Effective Neuronal Learning with Ineffective Hebbian Learning Rules

Gal Chechik
The Center for Neural Computation
Hebrew University in Jerusalem, Israel
and the School of Mathematical Sciences
Tel-Aviv University Tel Aviv, 69978, Israel
ggal@math.tau.ac.il

Isaac Meilijson
School of Mathematical Sciences
Tel-Aviv University Tel Aviv, 69978, Israel
isaco@math.tau.ac.il

Eytan Ruppin
Schools of Medicine and Mathematical Sciences
Tel-Aviv University Tel Aviv, 69978, Israel
ruppin@math.tau.ac.il

Correspondence should be addressed to:
Gal Chechik
The lab of Prof. Ruppin
School of Mathematical Sciences
Tel-Aviv University Tel Aviv, 69978, Israel
ggal@math.tau.ac.il

October 26, 2000

Abstract

In this paper we revisit the classical neuroscience paradigm of Hebbian learning. We find that it is difficult to achieve effective associative memory storage by Hebbian synaptic learning, since it requires network level information at the synaptic level or sparse coding level. Effective learning can yet be achieved even with non-sparse patterns by a neuronal process that maintains a zero sum of the incoming synaptic efficacies. This weight correction improves the memory capacity of associative networks from an essentially bounded to one that scales linearly with network size. It also enables the effective storage of patterns with multiple levels of activity within a single network. Such neuronal weight correction can be successfully carried out by activity-dependent homeostasis of the neuron's synaptic efficacies, which was recently observed in cortical tissue. Thus, our findings suggest that associative learning by Hebbian synaptic learning should be accompanied by continuous remodeling of neuronally-driven regulatory processes in the brain.

1 Introduction

Synapse-specific changes in synaptic efficacies, carried out by long-term potentiation (LTP) and depression (LTD) (Bliss & Collingridge, 1993), are thought to underlie cortical self-organization and learning in the brain. In accordance with the Hebbian paradigm, LTP and LTD modify synaptic efficacies as a function of the firing of pre and post synaptic neurons. In this paper we revisit the Hebbian paradigm, studying the role of Hebbian synaptic changes in associative memory storage, and their interplay with neuronally driven processes that modify the synaptic efficacies.

Hebbian synaptic plasticity has been the major paradigm for studying memory and self organization in computational neuroscience. Within the associative memory framework, numerous Hebbian learning rules were suggested and their memory performance was analyzed (Amit, 1989). Restricting their attention to networks of binary neurons, (Dayan & Willshaw, 1991; Palm & Sommer, 1996) derived the learning rule that maximizes the network memory capacity, and showed its relation to the covariance learning rule first described by (Sejnowski, 1977). In a series of papers (Palm & Sommer, 1988; Palm, 1992; Palm & Sommer, 1996), Palm and Sommer have further studied the space of possible learning rules. They identified constraints that must be fulfilled to achieve non-vanishing asymptotic memory capacity, and showed that effective learning can be achieved if either the mean output value of the neuron or the correlation between synapses are zero. The current paper extends their work in two ways: by providing a computational procedure that enforces the constraints on effective learning while obeying constraints on locality of information, and by showing how this procedure may be carried out in a biologically plausible manner.

Synaptic changes sub-serving learning have traditionally been complemented by neuronally driven normalization processes in the context of self-organization of receptive fields and cortical maps (von der Malsburg, 1973; Miller & MacKay, 1994; Goodhill & Barrow, 1994; Sirosh & Miikkulainen, 1994) and continuous unsupervised learning as in principal-component-analysis networks (Oja, 1982). In these scenarios normalization is necessary to prevent the excessive growth of synaptic efficacies that occurs when learning and neuronal activity are strongly coupled. This paper focuses on associative memory learning where this excessive synaptic runaway growth is mild (Grinstein-Massica & Ruppin, 1998), and shows that normalization processes are essential even in this simpler learning paradigm. Moreover, while other normalization procedures can prevent synaptic runaway, our analysis shows that a specific neuronally-driven correction procedure that preserves the total sum of synaptic efficacies is essential for effective associative memory storage.

The following section describes the associative memory model and establishes constraints that lead to effective synaptic learning rules. Section 3 describes the main result of this paper, a neuronal weight correction procedure that can modify synaptic efficacies towards maximization of memory capacity. Section 4 studies the robustness of this procedure when storing memory patterns with heterogeneous coding levels. Section 5 presents a biologically plausible realization of the neuronal normalization mechanism in terms of neuronal regulation. Finally, these results are discussed in section 6.

2 Effective Synaptic Learning rules

We study the computational aspects of associative learning in low-activity associative memory networks with binary firing $\{0, 1\}$ neurons. M uncorrelated memory patterns $\{\xi^\mu\}_{\mu=1}^M$ with coding level p (fraction of firing neurons) are stored in an N-neuron network. The ith neuron updates its firing state X_i^t at time t by

$$X_i^{t+1} = \theta(f_i^t), \quad f_i^t = \frac{1}{N}\sum_{j=1}^{N} W_{ij}X_j^t - T, \quad \theta(f) = \frac{1 + sign(f)}{2} \quad , \tag{1}$$

where f_i is its input field (postsynaptic potential) and T is its firing threshold. The synaptic weight W_{ij} between the jth (presynaptic) and ith (postsynaptic) neurons is determined by a general additive synaptic learning rule that depends on the neurons' activity in each of the M stored memory patterns ξ^η

$$W_{ij} = \sum_{\eta=1}^{M} A(\xi_i^\eta, \xi_j^\eta) \quad , \tag{2}$$

where $A(\xi_i^\eta, \xi_j^\eta)$ is a two-by-two synaptic learning matrix that governs the incremental modifications to a synapse as a function of the firing of the presynaptic (column) and postsynaptic (row) neurons

$$A(\xi_i, \xi_j) = \quad \begin{array}{r|c|c|c|} \multicolumn{4}{r}{\text{presynaptic } (\xi_j)} \\ \hline & & 1 & 0 \\ \hline \text{postsynaptic } (\xi_i) & 1 & \alpha & \beta \\ \hline & 0 & \gamma & \delta \\ \hline \end{array} \quad .$$

In conventional biological terms, α denotes an increment following a long-term potentiation (LTP) event, β denotes a heterosynaptic long-term depression (LTD) event, and γ a homosynaptic LTD event.

The parameters $\alpha, \beta, \gamma, \delta$ define a four dimensional space in which all linear additive Hebbian learning rules reside. In order to study this four dimensional space, we conduct a signal-to-noise analysis, as in (Amit, 1989). Such analysis assumes that the synaptic learning rule obeys a zero mean constraint ($E(A) = 0$), otherwise the synaptic values diverge, the noise overshadows the signal and no retrieval in possible (Dayan & Willshaw, 1991). The signal-to-noise ratio of the neuronal input field f_i during retrieval is (see Appendix A.1)

$$\begin{aligned}
\frac{Signal}{Noise} &\equiv \frac{E(f_i|\xi_i = 1) - E(f_i|\xi_i = 0)}{\sqrt{Var(f_i)}} \qquad (3) \\
&= \frac{[A(1,1) - A(0,1)](1-\epsilon) + [A(1,0) - A(0,0)]\epsilon}{\sqrt{\frac{p}{N}}\sqrt{Var\left[W_{ij}\right] + NpCOV\left[W_{ij}, W_{ik}\right]}} = \\
&= \sqrt{\frac{N}{M}} \frac{[A(1,1) - A(0,1)](1-\epsilon) + [A(1,0) - A(0,0)]\epsilon}{\sqrt{pVar\left[A(\xi_i, \xi_j)\right] + Np^2 COV\left[A(\xi_i, \xi_j), A(\xi_i, \xi_k)\right]}},
\end{aligned}$$

where ϵ is a measure of the overlap between the initial activity pattern X^0 and the memory pattern retrieved. As evident from equation (3) and already pointed out by (Palm & Sommer, 1996), when the **postsynaptic covariance** $COV\left[A(\xi_i, \xi_j), A(\xi_i, \xi_k)\right]$ (determining the covariance between the incoming synapses of the postsynaptic neuron) is positive and p does not vanish for large N, the network's memory capacity is bounded, i.e., it does not scale with the network size. As the postsynaptic covariance is non negative (see Appendices A3 and B), **effective learning rules** that obtain linear scaling of memory capacity as a function of the network's size require a vanishing postsynaptic covariance. Intuitively, when the synaptic weights are correlated, adding any new synapse contributes only little new information, thus limiting the number of beneficial synapses that help the neuron estimate whether it should fire or not.

(Palm & Sommer, 1996) have shown that the catastrophic correlation can be eliminated in the general case where the output values of neurons are $\{a, 1\}$ instead of $\{0, 1\}$, by setting $a = p/(1-p)$ or when p asymptotically vanishes for $N \to \infty$. In the biological case, however, it is difficult to think of the output value a as anything but a hardwired characteristic of neurons, while the coding level p may vary from network to network, or even between different patterns.

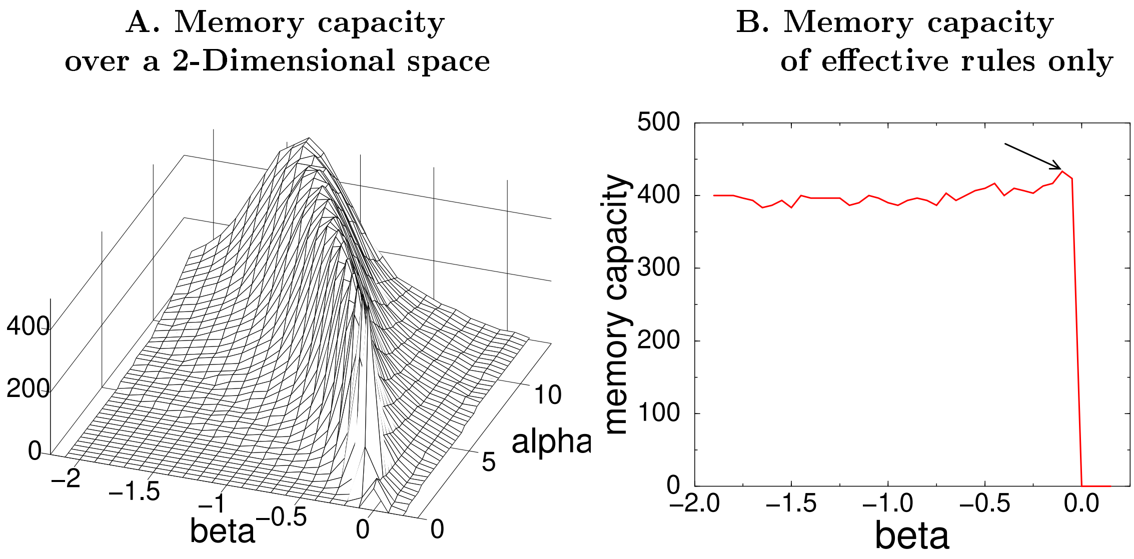

Figure 1: A. Memory capacity of a 1000-neuron network for different values of the free parameters α and β as obtained in computer simulations. The two remaining learning-rule parameters were determined by utilizing a scaling invariance constraint and the zero-synaptic mean constraint, leaving together only two free parameters (See Appendix A3). Capacity is defined as the maximal number of memories that can be retrieved with average overlap bigger than $m = 0.95$ after a single step of the network's dynamics when presented with a degraded input cue with overlap $m_0 = 0.8$. The overlap m^η (or similarity) between the current network's activity pattern X and the memory pattern ξ^η serves to measure retrieval acuity and is defined as $m^\eta = \frac{1}{p(1-p)N}\sum_{j=1}^{N}(\xi_j^\eta - p)X_j$. The coding level is $p = 0.05$. B. Memory capacity of the effective learning rules. The peak values on the ridge of Figure A, are displayed by tracing their projection on the β coordinate. The optimal learning rule (marked with an arrow) performs only slightly better than other effective learning rules.

Figure 1A illustrates the critical role of the postsynaptic covariance by plotting the memory capacity of the network as a function of the learning rule. This is done by focusing on a two dimensional subspace of possibly effective learning rules parametrized by the two parameters α and β, while the two other parameters are set by utilizing a scaling invariance constraint and the requirement that the synaptic matrix should have a zero mean (see the table in appendix A.13). As evident, effective learning rules lie on a narrow ridge (characterized by zero postsynaptic covariance), while other rules provide negligible capacity. Figure 1B depicts the memory capacity of the effective synaptic learning rules that lie on the essentially one-dimensional ridge observed in Figure 1A. It shows that all these effective rules are only slightly inferior to the optimal synaptic learning rule $A(\xi_i, \xi_j) = (\xi_i - p)(\xi_j - p)$ (Dayan & Willshaw, 1991), which maximizes memory capacity.

The results above show that effective learning requires a vanishing postsynaptic covariance constraint, yielding a requirement for a balance between synaptic depression and facilitation, $\beta = \frac{-p}{1-p}\,\alpha$ (see Eqs. 19 - 20). Thus, effective memory storage requires a delicate balance between LTP (α) and heterosynaptic depression (β). This constraint make effective memory storage explicitly dependent on the coding level p which is a global property of the network. It is thus difficult to see how effective rules can be implemented at the synaptic level. Moreover, as shown in Figure 1A, Hebbian learning rules lack robustness as small perturbations from the effective rules may result in a large decrease in memory capacity. Furthermore, these problems cannot be circumvented by introducing a nonlinear Hebbian learning rule of the form $W_{ij} = g\left(\sum_\eta A(\xi_i^\eta, \xi_j^\eta)\right)$ as even for a nonlinear function g the covariance $Cov\left[g(\sum_\eta A(\xi_i^\eta, \xi_j^\eta)), g(\sum_\eta A(\xi_i^\eta, \xi_k^\eta))\right]$ remains positive if $Cov(A(\xi_i, \xi_j), A(\xi_i, \xi_k))$ is positive (see Appendix B). In section 4 we further show that the problem cannot be avoided by using some predefined average coding level for the learning rule.

These observations put forward the idea that within a biological standpoint requiring locality of information and non vanishing coding level, effective associative learning are difficult to realize with Hebbian rules alone.

3 Effective Learning via Neuronal Weight Correction

The above results show that in order to obtain effective memory storage, the postsynaptic covariance must be kept negligible. How then may effective storage take place in the brain with Hebbian learning? We now proceed to show that a neuronally-driven procedure (essentially similar to that assumed by (von der Malsburg, 1973; Miller & MacKay, 1994) to take place during self-organization) can maintain a vanishing covariance and enable effective memory storage by acting upon ineffective Hebbian synapses and turning them into effective ones.

3.1 The Neuronal Weight Correction Procedure

The solution emerges when rewriting the signal-to-noise equation (Eq. 3) as

$$\frac{Signal}{Noise} = \frac{[A(1,1) - A(0,1)](1-\epsilon) + [A(1,0) - A(0,0)]\epsilon}{\sqrt{\frac{p(1-p)}{N} Var\left[W_{ij}\right] + \frac{p^2}{N^2} Var(\sum_{j=1}^{N} W_{ij})}} \quad , \tag{4}$$

showing that the post synaptic covariance is small only if the sum of incoming synapses is nearly constant (see Eq (18) in appendix A2). We thus propose that during learning, as a synapse is modified, its postsynaptic neuron additively modifies all its synapses to maintain the sum of their efficacies at a baseline zero level. As this **neuronal weight correction** is additive, it can be performed either after each memory pattern is stored or at a later time after several memories have been stored.

Interestingly, the joint operation of weight correction over a linear Hebbian learning rule is equivalent to the storage of the same set of memory patterns with another Hebbian learning rule. This new rule has both a zero synaptic mean and a zero postsynaptic covariance, as follows

$$\begin{array}{|c||c|c|} \hline & 1 & 0 \\ \hline 1 & \alpha & \beta \\ \hline 0 & \gamma & \delta \\ \hline \end{array} \quad \Longrightarrow \quad \begin{array}{|c||c|c|} \hline & 1 & 0 \\ \hline 1 & (\alpha-\beta)(1-p) & (\alpha-\beta)(0-p) \\ \hline 0 & (\gamma-\delta)(1-p) & (\gamma-\delta)(0-p) \\ \hline \end{array} \quad .$$

To prove this transformation, focus on a firing neuron in the current memory pattern. When an LTP event occurs, the pertaining synaptic efficacy is strengthened by α, thus all other synaptic efficacies must be reduced by $\frac{\alpha}{N}$ to keep their sum fixed. As there are on average Np LTP events for each memory, all incoming synaptic efficacies will be reduced by αp. This and a similar calculation for quiescent neurons yields the synaptic learning matrix displayed on the right. It should be emphasized that the matrix on the right is not applied at the synaptic level but is the emergent result of the operation of the neuronal mechanism on the matrix on the left, and is used here as a mathematical tool to analyze network performance. Thus, using a neuronal mechanism that maintains the sum of incoming synapses fixed enables the same level of effective performance as would have been achieved by using a zero-covariance Hebbian learning rule, but without the need to know the memories' coding level. Note also that neuronal weight correction applied to the matrix on the right will result in the same matrix, thus no further changes will occur with its re-application.

3.2 An Example

To demonstrate the beneficial effects of neuronal weight correction we apply it to a non-effective rule (having non-zero covariance): we investigate a common realization of the Hebb rule $A(\xi_i, \xi_j) = \xi_i \xi_j$ with inhibition added to obtain a zero-mean input field (otherwise the capacity vanishes) yielding $A(\xi_i, \xi_j) = \xi_i \xi_j - p^2$ (Tsodyks, 1989), or in matrix form

Zero-mean Hebb rule

	1	0
1	$1 - p^2$	$-p^2$
0	$-p^2$	$-p^2$

.

As evident, this learning rule employs both homo-synaptic and hetero-synaptic LTD to maintain a zero mean synaptic matrix, but its postsynaptic covariance is non-zero and is thus still an ineffective rule. Applying neuronal weight correction to the synaptic matrix formed by this rule results in a synaptic matrix which is identical to the one generated without neuronal correction by the following rule

Neuronally corrected Hebb rule

	1	0
1	$1 - p$	$-p$
0	0	0

which has both zero mean and zero postsynaptic covariance. Figure 2 plots the memory capacity obtained with the zero mean Hebb rule, before and after neuronal weight correction, as a function of network size. The memory capacity of the original zero-mean Hebb rule is essentially bounded, while after applying neuronal weight correction it scales linearly with network size.

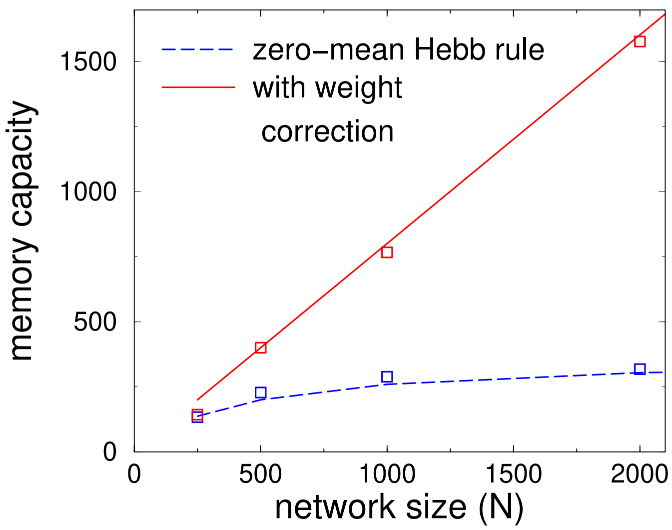

Figure 2: Network capacity as a function of network size. While the original zero-mean learning rule has bounded memory capacity, the capacity becomes linear in network size when the same learning rule is coupled with neuronal weight correction. The lines plot analytical results and the squares designate simulation results ($p = 0.05$).

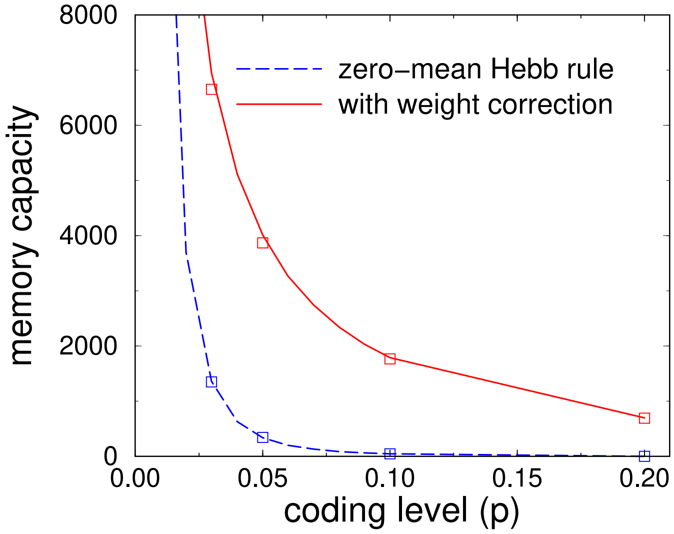

Figure 3: Comparison of network memory capacity for memory patterns with different values of the coding level p, in a network of $N = 5000$ neurons. The effect of neuronal correction is marked for a wide range of the p values, especially in the low coding levels observed in the brain. The lines plot analytical results and the squares designate simulation results.

Figure 3 shows that the beneficial effect of the neuronal correction remains marked for a wide range of coding level values p.

4 Effective learning with Variable Coding Levels

The identification of the post-synaptic covariance as the main factor limiting memory storage, and its dependence on the coding level of the stored memory patterns, naturally raises the common problem of storing memory patterns with heterogeneous coding levels within a single network. The following subsection explores this question by studying networks with a learning rule that was optimally adjusted for some average coding level, but actually store memory patterns that are distributed around it. Due to the dependence of effective learning rules on the coding level of the stored pattern, there exists no single learning rule that can provide effective memory storage of such patterns. However, as we now show in this section, the application of neuronal weight correction provides effective memory storage also under heterogeneous coding levels, testifying for the robustness of the neuronal weight correction procedure.

4.1 Analysis

Bearing in mind the goal of storing patterns with heterogeneous coding levels, we slightly modify the model presented in section 2: The memory pattern ξ^μ $(1 \le \mu \le M)$ has now a coding level p_μ (that is, it has $\mathbf{1}^T \xi^\mu = p_\mu N$ firing neurons out of the N neurons). For simplicity, we focus on a single learning rule

$$W_{ij} = \sum_{\mu=1}^{M} (\xi_i^\mu - a)(\xi_j^\mu - a) \quad , \tag{5}$$

where a is a parameter of the learning rule to be determined. As noted earlier, this rule is the optimal rule for the case of a *homogeneous coding level*, (when all memories share exactly the same coding level), if a is set to their coding level. The overlap m^μ between the current network's activity pattern X and the memory ξ^μ is now defined in terms of the coding level p_μ as $m^\mu = \frac{1}{p_\mu(1-p_\mu)N} \sum_{j=1}^{N} (\xi_j^\mu - p_\mu) X_j$.

A signal-to-noise analysis for the case of heterogeneous coding levels (see appendix C), reveals that both the mean synaptic weight $(E(W) = \sum_{\mu=1}^{M} (p_\mu - a)^2)$ and the postsynaptic covariance $(COV\,[W_{ij}, W_{ik}] = \sum_{\mu=1}^{M} p_\mu(1-p_\mu)(p_\mu - a)^2)$ are strictly positive, yielding

$$\frac{Signal}{Noise} \approx \sqrt{\frac{N}{M}} \frac{(1-a-\epsilon)\sqrt{p_1}}{\sqrt{\frac{1}{M}\sum_{\mu=1}^{M} p_\mu^2 (1-p_\mu)^2 + (2+Np_1)\frac{1}{M}\sum_{\mu=1}^{M} p_\mu(1-p_\mu)(p_\mu - a)^2}}. \tag{6}$$

Moreover, this analysis assumes that the neuronal threshold is optimally set to maximize memory retrieval. Such optimal setting requires that the threshold is set to

$$\begin{aligned} T^{Optimal}(\xi^1) &= \frac{E(f_i|\xi_i^1=1)+E(f_i|\xi_i^1=0)}{2} \\ &= (\frac{1}{2}-a)(1-a-\epsilon)p_1 + p_1\sum_{\mu=1}^{M}(p_\mu - a)^2 \end{aligned} \tag{7}$$

during the retrieval of the memory pattern ξ^1 (Chechik, Meilijson, & Ruppin, 1998). The optimal threshold thus depends both on the coding level of the retrieved pattern p_1 and on the variability of the coding levels p_μ. These parameters are global properties of the network that may be unavailable at the neuronal level, making optimal setting of the threshold biologically implausible.

To summarize, the signal-to-noise analysis reveals three problems that prevent effective memory storage of patterns with varying coding level using a single synaptic learning rule. First, the mean synaptic efficacy is no longer zero, and depends on the coding level variability (Eq. 27). Second, the postsynaptic covariance is non-zero (Eq. 29). Third, the optimal neuronal threshold explicitly depends on the coding level of the stored memory patterns (Eq. 7). As shown in the previous sections, these problems are inherent to all Hebbian additive synaptic learning rules, and are not limited to the learning rule of Eq. (5).

To demonstrate the effects of these problems on the network's memory performance we have stored memory patterns with coding levels that are normally distributed around a, in a network that uses the optimal learning rule and the optimal neuronal threshold for the coding level a (Eqs. 5,7). The memory capacity of such networks as a function of the network size is depicted in Figure 4, for various values of coding level variability. Clearly, even small perturbations from the mean coding level a result in considerable deterioration of memory capacity. Moreover, this deterioration becomes more pronounced for larger networks, revealing a bounded network memory capacity.

4.2 Effective Memory Storage

The above results show that when the coding levels are heterogeneous, both synaptic mean and synaptic covariance are non-zero. However, when applying neuronal weight correction over the learning rule of Eq. (5), the combined effect is equivalent to

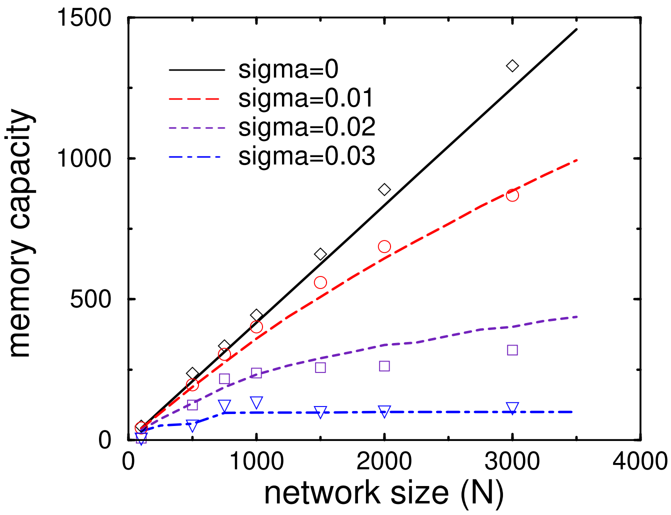

Figure 4: Memory capacity (defined in figure 1) as a function of network's size for various coding level distributions. Coding levels are normally distributed $p_\mu \sim N(a, \sigma^2)$ with mean of $a = 0.1$ and standard deviations of $\sigma = 0, 0.01, 0.02, 0.03$, but clipped to $(0, 1)$. Lines designate analytical results and icons display simulations results.

memory storage using the following synaptic learning rule

$$W_{ij} = \sum_{\mu=1}^{M} (\xi_i^\mu - a)(\xi_j^\mu - p_\mu) \quad . \tag{8}$$

This learning procedure is not equivalent to any predefined learning matrix, as a different learning rule emerges for each stored memory pattern, in a way that is adjusted to its coding level p_μ. Similarly to the case of homogeneous memory patterns with non-optimal learning rule , the application of neuronal weight correction results in a vanishing postsynaptic covariance and mean (Appendix C) and yields

$$\frac{Signal}{Noise} = \sqrt{\frac{N}{M}} \frac{(1 - a - \epsilon)\sqrt{p_1}}{\sqrt{\frac{1}{M}\sum_{\mu=1}^{M} p_\mu^2 (1 - p_\mu)^2 + \frac{1}{M}\sum_{\mu=1}^{M} p_\mu (1 - p_\mu)(a - p_\mu)^2}} \quad . \tag{9}$$

A comparison of Eq. (9) with Eq. (6) readily shows that the dependence of the noise term on network size (evident in Eq. 6) is now eliminated. Thus, a neuronal mechanism that maintains a fixed sum of incoming synapses effectively calibrates to zero the synaptic mean and postsynaptic covariance, providing a memory storage capacity that grows linearly with the size of the network. This is achieved without the need to explicitly monitor the actual coding level of the stored memory patterns.

The neuronal weight correction mechanism solves two of the three problems of storing memory patterns with variable coding level, setting to zero the synaptic mean and postsynaptic covariance. But even after neuronal weight correction is applied the optimal threshold is

$$T^{Optimal}(\xi^1) = (\frac{1}{2} - p_1)(1 - a - \epsilon)p_1 \quad , \tag{10}$$

retaining dependence on the coding level of the retrieved pattern. This difficulty may be partially circumvented by replacing the neuronal threshold with a global inhibitory term (as in (Tsodyks, 1989)). To this end, Eq. (1) is substituted with

$$X_i^{t+1} = \theta(f_i) \ , \quad f_i = \frac{1}{N}\sum_{j=1}^{N}(W_{ij} - I)X_j^t = \frac{1}{N}\sum_{j=1}^{N} W_{ij}X_j^t - \frac{I}{N}\sum_{j=1}^{N} X_j^t \quad , \tag{11}$$

where I is the global inhibition term set to $I = (\frac{1}{2} - a)(1 - a - \epsilon)$. When $E\left[X_j^t\right] = p_\mu$ [1], the mean neuronal field corresponds to a network without global inhibition that uses a neuronal threshold $T = (\frac{1}{2} - a)(1 - a - \epsilon)p_1$. For small p_1 this yields a fair approximation to the optimal threshold of (Eq. 10).

To demonstrate the beneficial effect of neuronal weight correction and activity-dependent inhibition, we turn again to store memory patterns whose coding levels are normally distributed as in Figure 4 using the learning rule of Eq. (5). Figure 5 compares the memory capacity of networks with and without neuronal weight correction and activity-dependent inhibition. The memory capacity is also compared to the case were all memories have the same coding level (dot-dashed line), showing that the application of neuronal weight correction and activity-dependent inhibition (long dashed line) successfully compensates for the coding level variability, obtaining almost the same capacity as the capacity achieved with a homogeneous coding level. Figure 6 plots the network's memory capacity as a function of coding level variability. While the original learning rule provides effective memory storage only when coding levels are close to the mean, the application of a neuronal correction mechanism provides effective memory storage even when storing an ensemble of patterns with high variability of coding levels. The addition of activity-dependent inhibition is mainly needed when the coding level variability is very high.

[1] This assumption is precise when considering a single step of the dynamics and the network is initialized in a state with activity p_μ, as considered here.

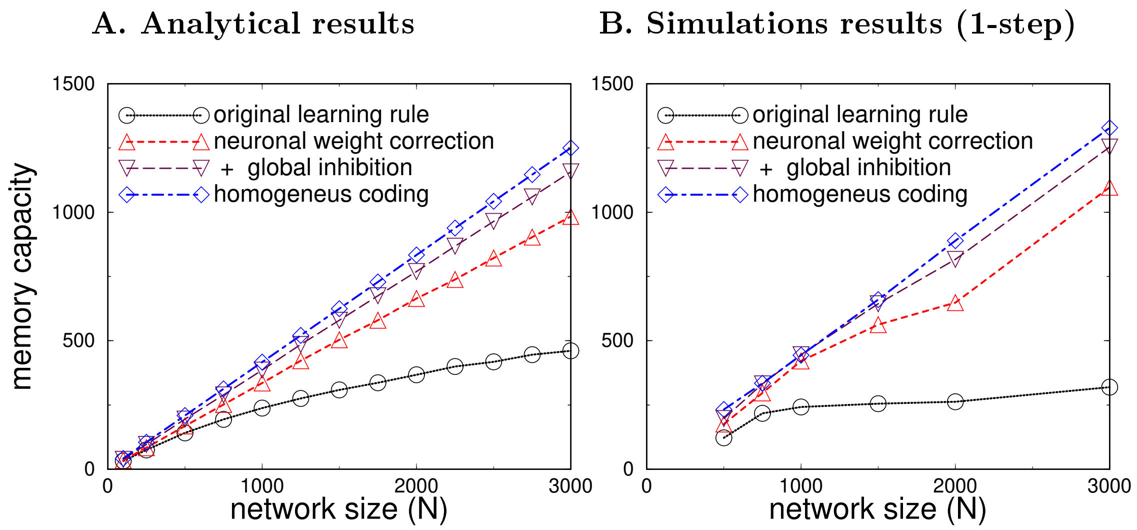

Figure 5: Memory capacity as a function of network size when storing patterns with normally distributed coding levels $p_\mu \sim N(0.1, 0.02^2)$ using the learning rule of Eq. (5). The four curves correspond to: no correction at all (solid line), neuronal weight correction (dashed line), neuronal weight correction with activity dependent inhibition (long dashed line), and homogeneous coding ($p_\mu \equiv a$) (dot dashed line). **A**. Analytical results, **B**. Simulations results of a network performing one step of the dynamics.

5 Neuronal Regulation Implements Weight Correction

The proposed neuronal weight correction algorithm relies on explicit information about the total sum of synaptic efficacies at the neuronal level. Several mechanisms for conservation of the total synaptic strength have been proposed (Miller, 1996). However, as explicit information on the synaptic sum may not be available, we turn to study the possibility of indirectly regulating the total synaptic sum by estimating the neuronal average postsynaptic potential with a Neuronal Regulation (NR) mechanism (Horn, Levy, & Ruppin, 1998). NR maintains the homeostasis of neuronal activity by regulating the postsynaptic activity (input field f_i) of the neuron around a fixed baseline. This homeostasis is achieved by multiplying the neuron's incoming synaptic efficacies by a common factor such that changes in the postsynaptic potential are counteracted by inverse changes in the synaptic efficacies. Such activity-dependent scaling of quantal amplitude of excitatory synapses, which acts to maintain the homeostasis of neuronal firing in a multiplicative manner, has already been observed in cortical tissues by (Turrigiano, Leslie, Desai, & Nelson, 1998; Rutherford,

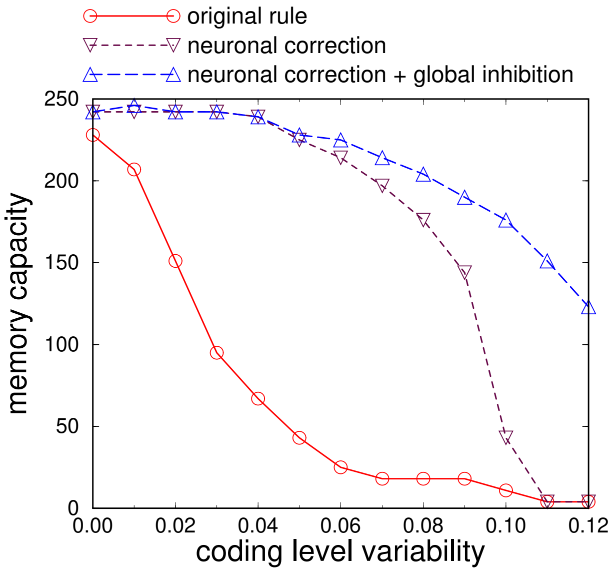

Figure 6: Memory capacity as a function of coding level variability of the stored memory patterns. Patterns were stored in a 1000-neuron network using the learning rule of Eq (1) with $a = 0.2$, but actually storing patterns with clipped normally distributed coding levels $p_\mu \sim N(0.2, \sigma^2)$, for several values of the standard deviation σ.

Nelson, & Turrigiano, 1998). These studies complement their earlier studies showing that neuronal postsynaptic activity can be kept at fixed levels via activity-dependent regulation of synaptic conductances (LeMasson, Marder, & Abbott, 1993; Turrigiano, Abbott, & Marder, 1994).

We have studied the performance of NR-driven correction in an excitatory-inhibitory memory model where excitatory neurons are segregated from inhibitory ones in the spirit of Dale's law (Horn et al., 1998; Chechik, Meilijson, & Ruppin, 1999). This model is similar to our basic model, except that Hebbian learning takes place on the excitatory synapses

$$W_{ij}^{excit} = \sum_{\eta=1}^{M} A(\xi_i^\eta, \xi_j^\eta) \quad , \tag{12}$$

with a learning matrix A that has a positive mean $E(A) = a$. The input field is now

$$f_i^t = \frac{1}{N} \sum_{j=1}^{N} W_{ij}^{excit} X_j^t - W_i^{inhib} \sum_{j=1}^{N} X_j^t \quad , \tag{13}$$

instead of the original term in Equation (1). When $W_{inhib} = Ma$, this model is mathematically equivalent to the model described above in Eqs. (1)-(2).

NR is performed by repeatedly activating the network with random input patterns, and letting each neuron estimate its input field. During this process, each neuron continuously gauges its average input field f_i^t around a zero mean by slowly modifying its incoming excitatory synaptic efficacies in accordance with

$$\kappa \frac{dW_{ij}^{excit}(t')}{dt'} = -W_{ij}^{excit}(t') f_i^t \ . \tag{14}$$

When all W^{excit} are close to a large mean value, multiplying all weights by a common factor approximates an additive change [2]. Figure 7 plots the memory capacity of networks storing memories according to the Hebb rule $W_{ij}^{excit} = \sum_{\eta=1}^{M} A(\xi_i^\eta, \xi_j^\eta) = \sum_{\eta=1}^{M} \xi_i^\eta \xi_j^\eta$, showing how NR, which approximates the additive neuronal weight correction, succeeds in obtaining a linear growth of memory capacity as long as the inhibitory synaptic weights are close to the mean excitatory synaptic values (i.e., the zero synaptic mean constraint is obeyed). Figure 8 plots the temporal evolution of the retrieval acuity (overlap) and the average postsynaptic covariance, showing that NR slowly removes the interfering covariance, improving memory retrieval.

6 Discussion

This paper highlights the role of neuronally-driven synaptic plasticity in remodeling synaptic efficacies during learning. We first showed that effective associative memory learning requires a delicate balance between synaptic potentiation and depression. This balance depends on network-level information thus making it difficult to be implemented at the synaptic level in a distributed memory system. However, our findings show that the combined action of synaptic-specific and neuronally-guided synaptic modifications yields an effective learning system. This allows for the usage of biologically feasible but ineffective synaptic learning rules, as long as they are further modified and corrected by neurally driven weight correction. The learning system obtained is highly robust and enables effective storage of memory patterns with heterogeneous coding level.

[2]The above learning rule results in synapses that are normally distributed, $N(Mp^2, (\sqrt{M}p(1-p))^2)$. Therefore all synapses reside relatively close to their mean when M is large. We may thus substitute $W_{ij}(t') = Mp^2 + \epsilon$ in Eq. (14) yielding $W_{ij}(t'+1) = W_{ij}(t') + \frac{d}{dt'}W_{ij}(t')/\kappa = (Mp^2 + \epsilon)(1 - f_i/\kappa)$. As f_i/κ and ϵ are small, this is well approximated by $W_{ij}(t`) - Mp^2 f_i/k$.

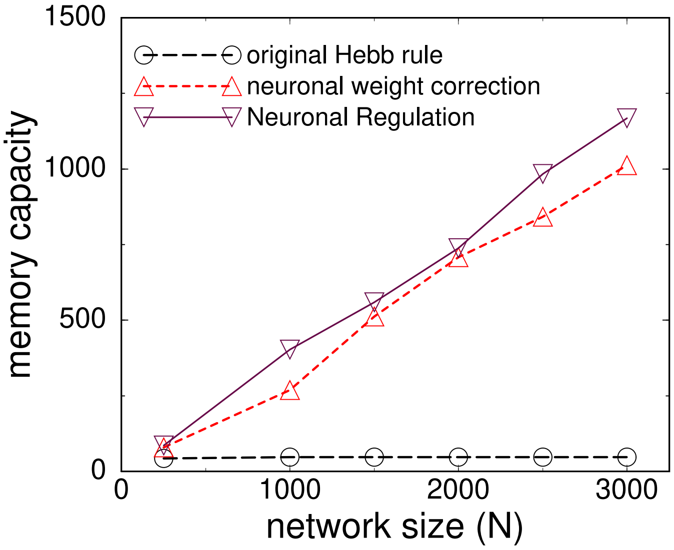

Figure 7: Memory capacity of networks storing patterns via the Hebb rule. Applying NR achieves a linear scaling of memory capacity with a slightly inferior capacity compared with that obtained with neuronal weight correction. Memory capacity is measured as in Figure 2, after the network has reached a stable state. W_i^{inhib} is normally distributed with mean $E(W^{excit}) = p^2 M$ and standard deviation $0.1p^2 M^{0.5}$, where $p = 0.1$.

The space of Hebbian learning rules was already studied by Palm and his colleagues (Palm & Sommer, 1988; Palm, 1992; Palm & Sommer, 1996). They derived the vanishing covariance constraint determining effective learning, and characterized cases where this covariance is vanishing (Palm & Sommer, 1996). The current paper complements their work by describing the procedure of neuronal weight correction that is capable of achieving vanishing postsynaptic covariance, and a possible biological realization of this procedure by the mechanism of Neuronal Regulation.

The characterization of effective synaptic learning rules relates to the discussion of the computational role of heterosynaptic and homosynaptic depression. Previous studies have shown that long-term synaptic depression is necessary to prevent saturation of synaptic values (Sejnowski, 1977), and to maintain zero mean synaptic efficacies (Willshaw & Dayan, 1990). Our study shows that proper heterosynaptic depression is needed to enforce zero postsynaptic covariance - an essential prerequisite of effective learning. The zero covariance constraint implies that the magnitude of heterosynaptic depression should be smaller than that of homosynaptic potentiation by a factor of $(1 - p)/p$. However, effective learning can be obtained regardless of

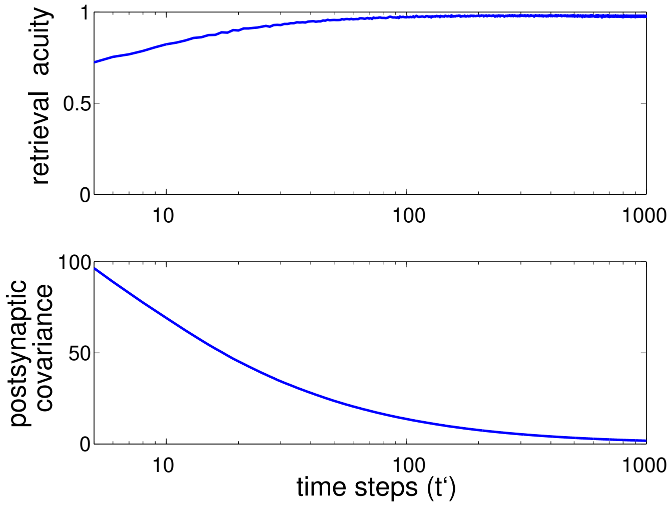

Figure 8: The temporal evolution of retrieval acuity and average postsynaptic covariance in a 1000-neuron network. 250 memories are first stored in the network using the Hebb rule, resulting in a poor retrieval acuity ($m \approx 0.7$ at $t = 0$ in the upper figure). However, as NR is iteratively applied to the network, the retrieval acuity gradually improves as the post-synaptic covariance vanishes. $p = 0.1$, $\kappa = 0.1$, other parameters as in Figure 7.

the magnitude of the homosynaptic depression changes, as long as the zero mean constraint stated above is satisfied. The terms potentiation/depression used in the above context should be cautiously interpreted: As neuronal weight correction may modify synaptic efficacies in the brain, the apparent changes in synaptic efficacies measured in LTD/LTP experiments may involve two kinds of processes: Synaptic-driven processes, changing synapses according to the covariance between pre and post synaptic neurons, and neuronally-driven processes, operating to zero the covariance between incoming synapses of the neuron. Although our analysis pertains to the combined effect of these processes, they may be experimentally segregated as they operate on different time scales and modify different ion channels (Bear & Abraham, 1996; Turrigiano et al., 1998). Thus, the relative weights of neuronal versus synaptic processes can be experimentally tested by studying the temporal changes in synaptic efficacy following LTP/LTD events, and comparing them with the theoretically predicted potentiation and depression end values.

Several forms of synaptic constraints were previously suggested in the literature to improve the stability of Hebbian learning - such as preserving the sum of synaptic strengths or the sum of their squares (von der Malsburg, 1973; Oja, 1982). Our analysis shows that in order to obtain effective memory storage it is the sum of synaptic strengths which must be preserved, thus predicting that it is this specific form of normalization that occurs in the brain. Interestingly, recent studies of spike triggered synaptic plasticity changes both in hippocampal and tectal neurons (Bi & Poo, 1998; Zhang, Tao, Holt, Harris, & Poo, 1998), were shown to result in a normalization mechanism that preserves the total synaptic sum (Kempter, Gerstner, & Hemmen, 1999).

The effects of various normalization procedures were thoroughly studied in the context of self organization of receptive fields and cortical maps (Miller & MacKay, 1994; Goodhill & Barrow, 1994). In these domains, it was shown that subtractive and additive enforcement of normalization constraints may lead to different representations and network architectures. It is thus interesting to note that in the case of associative memory network studied here, the procedure of enforcing the fixed sum constraint has only a minor effect on memory performance. As shown in Figure 1B, memory capacity only slightly depends on the underlying synaptic learning rule, as long as the learning rules constraints are met. Moreover, studying additive and non-additive normalization procedures (e.g. multiplicative normalization after adding some positive constant to the synaptic weights) we found that both procedures yielded almost the same memory performance.

Our results, obtained within the paradigm of auto associative memory networks, apply also to hetero-associative memory networks. More generally, neuronal weight correction qualitatively improves the ability of a neuron to correctly discriminate between a large number of input patterns. It thus enhances the computational power of the single neuron and may be applied in other learning paradigms. This interplay between cooperative and competitive synaptic changes is likely to plays a fundamental computational role in a variety of brain functions such as sensory processing and associative learning.

Appendix

A Signal-To-Noise Analysis For a General Learning Matrix

A.1 Signal-To-Noise Ratio of the Neuronal Input Field

We calculate the signal-to-noise ratio of a network storing memory patterns according to a learning matrix A with zero mean $E(A(\xi_i, \xi_j)) = p^2 A(1,1) + p(1-p)A(1,0) + (1-p)pA(0,1) + (1-p)^2 A(0,0) = 0$. Restricting attention, without loss of generality, to the retrieval of memory pattern ξ^1, the network is initialized in a state X generated independently of the other memory patterns. This state is assumed to have activity level p (thus equal to the coding level of ξ^1), and overlap $m_0^1 = \frac{(1-p-\epsilon)}{(1-p)}$ with ξ^1, where $\epsilon = P(X_i = 0|\xi_i^1 = 1) = (\frac{1-p}{p})P(X_i = 1|\xi_i^1 = 0)$. Denoting $W_{ij}^* = \sum_{\eta=2}^{M} A(\xi_i^\eta, \xi_j^\eta) = W_{ij} - A(\xi_i^1, \xi_j^1)$, we use the fact that W^* and X are independent and that $E[W^*] = 0$, and write the conditional mean of the neuron input field

$$
\begin{aligned}
E\left[f_i|\xi_i^1\right] &= E\left[\frac{1}{N}\sum_{j=1}^{N} W_{ij}X_j|\xi_i^1\right] - T = \qquad (15)\\
&= E\left[\frac{1}{N}\sum_{j=1}^{N} A(\xi_i^1, \xi_j^1)X_j|\xi_i^1\right] + E\left[\frac{1}{N}\sum_{j=1}^{N} W_{ij}^* X_j|\xi_i^1\right] - T = \\
&= E\left[\frac{1}{N}\sum_{j=1}^{N} A(\xi_i^1, \xi_j^1)X_j|\xi_i^1\right] + \frac{1}{N}\sum_{j=1}^{N} E[W_{ij}^*]E[X_j] - T = \\
&= E\left[\frac{1}{N}\sum_{j=1}^{N} A(\xi_i^1, \xi_j^1)X_j|\xi_i\right] - T = \\
&= A(\xi_i^1, 1)P(X_j = 1|\xi_j^1 = 1)P(\xi_j^1 = 1) + \\
&\quad + A(\xi_i^1, 0)P(X_j = 1|\xi_j^1 = 0)P(\xi_j^1 = 0) - T = \\
&= A(\xi_i^1, 1)(1-\epsilon)p + A(\xi_i^1, 0)\epsilon\frac{p}{1-p}(1-p) - T \quad .
\end{aligned}
$$

Since W^* and X are independent and the noise contribution of single memories is negligible, the variance is

$$
\begin{aligned}
Var\left[f_i|\,\xi_i^1\right] &= Var\left[\frac{1}{N}\sum_{j=1}^{N} W_{ij}X_j\right] \approx Var\left[\frac{1}{N}\sum_{j=1}^{N} W_{ij}^* X_j\right] \approx \qquad (16)\\
&\approx \frac{1}{N}Var\left[W_{ij}^* X_j\right] + COV\left[W_{ij}^* X_j, W_{ik}^* X_k\right] = \\
&= \frac{p}{N}Var\left[W_{ij}^*\right] + p^2 COV\left[W_{ij}^*, W_{ik}^*\right] \approx
\end{aligned}
$$

$$\approx \frac{M}{N} p \, Var\left[A(\xi_i, \xi_j)\right] + Mp^2 \, Cov\left[A(\xi_i, \xi_j), A(\xi_i, \xi_k)\right] \quad .$$

The neuronal field's noise is thus dominated by the covariance between its incoming synaptic weights. Combining equations (16)-(15), the signal-to-noise ratio of the neurons input field is

$$\begin{aligned} \frac{Signal}{Noise} &= \frac{E(f_i|\xi_i = 1) - E(f_i|\xi_i = 0)}{\sqrt{Var(f_i|\xi_i)}} = \qquad (17) \\ &= \sqrt{\frac{N}{M}} \frac{[A(1,1) - A(0,1)](1-\epsilon) + [A(1,0) - A(0,0)]\epsilon}{\sqrt{pVar\left[A(\xi_i, \xi_j)\right] + Np^2 Cov\left[A(\xi_i, \xi_j), A(\xi_i, \xi_k)\right]}} . \end{aligned}$$

When the postsynaptic covariance is zero, the signal-to-noise ratio remains constant as M grows linearly with N, thus implying a linear memory capacity. However, when the covariance is a positive constant, the term on the right is almost independent of N, and the memory capacity is bounded.

A.2 Obtaining Vanishing Covariance

The variance of the input field can also be expressed as

$$\begin{aligned} Var\left[f_i | \, \xi_i^1\right] &\approx Var\left[\frac{1}{N} \sum_{j=1}^{N} W_{ij}^* X_j\right] = \qquad (18) \\ &= \frac{1}{N^2}\left[\sum_{j=1}^{N} \sum_{k=1, k\neq j}^{N} Cov(W_{ij}^*, W_{ik}^*)\right] E^2(X_j) + \frac{1}{N^2}\left[\sum_{j=1}^{N} Var(W_{ij}^*)\right] E(X_j) = \\ &= \frac{p^2}{N^2} Var(\sum_{j=1}^{N} W_{ij}^*) - \frac{p^2}{N} Var(W_{ij}^*) + \frac{p}{N} Var(W_{ij}^*) p = \\ &= \frac{1}{N} p(1-p) Var\left[W_{ij}^*\right] + \frac{p^2}{N^2} Var(\sum_{j=1}^{N} W_{ij}^*) \quad . \end{aligned}$$

Thus, keeping the sum of the incoming synapses fixed results in a beneficial effect similar to that of removing postsynaptic covariance, and further improves the signal-to-noise ratio by a factor $\frac{1}{\sqrt{1-p}}$. The postsynaptic sum $(\sum W_{ij}^*)$ remains fixed if each memory pattern has exactly pN firing neurons out of the N neurons of the network.

A.3 Effective Learning Rules in Two Dimensional Space

The parameters $\alpha, \beta, \gamma, \delta$ of the table in section 2 define a four dimensional space in which all linear additive Hebbian learning rules reside. To visualize this space, figure 1 focuses on a reduced, two-dimensional space utilizing a scaling invariance constraint and the requirement that the synaptic matrix should have a mean zero. These yield the following rule, having two free parameters (α, β) only

$A(\xi_i, \xi_j) =$ postsynaptic (ξ_i), presynaptic (ξ_j):

	1	0
1	α	β
0	c	$f(\alpha, \beta, c)$

where c is a scaling constant and $f(\alpha, \beta, c) = \frac{-1}{(1-p)^2}\left[p^2\alpha + p(1-p)(c+\beta)\right]$ is set to enforce the zero mean constraint. The covariance of this learning matrix when setting $c = 1$ is

$$Cov\left[A(\xi_i, \xi_j), A(\xi_i, \xi_k)\right] = \frac{p}{(1-p)}\left[p\,\alpha + (1-p)\beta\right]^2 \tag{19}$$

that is always non-negative and equals zero only when

$$\beta = \frac{-p}{1-p}\,\alpha \quad . \tag{20}$$

B The Postsynaptic Covariance of Non-Additive Learning Rules

In this section we show that the postsynaptic covariance cannot be zeroed by introducing a non-additive learning rule of the form

$$W_{ij} = g\left(\sum_{\eta=1}^{M} A(\xi_i^\eta, \xi_j^\eta)\right) \tag{21}$$

for some nonlinear function g.

To show that, note that when X, Y are positively correlated random variables with marginal standard normal distribution and $E(g(X)) = 0$, we can write (using independent normally distributed random variables U, V, W)

$$\begin{aligned} E\left[g(X)g(Y)\right] &= E\left[g(U+V)g(W+V)\right] = E\left[E(g(U+V)g(W+V)|V)\right] \\ &= E\left[E(g(U+V)|V)^2\right] = Var\left[E(g(U+V)|V)\right] \geq 0. \end{aligned} \tag{22}$$

Equality holds only when $\phi(v) = E(g(U+V)|V=v) = E(g(U+v))$ is a constant function, and as such it must be zero because $E(g(X)) = 0$. To further show that the equality holds only when g is constant, we look at

$$\begin{aligned} 0 &= E(g(U+v)) = \int \frac{1}{\sqrt{2\pi}\sigma} g(v+u) e^{\frac{-u^2}{2\sigma^2}} du = \\ &= e^{-\frac{v^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\sigma} \int e^{\frac{vt}{\sigma^2}} g(t) e^{-\frac{t^2}{2\sigma^2}} dt \quad , \end{aligned} \tag{23}$$

or $\Psi(v) = \int e^{\frac{vt}{\sigma^2}} g(t) e^{-\frac{t^2}{2\sigma^2}} dt \equiv 0$. As $\Psi(v) = 0$ is the Laplace transform of $g(t)e^{-\frac{t^2}{2\sigma^2}}$, g is almost everywhere uniquely determined and the solution $g = 0$ is essentially the only solution.

C Signal-to-Noise Analysis for Heterogeneous Coding Levels

Analysis follows similar lines to those of section A.1, but this time the initialization state X has activity p_1 and overlap $m_0^1 = \frac{(1-p_1-\epsilon)}{(1-p_1)}$ with ξ^1, where $\epsilon = P(X_i = 0|\xi_i^1 = 1) = (\frac{1-p_1}{p_1})P(X_i = 1|\xi_i^1 = 0)$. The conditional mean of the neuronal input field in this case is

$$E\left[f_i|\xi_i^1\right] = (\xi_i^1 - a)(1 - a - \epsilon)p_1 + p_1 \sum_{\mu=2}^{M} (p_\mu - a)^2 - T \quad , \tag{24}$$

and its variance is

$$V\left[f_i\right] \approx \frac{1}{N} p_1 \, V\left[W_{ij}\right] + p_1^2 \, COV\left[W_{ij}, W_{ik}\right] \quad , \tag{25}$$

yielding

$$\frac{Signal}{Noise} \approx \frac{(1-a-\epsilon)p_1}{\sqrt{\frac{1}{N} p_1 V\left[W_{ij}\right] + p_1^2 COV\left[W_{ij}, W_{ik}\right]}}. \tag{26}$$

For large enough networks the noise term in the neuronal input field is dominated by the postsynaptic covariance $COV\left[W_{ij}, W_{ik}\right]$ between the efficacies of the incoming synapses. To obtain the signal-to-noise ratio as a function of the distribution of the coding levels $\{p_\mu\}_{\mu=1}^M$, we calculate the relevant moments of the synaptic weights' distribution

$$E(W_{ij}) = \sum_{\mu=1}^{M} E\left[W_{ij}^\mu\right] = \sum_{\mu=1}^{M} (p_\mu - a)^2 \quad , \tag{27}$$

$$\begin{aligned} V(W_{ij}) &= \sum_{\mu=1}^{M} V\left[W_{ij}^\mu\right] = \sum_{\mu=1}^{M} E\left[(W_{ij}^\mu)^2\right] - E^2\left(W_{ij}^\mu\right) = \\ &= \sum_{\mu=1}^{M} \left[p_\mu - 2p_\mu a + a^2\right]^2 - \sum_{\mu=1}^{M} (p_\mu - a)^4 \\ &= \sum_{\mu=1}^{M} p_\mu (1-p_\mu) \left[p_\mu(1-p_\mu) + 2(p_\mu - a)^2\right] \quad , \end{aligned} \tag{28}$$

and

$$
\begin{aligned}
COV \qquad & [W_{ij}, W_{ik}] = \qquad (29)\\
\approx \quad & \sum_{\mu=1}^{M} E\left[(\xi_i^\mu - a)^2(\xi_j^\mu - a)(\xi_k^\mu - a)\right] - \sum_{\mu=1}^{M} E^2\left[(\xi_i^\mu - a)(\xi_j^\mu - a)\right] =\\
= \quad & \sum_{\mu=1}^{M} p_\mu(1-p_\mu)(p_\mu - a)^2 \quad .
\end{aligned}
$$

Substituting Eqs. (27)-(29) in the signal-to-noise ratio (Eq. 26) one obtains

$$
\frac{Signal}{Noise} \approx \sqrt{\frac{N}{M}} \frac{(1-a-\epsilon)\sqrt{p_1}}{\sqrt{\frac{1}{M}\sum_{\mu=1}^{M} p_\mu^2(1-p_\mu)^2 + (2+Np_1)\frac{1}{M}\sum_{\mu=1}^{M} p_\mu(1-p_\mu)(p_\mu - a)^2}} . \tag{30}
$$

When applying neuronal weight correction, the resulting learning procedure is equivalent to using the rule

$$
W_{ij} = \sum_{\mu=1}^{M} (\xi_i^\mu - a)(\xi_j^\mu - p_\mu) \quad , \tag{31}
$$

that has the following moments

$$
E(W_{ij}) = 0 \quad , \tag{32}
$$

$$
V(W_{ij}) = \sum_{\mu=1}^{M} p_\mu^2(1-p_\mu)^2 + \sum_{\mu=1}^{M} (p_\mu - a)^2 p_\mu(1-p_\mu) \quad , \tag{33}
$$

$$
COV\,[W_{ij}, W_{ik}] = 0 \quad , \tag{34}
$$

and results in a signal-to-noise ratio

$$
\frac{Signal}{Noise} = \sqrt{\frac{N}{M}} \frac{(1-a-\epsilon)\sqrt{p_1}}{\sqrt{\frac{1}{M}\sum_{\mu=1}^{M} p_\mu^2(1-p_\mu)^2 + \frac{1}{M}\sum_{\mu=1}^{M} p_\mu(1-p_\mu)(a - p_\mu)^2}} \quad . \tag{35}
$$

References

Amit, D. J. (1989). *Modeling brain function. the world of attractor neural networks.* Cambridge.

Bear, M. F., & Abraham, W. C. (1996). Long term depression in hippocampus. *Annu. Rev. Neurosci.*, *19*, 437-462.

Bi, Q., & Poo, M. M. (1998). Precise spike timing determines the direction and extent of synaptic modifications in cultured hippocampal neurons. *J. Neurosci*, *18*, 10464-10472.

Bliss, T. V. P., & Collingridge, G. L. (1993). Syanptic model of memory: long-term potentiation in the hippocampus. *Nature*, *361*, 31-39.

Chechik, G., Meilijson, I., & Ruppin, E. (1998). Synaptic pruning during development: A computational account. *Neural Computation*, *10*(7), 1759-1777.

Chechik, G., Meilijson, I., & Ruppin, E. (1999). Neuronal regulation: A mechanism for synaptic pruning during brain maturation. *Neural Computation*, *11*(8), 2061-2080.

Dayan, P., & Willshaw, D. J. (1991). Optimizing synaptic learning rules in linear associative memories. *Biol. Cyber.*, *65*, 253.

Goodhill, G. J., & Barrow, H. G. (1994). The role of weight normalization in competitive learning. *Neural Computation*, *6*, 255-269.

Grinstein-Massica, A., & Ruppin, E. (1998). Synaptic runaway in associative networks and the pathogenesis of schizophrenia. *Neural Computation*, *10*, 451-465.

Horn, D., Levy, N., & Ruppin, E. (1998). Synaptic maintenance via neuronal regulation. *Neural Computation*, *10*(1), 1-18.

Kempter, R., Gerstner, W., & Hemmen, J. L. van. (1999). Hebbian learning and spiking neurons. *Physical Review E*, *59*, 4498-4514.

LeMasson, G., Marder, E., & Abbott, L. F. (1993). Activity-dependent regulation of conductances in model neurons. *Science*, *259*, 1915-1917.

Miller, K. D. (1996). Synaptic economics: Competition and cooperation in synaptic plasticity. *Neuron*, *17*, 371-374.

Miller, K. D., & MacKay, D. J. C. (1994). The role of constraints in Hebbian learning. *Neural Computation*, *6*(1), 100-126.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, *15*, 267–273.

Palm, G. (1992). On the asymptotic information storage capacity of neural networks. *Neural Computation*, *4*, 703-711.

Palm, G., & Sommer, F. (1988). On the information capacity of local learning rules. In R. Eckmiller & C. von der Malsburg (Eds.), *Neural computers* (p. 271-280). Springer-Verlag.

Palm, G., & Sommer, F. (1996). Associative data storage and retrielval in neural networks. In E. Domani, J. vanHemmen, & K. Schulten (Eds.), *Models of neural networks III. association, generalization and represantation* (p. 79-118). Springer.

Rutherford, L. C., Nelson, S. B., & Turrigiano, G. G. (1998). BDNF has opposite effects on the quantal amplitude of pyramidal neuron and interneuron excitatory synapses. *Neuron*, *21*, 521-530.

Sejnowski, T. J. (1977). Statistical constraints on synaptic plasticity. *J. Theo. Biol.*, *69*, 385-389.

Sirosh, J., & Miikkulainen, R. (1994). Cooperative self-organization of afferent and lateral connections in cortical maps. *Biological Cybernetics*, *71*, 66-78.

Tsodyks, M. V. (1989). Associative memory in neural networks with Hebbian learning rule. *Modern Physics letters*, *3*(7), 555-560.

Turrigiano, G. G., Abbott, L. F., & Marder, E. (1994). Activity-dependent changes in the intrinsic properties of neurons. *Science*, *264*, 974-977.

Turrigiano, G. G., Leslie, K., Desai, N., & Nelson, S. B. (1998). Activity dependent scaling of quantal amplitude in neocoritcal pyramidal neurons. *Nature*, *391*(6670), 892-896.

von der Malsburg, C. (1973). Self organization of orientation sensitive cells in the striate cortex. *Kybernetik*, *14*, 85-100.

Willshaw, D. J., & Dayan, P. (1990). Optimal plasticity from matrix memories: What goes up must come down. *Neural Computation*, *2*(1), 85-93.

Zhang, L., Tao, H. Z., Holt, C., Harris, W., & Poo, M. M. (1998). A critical window in the cooperation and competition among developing retinotectal synapses. *Nature*, *395*, 37-44.