f and p the smaller $\alpha_{c\xi}$ is. Somewhat similar behavior can be seen in Figure 4(b), where the critical capacity of the low activity memories is plotted. In general, the wider the distribution of coding levels the lower are the critical capacities.

The critical capacities of memories stored in a linear modular network are lower than the ones observed in a single module network, but again, memories with higher modular coding have a much lower critical capacity than low-coded memories. An example of this behavior can be seen in Figure 5, comparing a single module network with both linear and non-linear modular networks, all storing the same population of memories. It should be noted that critical capacities depend strongly on both p and N, therefore we have made sure that these values are the same for all the networks that we compare with one another.

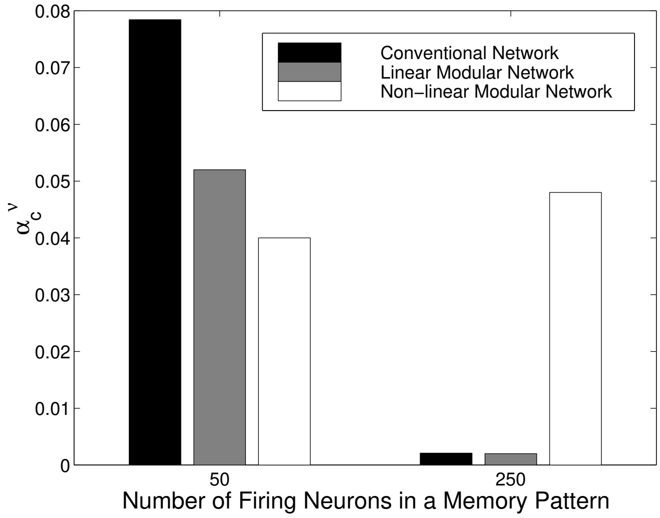

Figure 5: Critical capacities for conventional (single module), linear and non-linear modular networks storing two populations of memories. All three networks have 5000 neurons, and the activities of the two memory populations are 50 and 250 neurons correspondingly (that is, diffused activities of 0.01 and 0.05 in the single module network and two populations of $\Omega = 1, 5$ and $p = 0.05$ in the modular networks). Other parameters are $L = 5$, $\theta = \theta_n = 0.7$, $\theta_d = 2$ and $\lambda = 0.5$.

As evident, the non-linear modular network has a lower critical capacity of the small memories but is significantly better for the large memories. As a result, the information capacity [Frolov *et al.*, 1997] of the non-linear modular network is considerably larger than that of the other networks.

The multi-modular network has the advantage that it can successfully encode memories with different modular coding levels. However its overall capacity is limited by that of a single module. Since we deal with a network of large N and limited L this is not a problem. However, when comparing its memory capacity to that of a single module of size NL while disregarding the modular coding issue, clearly the single module network will do much better in the limit of large L.

4 Fault Tolerance in Multi-Modular Networks

We turn now to the effects of pathological input alterations on the retrieval of memories with different modular codings in a network characterized by a non-linear dendritic transmission function $\mathcal{G}[x] = \lambda\Theta[x]$. We start by using randomly distorted input cues in which all modules have the same probability for error. We find that regardless of their modular codings Ω, all memories are equally susceptible to this kind of homogeneous random input noise. The situation is fundamentally different if spatially localized damage is introduced, by setting the input cues to some of the modules to zero. Figure 6 shows the pattern completion quality of a network in which 6 of the 10 modules have their inputs shut off.

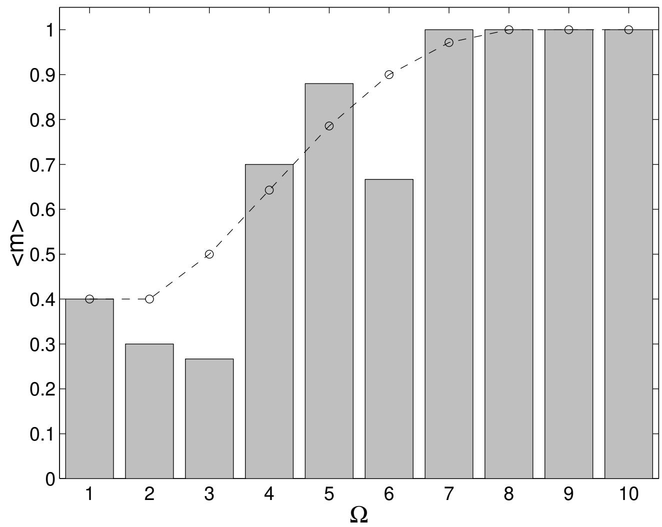

Figure 6: The performance of a multi-modular network when 6 of the 10 modules do not receive input cues. The mean overlap is plotted versus modular coding. The bars represent results of one simulation. The dashed line represents average over many simulations that is estimated analytically. The simulation parameters are: $L = 10$, $N = 500$, $M = 50$, $p = 0.05$, $\lambda = 0.7$, $\theta_d = 2$ and $\theta_n = 0.6$.

The choice of $\theta_d = 2$ in this simulation ensures that the presence of two active external modules is sufficient for the activation of any other module in a memory. Thus, as demonstrated by the shaded bars, memories with modular codings $\Omega = 8$ and above will always be successfully retrieved. The probability for successful retrieval decreases with the modular coding level.

Figure 6 shows that memories that have larger modular codings are more resilient to afferent damage, since their chances of retaining the minimal number of modular inputs required for activating all modules in a memory pattern are significantly higher. Interestingly, these results can account for the psychological findings that concrete items are more resilient to brain damage than abstract ones [Jones, 1985, Warrington and Shallice, 1984]. This follows if one accepts the assumption that concrete items have more attributes than abstract ones (see [Hinton and Shallice, 1991] and [Friedemann, 1998] for a review), which implies that concrete items have higher modular coding, and hence are more resilient.

The use of non-linear step functions such as $\mathcal{G}_1[x] = \lambda\Theta[x]$ enabled us to model non-linear processing of distal inputs. Once we allow for non-linear dendritic processing we may envision more complicated structures of $\mathcal{G}[x]$. An extreme case is one of hierarchical inter-modular connectivity resulting in nested step-functions, such as

$$\mathcal{G}_2 = \lambda\Theta\left[\sum_{k\neq l}^{L}\Theta\left[\sum_{j}^{N} J_{ij}{}^{lk} V_j{}^{k}(t) - \gamma_d \mathcal{Q}_k(t) - \theta_k\right] - \theta_d\right] . \tag{9}$$

This would correspond to a situation in which every external module connects to a different branch of the dendritic tree, and each branch produces its own non-linear processing that is further modulated on the way to the soma by an additional non-linearity. Such semi-local dendritic processing gains biological support from the dynamical behavior of dendritic voltage-gated channels. Multiple EPSPs occurring on the same branch and within a narrow time window may activate voltage-gated channels and produce a much larger response than would occur if they were on separate branches or occurred outside this time window [Johnston *et al.*, 1996]. Currently there exists no direct biological evidence in cortex for focal spatial connectivity of the sort implied by $\mathcal{G}_2$, but the specificity of synaptic connections in other brain regions such as the cerebellum [Ito, 1984] and the olfactory bulb [Shepherd, 1990] may indicate that such connectivity may be plausible.

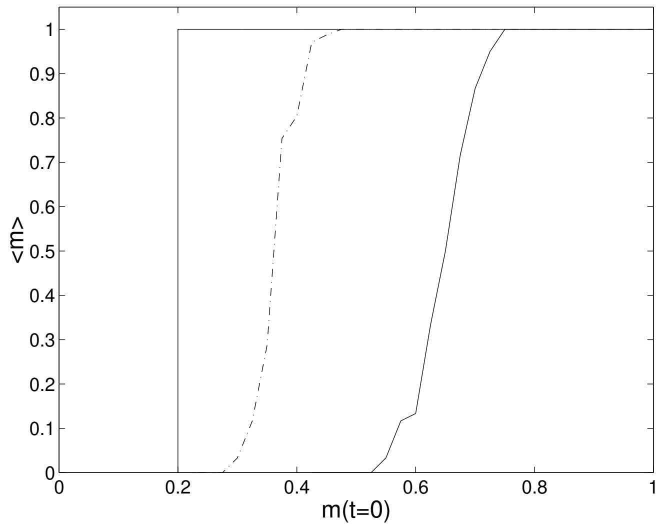

Figure 7: The pattern completion performance of modular networks with different types of non-linear dendritic processing. The mean overlap is plotted *vs.* the overlap of the input cue. The solid line represents the performance of the $\mathcal{G}_2$ network and the dash-dot line represents $\mathcal{G}_1$. The left curve of $\mathcal{G}_2$ corresponds to the case when full input is presented to only one module (out of the 5 that comprise a memory), while the right solid curve corresponds to partial input to all modules. The $\mathcal{G}_1$ network cannot retrieve the memory under the first condition, but performs better for the case of partial inputs to all modules, as shown by the dash-dot curve. Parameters are $L = 5$, $N = 1000$, $p = 0.05$, $\lambda = 0.8$, $\Omega = 5$ $\theta_d = 1.5$ and $\theta_k = 0.7$.

It is interesting to compare the results of the two functions $\mathcal{G}_1[x]$ and $\mathcal{G}_2[x]$. We tested pattern completion under two forms of external partial inputs: 1. an uncorrupted input cue was presented to a single module within the subset of modules in which the memory was encoded, and no input was provided to all other modules. 2. an input cue was given at the same level of partial activity to all modules. As evident from Figure 7, the nested nonlinearities of $\mathcal{G}_2$ enable retrieval even in the extreme case in which only one module is activated by the cue.

5 Summary

In this paper we use the modular structure of the brain as the basis for modeling an associative memory network that can accommodate memories with very different activity levels. All these memories are assumed, for simplicity, to have the same sparseness within a module, but they may have very different numbers of modules in which they are encoded.

To achieve our goal we have made a distinction between two types of synaptic connections, proximal ones that occur on basal dendrites and correspond to inputs from near-by neurons, and distal ones that occur on apical dendrites and transmit information from far-away neurons belonging to different modules. Moreover, we have made the strong assumption that the synaptic currents of the distal synapses are further processed non-linearly by the dendritic tree on their way to the soma. Representing this nonlinear processing by a Heaviside threshold function has the important advantage that the neuronal input that is due to activity in other modules will always have a standard strength. Thus a neuron activated during memory retrieval feels the input of all other neurons in its own module as well as a standard strong input reflecting the fact that several other modules of the same memory pattern are active, but it is not sensitive to the exact number of other active modules. This solves the problem that prevents standard neural networks from accommodating patterns with different activities (or different modular codings using our underlying modular structure). Since these neurons are sensitive to the signal to noise ratio, their parameters have to be tuned to a particular activity level, thus allowing only a limited range of variation. In our case this is no problem since the sparseness within a module is kept constant and the external signal is standardized.

We have worked with a binary neural network throughout this paper. We have checked that our conclusions hold also for a network with analog neurons, in which the step function of Eq. 2 is replaced by the sigmoid $f(x) = 1/(1 + e^{-x/T})$. A single module network of this kind is viable over a certain range of T. Combining L such networks into a multi-modular one does not change the critical temperature for memories coded in one module. Simulations show that memories with higher modular coding are more stable and survive for a larger range of T. The validity of our results in more realistic models of associative memory with spiking neurons is an interesting question for future study.

An interesting outcome of our model is its resilience to corrupted input, or afferent damage, of a focused nature. In conventional neural network models one studies basins of attraction by considering diffuse damage, i.e. random corruption of inputs. Here we have an ordered system, at least as far as its modular structure is concerned. Hence we can study its behavior under focused damage, when some of the modules are deprived of their afferent inputs while keeping all lateral connections intact. We find that memories that have large modular coding are more resilient to this kind of damage. This fact can be connected to the interesting psychological observation of resilience of natural memories in patients with focal damages. Moreover, we find that further dendritic nonlinearities, such as nested Heaviside functions specifying the information transfer between every pair of modules, lead to stronger resilience to focal damage and weaker resilience to diffuse damage.

We believe that the concept of multi-modular networks is necessary for a meaningful implementation of associative neural networks in models of the cortex. A module is naturally described in terms of neurons that share with one another a large number of synaptic connections through the grey matter. We have labeled them as proximal synapses in our work, introducing a distinction between them and the distal apical connections. Whereas in Nature the situation may be much more complicated, it stands to reason that this simple dichotomy should play a leading role, distinguishing between near-by and far-away inputs. It remains to be seen if, indeed, the latter are subject to further nonlinear processing, and if our simple representation can do justice to the complicated dynamical mechanisms that go on in real neural circuits. In any case, we have seen that it leads to interesting consequences and allows for simple modeling that was not possible without it.

Acknowledgment: We thank Isaac Meilijson for very helpful discussions.

References

[Amit, 1995] D. J. Amit. The Hebbian paradigm reintegrated: local reveberations as internal representations. *Behavioural and Brain Science*, 18:617, 1995.

[Friedemann, 1998] P. Friedemann. Words in the brain's language. *Behavioural and Brain Science*, 1998. to be published.

[Frolov *et al.*, 1997] A. A. Frolov, D. Husek, and I. P. Muraviev. Informational capacity and recall quality in sparsely encoded hopfield-like neural network - analytical approaches and computer simulation. *Neural Networks*, 10(5):845 – 855, 1997.

[Herrmann *et al.*, 1995] M. Herrmann, J. A. Hertz, and A. Prugel-Bennet. Analysis of synfire chains. *Network*, 6:403–414, 1995.

[Hetherington and Shapiro, 1993] P. A. Hetherington and L. M. Shapiro. Simulating Hebb cell assemblies: the necessity for partitioned dendritic trees and a post-not-pre LTD rule. *Network*, 4:135–153, 1993.

[Hevner, 1993] R. F. Hevner. More modules. *TINS*, 16(5):178, 1993.

[Hinton and Shallice, 1991] G. E. Hinton and T. Shallice. Lesioning at attractor network: investigations of acquired dyslexia. *Psychological Review*, 98(1):74–95, 1991.

[Hubel and Wiesel, 1977] D. H. Hubel and T. N. Wiesel. Functional architecture of macaque visual cortex. *Proc. R. Soc. Lond. B*, 198:1–59, 1977.

[Ito, 1984] M. Ito. *The cerebellum and neural control.* Raven Press, 1984.

[Johnston *et al.*, 1996] D. Johnston, J. C. Magee, C. M. Colbert, and B. R. Christie. Active properties of neuronal dendrites. *Annu. Rev. Neurosci.*, 19:165 – 186, 1996.

[Jones, 1985] G. V. Jones. Deep dyslexia, imageability, and ease of predication. *Brain and Language*, 24:1–19, 1985.

[Lauro-Grotto *et al.*, 1994] R. Lauro-Grotto, S. Reich, and M. A. Virasoro. The computational role of conscious processing in a model of semantic memory. In *Proceedings of the IIAS Symposium on Cognition Computation and Consciousness*, 1994.

[Magee and Johnston, 1995] J. C. Magee and D. Johnston. Synaptic activation of voltage-gated channels in the dendrites of hippocampal pyramidal neurons. *Science*, 268:301 – 304, 1995.

[Markram *et al.*, 1997] H. Markram, J. Lübke, M. Frotscher, A. Roth, and B. Sakmann. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol.*, 500(2):409 – 440, 1997.

[Martin *et al.*, 1995] A. Martin, J. V. Haxby, F. M. Lalonde, C. L. Wiggs, and L. G. Ungerleider. Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, 270:102 – 105, 1995.

[Martin *et al.*, 1996] A. Martin, C. L. Wiggs, L. G. Ungerleider, and J. V. Haxby. Neural correlates of category-specific knowledge. *Nature*, 379:649 – 652, 1996.

[Mountcastle, 1957] V. B. Mountcastle. Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J. Neurophysiol*, 20:408–434, 1957.

[Mountcastle, 1997] V. B. Mountcastle. The columnar organization of the neocortex. *Brain*, 120:701 – 722, 1997.

[O'Kane and Sherrington, 1993] D. O'Kane and D. Sherrington. A feature retrieving attractor neural network. *Journal of Physics A: Mathematical and General*, 26:2333 – 2342, 1993.

[O'Kane and Treves, 1992] D. O'Kane and A. Treves. Short and long range connections in autoassociative memory. *Journal of Physics A: Mathematical and General*, 25:5055 – 5069, 1992.

[Shepherd, 1990] G. M. Shepherd. *The synaptic organization of the brain*. Oxford, 1990.

[Stuart *et al.*, 1997] G. Stuart, J. Schiller, and B. Sakmann. Action potential initiation and propagation in rat neocortical pyramidal neurons. *Journal of Physiology*, 505:617 – 632, 1997.

[Tsodyks, 1989] M. V. Tsodyks. Associative memory in neural networks with the hebbian learning rule. *Modern Physics Letters B*, 3(7):555–560, 1989.

[Viana and Martínez, 1995] L. Viana and C. Martínez. Dynamics of a neural network composed by two hopfield subnetworks interconnected unidirectionnaly. *Journal de Physique I*, 5:573 – 580, 1995.

[Warrington and Shallice, 1984] E. K. Warrington and T. Shallice. Category specific semantic impairments. *Brain*, 107:829–854, 1984.

[Yuste and Tank, 1996] R. Yuste and D. W. Tank. Dendritic integration in mammalian neurons, a century after cajal. *Neuron*, 16:701–716, 1996.

[Yuste *et al.*, 1994] R. Yuste, M. J. Gutnick, D. Saar, and K. R. Delaney D. Tank. Ca^{2+} accumulation in dendrites of neocortical pyramidal neurons: An apical band and evidence for two functional compartments. *Neuron*, 13:23 – 43, 1994.

A Appendix A: Capacity of a Multi-Modular Network

A.1 Linear Dendritic Processing ($\mathcal{G}[x] = x$)

Consider a modular network in which $\mathcal{G}[x] = x$. The network stores memory patterns with variable modular coding with a given distribution and $p \ll 1$. The degree of global inhibition is taken to be $\gamma_n = \frac{\langle\Omega\rangle}{L} Mp^2$ and $\gamma_d = \frac{\langle\Omega^2\rangle - \langle\Omega\rangle}{L(L-1)} Mp^2$, where $\langle\Omega\rangle$ and $\langle\Omega^2\rangle$ are the first two moments of the distribution of modular codings. We calculate the critical capacity of the modular network using mean-field analysis similar to [Herrmann *et al.*, 1995]. We assume that $pN \gg 1$ and $p \ll 1$, and that memory pattern η^ν has macroscopic overlap. We then derive the fixed-point equations for the modular overlap

$$m^\mu{}_l(t) = \frac{1}{pN} \sum_{i=1}^{N} \eta^\mu{}_{il} V_i^l(t) \ , \tag{10}$$

and activity $\mathcal{Q}_l(t)$ for each one of the L modules in the network[1].

Writing the local field of the i'th neuron in module l in terms of the modular overlap $m^\nu{}_l(t)$ and modular activity $\mathcal{Q}_l(t)$ we get

$$h_i{}^l(t) = \eta^\nu{}_{il} \sum_{k}^{L} m^\nu{}_k(t) + \phi^{int}{}_i + \phi^{ext}{}_i \ , \tag{11}$$

where $\phi^{int}{}_i$ and $\phi^{ext}{}_i$ are the crosstalk internal and external noise terms that result from the overlapping patterns. For large N and finite L, we assume that the sum of $\phi^{int}{}_i$ and $\phi^{ext}{}_i$ is nearly normally distributed with zero mean and variance $\phi^2(t) = \phi^2_{int}(t) + \phi^2_{ext}(t) + 2\phi_{int-ext}(t)$ (the central-limit theorem), where

$$\phi^2_{int}(t) = \alpha\langle\Omega\rangle p\mathcal{Q}_l(t)\left[1 + Np^2\mathcal{Q}_l(t)\right] \ , \tag{12}$$

$$\begin{aligned} \phi^2_{ext}(t) &= \alpha\tfrac{\langle\Omega^2\rangle - \langle\Omega\rangle}{L-1} p \sum_{k\neq l}^{L} \mathcal{Q}_k(t)\left[1 + Np^2\mathcal{Q}_k(t)\right] \\ &\quad + \alpha\tfrac{\langle\Omega^3\rangle - 3\langle\Omega^2\rangle + 2\langle\Omega\rangle}{(L-1)(L-2)} Np^3 \sum_{(k\ n)}^{L} \mathcal{Q}_k(t)\mathcal{Q}_n(t) \ , \end{aligned} \tag{13}$$

[1] Using both order parameters one can derive the overlap function given in Eq. 7.

$$\phi_{int-ext}(t) \quad = \quad \alpha\frac{\langle\Omega^2\rangle-\langle\Omega\rangle}{L-1}Np^3\mathcal{Q}_l(t)\sum_{k\neq l}^{L}\mathcal{Q}_k(t)\ . \tag{14}$$

The capacity, α, is $\frac{M}{LN}$ and the sum over $(k\ n)$ denotes summation over all pairs $k \neq n$ from the interval $[1,\cdots,l-1,l+1,\cdots,L]$. Terms of the order of αp^2 were neglected in both expressions.

The fixed point equations for the modular overlap and activity are calculated from Eq. 2, Eq. 10, Eq. 5 and Eq. 11 by replacing the average over the sites with an average over the Gaussian noise terms. The resulting equations for module l are:

$$m^{\nu}{}_l = \Phi\left(\frac{\theta_n+\theta_d-\sum_k^L m^{\nu}{}_k}{\phi}\right)\ , \qquad \mathcal{Q}_l = m^{\nu}{}_l + \frac{1}{p}\Phi\left(\frac{\theta_n+\theta_d}{\phi}\right)\ , \tag{15}$$

where

$$\Phi(x) = \int_x^\infty \exp\left(-\frac{z^2}{2}\right)\frac{dz}{\sqrt{2\pi}}\ . \tag{16}$$

Solving all $2L$ equations simultaneously for the overlap and activity of memory patterns with modular coding Ω we look for the critical capacity $\alpha_c(\Omega)$ at which these equations become unstable.

A.2 Non-linear Dendritic Processing with the Step Function ($\mathcal{G}[x] = \lambda\Theta[x]$)

Consider a modular network similar to the one introduced in Appendix A.1 in which $\mathcal{G}[x] = \lambda\Theta[x]$. The local field of the i'th neuron in module l in terms of the modular overlap $m^{\nu}{}_l(t)$ and modular activity $\mathcal{Q}_l(t)$ is:

$$h_i{}^l(t) = \eta^{\nu}{}_{il}m^{\nu}{}_l(t) + \phi^{int}{}_i + \lambda\Theta\left[\eta^{\nu}{}_{il}\sum_{k\neq l}^{L}m^{\nu}{}_k(t) + \phi^{ext}{}_i - \theta_d\right]\ , \tag{17}$$

For large N and finite L we assume that $\phi^{int}{}_i$ and $\phi^{ext}{}_i$ are nearly normally distributed with zero mean and variance $\phi^2_{int}(t)$ and $\phi^2_{ext}(t)$ given in Eq. 12 and Eq. 13 respectively.

The fixed point equations for the modular overlap and activity are calculated from Eq. 2, Eq. 7, Eq. 5 and Eq. 17. For $\mathcal{G}[x] = \lambda\Theta[x]$ the internal and external components of the local field are independent random variables[2], but this is not the case for a general $\mathcal{G}[x]$ as will be shown in Appendix B. Thus, by averaging over these two independent Gaussian noise terms we get for the modular overlap and activity

$$\begin{aligned} m^{\nu}{}_{l} &= \left[1-\Phi\left(\frac{\theta_d-\sum_{k\neq l}^{L} m^{\nu}{}_{k}}{\phi_{ext}}\right)\right]\Phi\left(\frac{\theta_n-m^{\nu}{}_{l}}{\phi_{int}}\right) \\ &\quad +\Phi\left(\frac{\theta_d-\sum_{k\neq l}^{L} m^{\nu}{}_{k}}{\phi_{ext}}\right)\Phi\left(\frac{\theta_n-m^{\nu}{}_{l}-\lambda}{\phi_{int}}\right) , \end{aligned} \tag{18}$$

and

$$\mathcal{Q}_l = m^{\nu}{}_{l} + \frac{1}{p}\left\{\left[1-\Phi\left(\frac{\theta_d}{\phi_{ext}}\right)\right]\Phi\left(\frac{\theta_n}{\phi_{int}}\right)+\Phi\left(\frac{\theta_d}{\phi_{ext}}\right)\Phi\left(\frac{\theta_n-\lambda}{\phi_{int}}\right)\right\} . \tag{19}$$

$\alpha_c(\Omega)$ is obtained as in Appendix A.1.

B Appendix B: The Overlap Function for a General Dendritic Function

We derive $m^{\nu}(1)$ for an arbitrary dendritic function $\mathcal{G}[x]$. We start when the network is in state η^{ν} at $t=0$ and we wish to find the network's overlap with this memory pattern at $t=1$. For this purpose we write the overlap function Eq. 7 as

$$\begin{aligned} m^{\nu}(1) &= \frac{L}{p(1-p)\Omega^{\nu}}\Big[(1-p)\Pr(\eta^{\nu}{}_{il}=1)\Pr(V_i^k(1)=1|\eta^{\nu}{}_{il}=1) \\ &\quad - p\Pr(\eta^{\nu}{}_{il}=0)\Pr(V_i^k(1)=1|\eta^{\nu}{}_{il}=0)\Big] \\ &= \int_{\theta_n}^{\infty}\Pr(h_i{}^{l}|\eta^{\nu}{}_{il}=1)dh_i{}^{l} - \frac{L-p\Omega^{\nu}}{(1-p)\Omega^{\nu}}\int_{\theta_n}^{\infty}\Pr(h_i{}^{l}|\eta^{\nu}{}_{il}=0)dh_i{}^{l} . \end{aligned} \tag{20}$$

The conditional probabilities $\Pr(h_i{}^{l}|\eta^{\nu}{}_{il}=1)$ and $\Pr(h_i{}^{l}|\eta^{\nu}{}_{il}=0)$ are the joint density functions of two random variables, the internal and external fields of neuron i in module l given the state of this neuron in pattern η^{ν}. To calculate these density functions we assume

[2] The correlation coefficient of the internal and external components can be calculated numerically by using the bivariate density function given in Eq. 24, and Eq. 21, Eq. 22 and Eq. 23. The calculation yields $\rho[Y,\Theta[X]]=0$, where X and Y are defined in Appendix B.

that the internal field, given in Eq. 4, is distributed normally and we will denote it as Y. Y has mean $\mu_Y = {\eta^\nu}_{il} - \theta_n$ and variance

$$ {\sigma_Y}^2 = \alpha \langle \Omega \rangle p \left(1 + N p^2 \right) \; . \tag{21} $$

We denote the external field, given in Eq. 6, as $\mathcal{G}[X]$ and assume that X is a random variable that distributes normally with mean $\mu_X = (\Omega^\nu - 1){\eta^\nu}_{il} - \theta_n$ and variance

$$ {\sigma_X}^2 = \alpha \tfrac{\langle \Omega^2 \rangle - \langle \Omega \rangle}{L-1} p (\Omega^\nu - 1) \left(1 + N p^2 \right) + \alpha \tfrac{\langle \Omega^3 \rangle - 3\langle \Omega^2 \rangle + 2\langle \Omega \rangle}{(L-1)(L-2)} \left(\begin{array}{c} \Omega^\nu - 1 \\ 2 \end{array} \right) N p^3 \; . \tag{22} $$

The two random variables X and Y have a non-zero correlation coefficient

$$ \rho = \frac{\alpha \frac{\langle \Omega^2 \rangle - \langle \Omega \rangle}{L-1} \left(\Omega^\nu - 1 \right) N p^3}{\sigma_X \sigma_Y} \; , \tag{23} $$

and together they possess a bivariate normal distribution with a joint density function

$$ f_{XY}(x,y) = \frac{\exp \left\{ - \frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_X}{\sigma_X} \right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 - 2\rho \left(\frac{x-\mu_X}{\sigma_X} \right) \left(\frac{y-\mu_Y}{\sigma_Y} \right) \right] \right\}}{2\pi \sigma_X \sigma_Y \left(1 - \rho^2 \right)^{1/2}} \; . \tag{24} $$

The density function of ${h_i}^l = Y + \mathcal{G}[X]$ for an arbitrary dendritic function $\mathcal{G}[x]$ is calculated using the following substitution of variables: $U = X$, $V = Y + \mathcal{G}[X]$ and $W = \mathcal{G}^{-1}[U - V]$ leading to

$$ f_V(v) = \int_{-\infty}^{\infty} f_{XY}(v - \mathcal{G}[w], w) dw \; . \tag{25} $$

The conditional probabilities: $\Pr({h_i}^l | {\eta^\nu}_{il} = 1)$ and $\Pr({h_i}^l | {\eta^\nu}_{il} = 0)$ are then calculated by choosing $\mathcal{G}[x]$, ${\eta^\nu}_{il}$ and Ω, and the overlap function is calculated by using these results as the integrands of Eq. 20.

C Appendix C: Memories with Different Levels of Activity in a Single Module Network

In order to compare the results of our multi-modular network with a standard associative memory model, we study here a single module network[Tsodyks, 1989] with N excitatory neurons. We assume that the network stores M_1 memory patterns η^μ of sparse coding level p and M_2 patterns ξ^ν with coding level f such that $p < f << 1$. The synaptic efficacy J_{ij} between the jth (presynaptic) neuron and the ith (postsynaptic) neuron is chosen in the Hebbian manner

$$J_{ij} = \frac{1}{Np}\sum_{\mu=1}^{M_1} \eta^\mu{}_i \eta^\mu{}_j + \frac{1}{Np}\sum_{\mu=1}^{M_2} \xi^\nu{}_i \xi^\nu{}_j \ . \tag{26}$$

The updating rule for the activity state V_i of the ith binary neuron is given by

$$V_i(t+1) = \Theta\left[h_i(t) - \theta\right] \ , \tag{27}$$

where θ is the threshold and

$$h_i(t) = h_i^e(t) - \frac{\gamma}{p}\mathcal{Q}(t) \tag{28}$$

is the local field. It includes the excitatory Hebbian coupling of all other excitatory neurons,

$$h_i^e(t) = \sum_{j\neq i}^{N} J_{ij} V_j(t) \ , \tag{29}$$

and global inhibition that is proportional to the total activity of the excitatory neurons

$$\mathcal{Q}(t) = \frac{1}{N}\sum_{j}^{N} V_j(t) \ . \tag{30}$$

The overlap $m(t)$ between the network activity and the memory patterns is defined for the two memory populations as

$$m_\xi{}^\nu(t) = \frac{1}{Nf}\sum_{j}^{N} \xi^\nu{}_j V_j(t) \ , \qquad m_\eta{}^\mu(t) = \frac{1}{Np}\sum_{j}^{N} \eta^\mu{}_j V_j(t) \ . \tag{31}$$

In a model with single sparse coding, the critical value of the storage capacity $\alpha = \frac{M}{N}$ is known to be a function of both N and the sparse coding level. Here we will therefore encounter two critical values: $\alpha_{c\xi}$ above which the population of ξ^ν patterns is unstable and $\alpha_{c\eta}$ above which the population of η^μ patterns is unstable. We calculate them for the case of $M_1 = M_2 = \frac{M}{2}$ and $\gamma = M_1 f^2 + M_2 p^2$, using a mean-field analysis similar to [Herrmann *et al.*, 1995] . They are given by the fixed points of the following equations:

$$m_\eta = \Phi\left(\frac{\theta - m_\eta}{\phi}\right) \ , \qquad \mathcal{Q} = p m_\eta + \Phi\left(\frac{\theta}{\phi}\right) \ , \tag{32}$$

and

$$m_\xi = \Phi\left(\frac{\theta - \frac{f}{p}m_\xi}{\phi}\right) \ , \qquad\qquad \mathcal{Q} = f m_\xi + \Phi\left(\frac{\theta}{\phi}\right) \ , \tag{33}$$

where

$$\phi^2 = \frac{1}{2}\alpha\mathcal{Q}\left(1 + \frac{f^2}{p^2}\right) + \frac{1}{2}\alpha N p \mathcal{Q}^2\left(1 + \frac{f^3}{p^3}\right) \tag{34}$$

and

$$\Phi(x) = \int_x^\infty \exp\left(-\frac{z^2}{2}\right)\frac{dz}{\sqrt{2\pi}} \ . \tag{35}$$

To find the critical capacity, $\alpha_{c\eta}$ that corresponds to the η^μ memory patterns, we look for the capacity at which the fixed-point equations become marginally stable, using the limit $N \gg 1$ and $fN \gg 1$. Thus, $\alpha_{c\eta}$ solves

$$\frac{\partial}{\partial\mathcal{Q}}\Phi\left(\frac{\theta}{\phi}\right)\Big|_{\mathcal{Q}=p} = 1 \ , \tag{36}$$

and $\alpha_{c\xi}$ solves

$$\frac{\partial}{\partial\mathcal{Q}}\Phi\left(\frac{\theta}{\phi}\right)\Big|_{\mathcal{Q}=f} = 1 \ . \tag{37}$$