Memory Capacity of Balanced Networks

Yuval Aviel[1,3], David Horn[2] and Moshe Abeles[1]

[1] Interdisciplinary Center for Neural Computation, Hebrew University, Jerusalem, Israel.

[2] School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel.

[3] To whom correspondence should be sent. aviel@cc.huji.ac.il

24 June 2004

Abstract

We study the problem of memory capacity in balanced networks of spiking neurons. Associative memories are represented by either synfire chains (SFC) or Hebbian cell assemblies (HCA). Both can be embedded in these balanced networks by a proper choice of the architecture of the network. The size w_E of a pool in a SFC, or of an HCA, is limited from below and from above by dynamical considerations. Proper scaling of w_E by $\sqrt{K}$, where K is the total excitatory synaptic connectivity, allows us to obtain a uniform description of our system for any given K. Using combinatorial arguments we derive an upper limit on memory capacity. The capacity allowed by the dynamics of the system, α_c, is measured by simulations. For HCA we obtain α_c of order 0.1, and for SFC we find values of order 0.065.

The capacity can be improved by introducing 'shadow patterns', inhibitory cell assemblies that are fed by the excitatory assemblies in both memory models. This leads to a doubly-balanced network, where, in addition to the usual global balancing of excitation and inhibition, there exists specific balance between the effects of both types of
assemblies on the background activity of the network.
For each one of the memory models, and for each network architecture, we obtain an allowed region (phase space) for $w_E / \sqrt{K}$ in which the model is viable.

1. Introduction

An interesting property of neural networks is their ability to function as memory devices. In this mode memories are embedded in the synaptic connections of the network, to be recalled later. Recalling is typically done by external ignition of part of the desired memory. The system, using the given hint, should then settle on an associated memory, hence it is called an 'associative memory' model. Reading out the network's activity, it is possible to decide if the input memory exists or not, and if it does, then what are its constituents.

One method of embedding memories is that of 'attractor neural networks' (Amit, 1989), where an initial input may lead to dynamical flow into an attractor, that serves as a Hebbian Cell Assembly (HCA) (Hebb, 1949) representing the recalled memory through elevated firing rates of its neurons. Another method is that of storing memories in *spatio-temporal patterns* of activity, such as Synfire Chains (SFC)

(Abeles, 1991). The memories are encoded in the precise firing patterns produced by the given external input. As this method involves fine temporal structure, it is capable of dynamic binding (Bienenstock, 1995). Applying binding (von der Malsburg & Schneider, 1986) to HCAs one has to resort to other temporal means (Horn, Sagi, & Usher, 1991).

A network memory model should allow a neuron to participate in more than one memory, i.e. have memories stored in a distributed manner. In this case a neuron may receive input also when it is not supposed to fire. This input is due to the overlap between memories and should be treated as noise. Moreover, background activity in the absence of any memory retrieval, should also be allowed, and may be regarded as being due to some other noise source. One should keep in mind that the statistics of an active cortical tissue is one of high variability. An irregular spike train is observed on the individual neuron level, together with an asynchronous activity on the global level (Shadlen & Newsome, 1994). Among the sources of that variability are the 50% inputs from other cortical areas (Braitenberg & Schuz, 1991) and probabilistic synaptic release (Huang & Stevens, 1997).

A network model capable of displaying stationary noisy activity of this kind is the *balanced network* (BN) model. A network is said to be balanced (Shadlen & Newsome, 1994; van Vreeswijk & Sompolinsky, 1998) if each neuron in the network receives equal amounts of excitation and inhibition. Its membrane potential will then fluctuate around some mean value and the firing process is noise driven, and therefore irregular (Abeles, 1982; Gerstein & Mandelbrot, 1964). BNs have been shown (Brunel, 2000) to mimic the *in-vivo* firing statistics of cortical tissue, and it is

therefore plausible that cortical neurons receive balanced input. BNs have also been shown (Brunel, 2000; van Vreeswijk & Sompolinsky, 1998) to have a stable asynchronous state (AS).

Generally, BNs assume sparse and random connectivity. It is possible to embed memories in the connections of a BN, but this would violate the random connectivity assumption. The main consequence of introducing ordered connectivity in an otherwise random connectivity matrix is the appearance of a new critical point beyond which the AS is unstable. This was studied in detail in (Aviel, Mehring, Horn, & Abeles, 2003).

Having set the background, we will now turn to our memory models of choice.

A milestone in the theory of attractor neural network is the Hopfield model (Hopfield, 1982), capable of storing and retrieving binary vectors by a clever construction of the connectivity matrix of the neural network. The models of Hopfield, as well as some predecessors (Willshaw, Buneman, & Longuet-Higgins, 1969) and followers (Tsodyks & Feigelman, 1988) are, however, based on binary rather than spiking neurons (but see (Treves, 1990) for a Hopfield model of linear threshold neurons). All attractor neural networks serve as some kind of implementation of the idea of *Hebbian cell assemblies* (HCA) (Hebb, 1949), i.e. they contain neuronal ensembles that represent a memory by elevated firing rates. The cell assembly may be characterized by denser intra-assembly connectivity that facilitates a sustained activity upon partial assembly excitation (Braitenberg, 1978). The Hopfield model is a step backwards from the biological complexity of HCAs but it allows detailed analysis. Gerstner (Gerstner & van Hemmen, 1992) showed how the Hopfield model

may be incorporated in a network of spiking neurons.

A network that reacts to external inputs with a reproducible firing patterns (over neurons and time), is said to code its input by spatio-temporal patterns. The model we use for producing spatio-temporal patterns is the *synfire chain* (SFC) (Abeles, 1982). For other interesting models, see (Izhikevich, Gally, & Edelman, 2003; Levy, Horn, Meilijson, & Ruppin, 2001; Miller, 1996).

The SFC dictates a well-defined connectivity pattern among neurons in the form of feed-forward connections between pools of neurons. Each of the w_E neurons in a pool receives L connections from neurons in the previous pool, thus creating a chain of pools. Other additional input connections as well as outputs are allowed.

If L is large enough, then a synchronized firing volley of most of the neurons in a pool may form a wave of activity that propagates along the chain (Diesmann, Gewaltig, & Aertsen, 1999). Also if L is large enough, and if the igniting volley is synchronized and strong enough, the waves are stable in the presence of background noise (Diesmann et al., 1999).

To avoid terminological confusion, the feed-forward connectivity schemes are referred to henceforth as *chains*, and the synchronized volley propagating along a chain as a *synfire wave*, or simply a *wave*. A wave can propagate in a synchronized manner along a chain, or it can lose its synchrony, dissolving into the background activity. A wave is said to be *stable* if it remains as a synchronized volley for more than 100ms.

Having introduced all the ingredients – asynchronous activity at the global level,

attractors or spatio-temporal patterns at the assemblies' level, and irregular spiking at the neuronal level – we may state our mission:

> *We seek a balanced network of spiking neurons that can serve as a high capacity memory device.*

To establish the goal, we use BNs of Integrate-and-Fire (IAF) neurons, similar to the model described in (Brunel, 2000). As memory models we use either a HCA or a SFC.

HCAs embedded in a recurrent network were suggested as a model (Amit & Brunel, 1997) of the persistent activity observed in delayed match-to-sample experiments performed in the inferotemporal and in the prefrontal cortex area (Miyashita & Chang, 1988). Within this model the network, in the absence of memory stimulation, exhibits sustained asynchronous activity. After learning, an elevated activity within a HCA is obtained, if a familiar stimulus is presented for a short while. Wang (Wang, 1999) found that in order to realize a stable, low-rate, persistent activity coexisting with a stable resting state, recurrent excitation should be primarily mediated by kinetically slow synapses of the NMDA type. Later, Compte et al (Compte, Brunel, Goldman-Rakic, & Wang, 2000) examined the synaptic mechanisms of selective persistent activity underlying spatial working memory in the prefrontal cortex. Their model reproduces the phenomenology of the occulomotor delayed-response experiment of Funahashi et al (Funahashi, Bruce, & Goldman-Rakic, 1989). Brunel & Wang (Brunel & Wang, 2001) studied the effects of external input and neuromodulation on persistent activity in a working memory model. Additional

reviews on HCA can be found in (Sommer & Wennekers, 2000, 2001).

In (Aviel, Mehring et al., 2003) we studied the embedding of SFCs in a BN of IAF neurons. The main obstacle that we found is the conflict between constraints due to embedded memories on one hand and the AS stability on the other hand. The conflicting demands are not easily resolved. This is in agreement with the study of (Mehring, Hehl, Kubo, Diesmann, & Aertsen, 2003), where a similar setup was used with added topological, cortex like, connectivity. The authors report a fairly small parameter regime where SFC can be embedded and recalled without destabilizing the AS.

None of these publications paid special attention to the question of high capacity. The focus of the present paper is capacity. We load the system with memory patterns, and look for conditions under which we can

(a) maximize the load

(b) keep the asynchronous state stable

(c) recall every memory in a stable manner.

Obtaining maximal capacity is, of course, a desirable goal. Many papers dealt with this kind of question in networks of binary neurons. While far from biological reality, binary neurons lend themselves to analytical examination.

The Hopfield model, for example, is useful because it can be analyzed theoretically. In particular its behavior with respect to memory load is analyzed in (Amit, Gutfreund, & Sompolinsky, 1985), where it is shown that the maximal number of uncorrelated memory patterns, P_{max}, is linear in N, the number of neurons:

$P_{\max} = \alpha_c N$, with α_c = 0.14.

Applying results obtained in a binary neural network to neural networks of spiking neurons is not straightforward. Spiking neurons involve membrane time constants and non-linear resetting which lead to much richer system dynamics. Little is known on capacity of neural networks of spiking neurons.

Sommers and Wennekers (Sommer & Wennekers, 2001) obtained a high capacity limit of HCA in a symmetrically coupled network of 100 Pinsky-Rinzel neurons. An inhibitory loop provided negative feedback leading to dynamic threshold control that can improve capacity. Their network operated in the oscillatory regime, and obtained high capacity. Here we limit ourselves to the asynchronous regime within a BN, which introduces other types of constraints.

The capacity of SFCs has also been investigated: Bienenstock (Bienenstock, 1995) used an *r*-winners-take-all model and applied to it signal-to-noise analysis. Hermann et al (Herrmann, Hertz & Prugel-Bennett, 1995; Hertz, 1999) reduced the IAF model to a binary model so statistical mechanics methods can be used. Both arrived at similar conclusions; again $P_{\max} = \alpha_c N$, but now $\alpha_c \cong 8$. Here P_{max} is the maximal number of pools.

In section 2 our model is presented in detail, in section 3 the scaling of the system is discussed, then in section 4 the results are reported and finally in section 5 we discuss the results.

2. The model

We employ three levels of modeling in our system; on the microscopic level – a neuronal model is required, for which we use an Integrate-and-Fire model. On the intermediate level – we model memory patterns in two different ways, through synfire chains and by Hebbian cell assemblies. On the macroscopic level we use a balanced network and discuss, in the following, its modification into a doubly balanced network.

For simulations we used the SYNOD environment (Diesmann, Gewaltig, & Aertsen, 1995) with the Paranel kernel (Morrison & Diesmann). The neuronal model was integrated with time steps of 0.1ms.

2.1 Single neuron model

Following Lapique (Tuckwell, 1988), we use an Integrate-and-Fire model, in which the *i*-th neuron's membrane potential, $V_i(t)$, obeys the equation:

(1) $$\tau \frac{dV_i(t)}{dt} = -V_i(t) + RI_i(t),$$

where $I_i(t)$ is the synaptic current arriving at the soma and R is the membrane resistance. Spikes as well as post-synaptic currents are modeled by delta functions; hence, the input is written as

(2) $$RI_i(t) = \sum_j \sum_{t_j^f} J_{ij}\delta(t - t_j^f - \tau_{delay}),$$

where the first sum is over different neurons, whereas the second sum represents their spikes arriving at times $t = t_j^f - \tau_{delay}$. t_j^f is the emission time of the f-th spike by neuron j, and τ_{delay} is a transmission delay, which we assume here to be the same for any pair of neurons. The sum is over all neurons that project their output to neuron i, both local and external afferents. The strength of the synapse that neuron j forms on neuron i is J_{ij}.

When $V_i(t)$ reaches the firing threshold θ, an action potential is emitted by neuron i, and after a refractory period τ_{rp}, during which the potential is insensitive to stimulation, the depolarization is reset to V_{reset}.

The following parameters were used in all simulations: The transmission delay τ_{delay} = 1.5ms, the threshold θ = 20mV, the membrane time constant τ = 10ms, the refractory period τ_{rp} = 2.5ms, the resetting potential V_{reset} = 0mV and the membrane resistance $R = 40M\Omega$.

The inhibitory and excitatory neurons have identical parameters.

2.2 Memory model

Two memory models are explored, constructed on the basis of either the *Hebbian Cell Assembly* (HCA) or the *SynFire Chain* (SFC).

HCAs form an associative-memory model in a biologically plausible network. An assembly is a group of w_E randomly chosen excitatory neurons. The assembly is distinguished by a dense inter-assembly connectivity. A neuron in an assembly

receives L connections from other neurons in the assembly. The dense connectivity allows for sustained high firing rate, once an assembly is ignited.

A neuron will typically participate in more than one assembly. In that case, it will have high connectivity with all these assemblies. If two neurons happen to participate together in two assemblies, we assign two different synapses to the connectivity induced by the two memories. All excitatory (inhibitory) synaptic connections are assumed to be of equal strength, J (J_I), and a stronger bond between a pair of neurons is brought about by multiple synapses. Each neuron will be assumed to have a fixed synaptic resource, made of K excitatory (inhibitory) "synaptic units" J (J_I). This constraint on synaptic resources will impose an upper bound on the number of memory patterns one can embed in the network.

SFCs (Abeles, 1991) are feed forward connections among neuronal pools, with a fixed number of excitatory neurons, w_E, in each pool. w_E is also referred to as the "chain width". To wire a SFC, we randomly pick pools of w_E neurons each and connect them in a feed-forward manner, thus obtaining a chain of pools. Converging-diverging connections form a link between two consecutive pools. Similar to HCAs, a neuron will typically participate in more than one pool. In contrast to HCAs, each neuron in a pool receives L specific connections from the previous pool, and not from the pool to which it belongs. If w_E is large enough, then a synchronized firing volley of most of the neurons in a pool may propagate along the chain forming a Synfire wave, a stable wave of activity (Diesmann, Gewaltig, & Aertsen, 1999). The wave signals an activity of a specific memory object. The wave can be (de-) synchronized

with other waves, allowing for a mechanism of dynamically (un-) binding objects (Hayon, 2002).

In these two memory models we have three parameters; w_E – the size of the memory pattern (either an assembly or a pool), L – the inter-pattern connectivity, and P – the number of memories (assemblies or pools).

Ignition of a memory pattern is performed by increasing the external input to all members of the pattern during 5ms. An HCA pattern will be said to be stable if after ignition it manifests a sustained elevated firing rate for at least 100ms. A Synfire wave will be said to be stable if a wave propagates along the chain for at least 100ms. Both patterns have a minimal value of w_E, w_{min}, below which patterns are unstable. The minimum is due to the condition that the firing of the assembly or pool guarantees the continuing firing of the same assembly or the next pool. In case of SFCs, (Aviel, Mehring et al., 2003; Tetzlaff et al., 2003) w has also a maximal value, above which spontaneous emergence of waves occur. This is also likely to be the case for HCAs.

Once wiring of all memory patterns is done, more connections are added such that if an excitatory neuron has less than K excitatory (K_I inhibitory) synapses, random sources are chosen from the network until it reaches exactly K excitatory (K_I inhibitory) synapses. The resulting connectivity is a mixture of random connectivity and ordered connectivity.

2.3 Network model

The excitatory population consists of N_E excitatory neurons, and the inhibitory one consists of $N_I \equiv \gamma N_E$ inhibitory neurons. A sparse connectivity is required, both to adhere as closely as possible to biological values and to induce a source of randomness. In addition to the external input (*K* excitatory afferents), each neuron in the network receives exactly *K* excitatory and $K_I = \gamma K$ inhibitory afferents, from the excitatory and inhibitory populations, respectively. In our simulations we will use $K = \varepsilon N_E$, with $\varepsilon = 0.1$. Our formalism allows other relations as well, e.g. *K=Const* when N_E grows. This will be further elaborated in the Discussion section below.

If a neuron of population *y* (either *E* or *I*) innervates a neuron of population *x* (either *E* or *I*) its synaptic strength J_{xy} is defined as follows:

(3) $$J \equiv J_{xE} = J_0 / \sqrt{K}, \qquad J_I \equiv J_{xI} = -gJ_0 / \sqrt{K_I}$$

where J_0 is a constant. Note that $J_I = -g/\sqrt{\gamma} \cdot J$, hence the constant $g/\sqrt{\gamma}$ is the relative strength of the inhibitory synapses.

A Poisson process with rate $\nu_{ext}K$ simulates the external input. $\nu_{ext} \equiv \nu \cdot \nu_{thre}$, where ν_{thre} is the minimal rate needed to emit a spike within τ milliseconds (on the average) in a neuron that gets balanced input (Brunel, 2000).

The network parameters used in this paper are:

$\gamma = 1/4, \varepsilon = 0.1,$ g=5, J_0=10 and ν=1/20.

The total input to a neuron is, therefore, nearly balanced. There is a small bias toward an excess of excitation, which controls the firing rates.

3. Scaling

Scaling is a crucial issue in modeling of complex systems. It tells us how parameters should be modified (scaled) while changing the number of elements (or the size) of a system. In our case, the "system" is the network with its connectivity matrix, the external input and the single neuron dynamics. The number of elements is the number of neurons or synapses in the network.

Complex systems often show complex behavior that depends on of the parameters of the system. For proper scaling the behavior may only weakly depend on the size of the system. In this regard it makes sense to consider the "thermodynamic limit", where the number of elements goes to infinity. In this section we propose a scaling scheme that allows us to capture the essential behavior of the model for all values of K.

We start by scaling the synaptic weight, J. The proportionality factor of the synaptic weight is typically chosen in model calculations as a function of the number of synapses, K:

(4) $$J = J_0 / K^{\beta}$$

where β gets values 0, 0.5 or 1.

The mean field approach using β=1 leads to weak coupling in the high K regime. Examples of such studies are the Hopfield model (Hopfield, 1982) and phase variable analysis (Hansel, Mato, & Munier, 1995).

Constant synaptic strength, i.e. β=0, was also used in the literature. Brunel (Brunel, 2000), for example, used fixed synapses together with the assumption that the constant post-synaptic potential (PSP) contribution is small relative to the distance between threshold and resting potential. Also in this regime the synaptic coupling is weak, and one can model the barrage of incoming PSPs as Gaussian noise. Since the contribution of each PSP is small, the overall change in the membrane potential is smooth and diffusion approximation can be used.

Following (van Vreeswijk & Sompolinsky, 1998), we propose using β=0.5. The reason underlying this choice comes from requiring the firing rate of each neuron to be independent of the number of synapses, K. The same logic applies to the inhibitory synaptic strength J_I. Choosing β=0.5, we obtain Eq. (3).

Our system operates under balanced conditions, i.e. the mean input to a neuron is near zero and fluctuations drive the spiking process. Let h_x be the field generated by synapses of population x (x=E or I) felt by some neuron, and let us assume that the system is in an asynchronous state (AS). Under these conditions all synaptic inputs can be modeled by a Poisson process with rate v. To simplify calculations, we assume that both populations fire at the same rate v. Using (3), the mean and variance can be described as follows:

(5) $$\mu \equiv h_E - h_I = v \cdot \left(JK - J_I K_I \right) = J \cdot v \cdot K \left(1 - g\sqrt{\gamma} \right) = J_0 \cdot v \cdot \sqrt{K} \left(1 - g\sqrt{\gamma} \right)$$

(6) $$\sigma^2 = \mathrm{var}(h_E) + \mathrm{var}(h_I) = \nu \cdot \left(J^2 K + J_I^2 K_I\right) = J^2 \cdot \nu \cdot K\left(1 + g^2\right) = J_o^2 \nu\left(1 + g^2\right)$$

From Eq. (5), we see that $g\sqrt{\gamma}$ controls the balance between the two populations. The mean of the field is zero if $g\sqrt{\gamma} = 1$, and therefore independent of *K*. But also the variance is independent of *K*. This would not have been the case for β=0 or 1.

On average, therefore, an input of $\sqrt{K}$ excitatory PSPs produces a spike. The minimum rate of external input that is needed in order to emit a spike within τ ms (on average) in a neuron that does not get other inputs (Brunel, 2000) is defined to be $\nu_{thre} \equiv \theta / \left(\tau \cdot J \cdot \sqrt{K}\right) = \theta / \left(\tau \cdot J_0\right)$. Again, ν_{thre} is independent of *K*.

In our simulations $g > 1/\sqrt{\gamma}$, so there is a potential excess of inhibition, but there is also an additional excitatory external input. Even if the total input is not exactly balanced, the feedback between the population, which is on the order of $\sqrt{K}$ according to Eq. (5), leads to an appropriate change of the firing rates that stabilize the AS (van Vreeswijk & Sompolinsky, 1998).

Another hint for setting β to 0.5 comes from our previous study (Aviel, Mehring et al., 2003), where we demonstrated that w_E has to obey $w_E = Const \cdot \sqrt{K}$. In this case an input from w_E synapses, Jw_E, should lead to firing with high probability. It means that $Jw_E \cong O(1)$, i.e. *J* is proportional to $1/\sqrt{K}$.

As a consequence, we set the pattern size $w_E = C_w \sqrt{K}$, and the intra-pattern connectivity $L = C_L \sqrt{K}$, where C_w and C_L are the scaling pre-factors. The scaling of

the external rate and the synaptic weights is given by $v_{ext} \equiv v \cdot v_{thre} = v \cdot \theta / (\tau \cdot J_0)$ and $J \equiv J_0 / \sqrt{K}$ respectively, as discussed in chapter "*The model*" above.

We use the following values:

Parameter	Values	
	HCA	SFC
J_0	10	10
C_w	3.3	4
C_L	$0.75C_w$	C_w
v	0.05	0.05

The usefulness of the scaling of pattern connectivity and external rate will become evident from the discussions below.

In the next section we discuss properties of the resulting balanced network.

4. Results

4.1 Balanced Network

In a previous paper (Aviel, Mehring et al., 2003) we studied the embedding of SFCs in the synaptic connectivity matrix of a balanced network. There, as we do here, we enforced two constraints: the asynchronous state (AS) has to be the stable background mode of the system, and the synfire wave has to propagate on top of it in a stable manner. Based on simplified models and simulations, we concluded that these two constraints could be met if K is large enough ($K > Const \cdot w_{\min}^2$). In this paper, we

take the network load $\alpha \equiv P/N_E$ into consideration, using synaptic scaling. This allows further refinement of our statements.

Brunel's network (Brunel, 2000) serves here as the starting point, in which we embed HCA or SFC patterns. But apart from introducing ordered patterns in its connectivity matrix, we also use a particular scaling procedure. Synaptic weights, instead of being constants, are now scaled like $1/\sqrt{K}$ (van Vreeswijk & Sompolinsky, 1998). This, in turn, leads to $v_{thre} = \theta/(\tau \cdot J_0)$ = *Const*. We verified by simulations that indeed the mean firing rate (in the AS) is linearly related to v_{ext} and is only weakly dependent on N_E.

The analysis of the Appendix in (Aviel, Mehring et al., 2003) supports square-root-scaling. We have shown there that w_E is limited by two constraints:

(7) $$w_{min} < w_E < w_{max}$$

where $w_{max} \equiv C_b\sqrt{K}$. The upper bound is due to requiring stability of the AS, and the lower bound is posed by wave stability demand. We can estimate w_{min} as the mean number of excitatory PSPs needed to evoke an action potential with high probability. Assuming a normal distribution of the membrane potential, w_{min} can be approximated by twice the distance between threshold and $\langle V \rangle$ in units of *J*, where $\langle V \rangle$ is the mean membrane potential and *J* is the synaptic unit:

(8) $$w_{\min} = 2\frac{\theta - \langle V \rangle}{J}.$$

Substituting Eq. (8) in Eq. (7), we notice that w_E is sandwiched between an upper bound that is of order $\sqrt{K}$ and a lower bound that is of order K^{β}. Unless $\beta \le 0.5$, w_E has no valid value in the limit of large *K*.

By choosing $\beta = 0.5$ and $w_{\min} \equiv C_a\sqrt{K}$ accordingly, Eq (7) can be rewritten as:

(9) $$C_a < C_w < C_b$$

Under these conditions, we expect the SFC to be stable in our system regardless of the number of synapses *K*.

In Figure 1a, we can see a wave propagating along a chain of 750 pools in a BN of 15,000 excitatory neurons. If, however, the chain width exceeds a critical value, or if the network load exceeds capacity, global oscillations appear. In Figure 1b, we increase the number of pools embedded in the network. This leads to emergence of global oscillations after wave ignition.

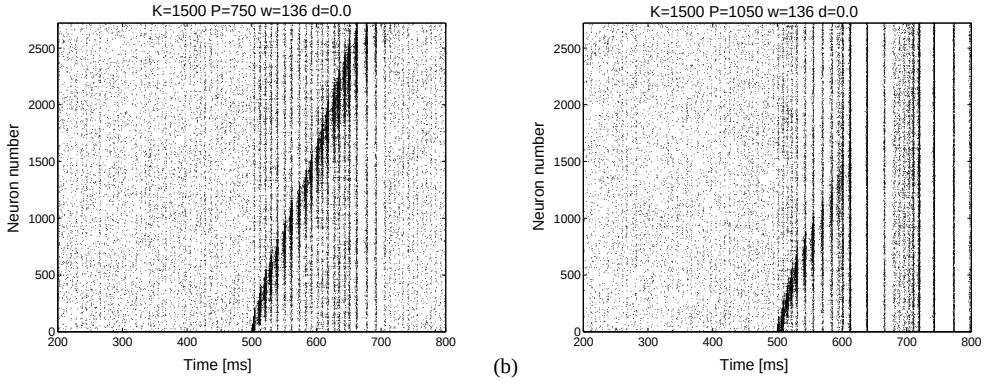


Figure 1: Raster plots of a SFC embedded in a BN with *K*=1500. Left: α = 0.05, Right: α = 0.07. Neurons that participate in more than one pool may appear more than once on the raster plot, whose y-axis is ordered according to pools,

and represents every second neuron in each pool. The first 40 pools of the SFC are presented on the y-axis.

We also embedded HCAs according to the protocol described in "*The model*" section. In a lightly loaded system, as in Figure 2a, an evoked assembly sustained its activity for hundreds of milliseconds without provoking strong global oscillations.

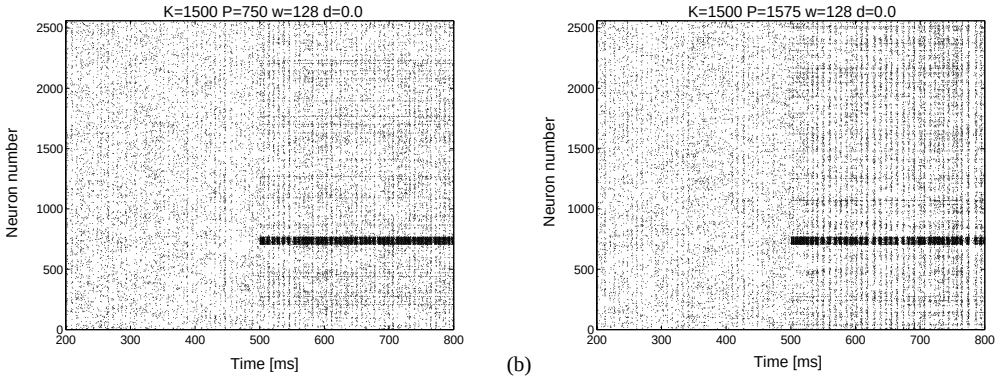


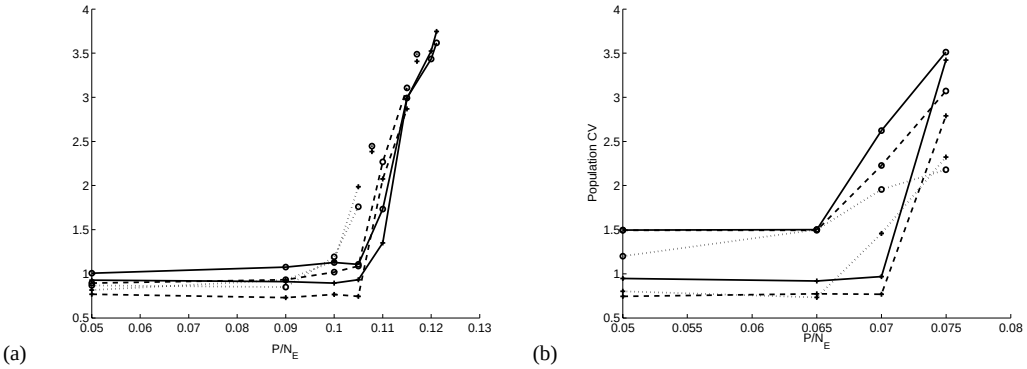
Figure 2:Raster plots of HCAs embedded in a BN with *K*=1500. Left: α = 0.05, Right: α = 0.105. The twelfth pattern is externally ignited at time t=500ms for 5ms. Neurons that participate in more than one HCA may appear more than once on the raster plot, whose y-axis is ordered according to HCAs, and represents every second neuron in each pattern. 40 HCAs are presented on the y-axis.

As in the SFC case, Figure 2b shows that the global oscillations become prominent as the load increases. A surprisingly high capacity is obtained in both cases.

In order to quantify the emergence of global oscillations, the values of the population rate, i.e. number of spikes of a population in one millisecond, are considered. An average and standard deviation (std) of these values are taken over a time window of 300 ms, in order to compute a population's coefficient of variance (CV). The CV is the std divided by the mean, and in our case it measures the amount of synchrony of the population activity. CV near one means a std that is close to the mean, which is in

agreement with an AS. High CV is a result of high std, which implies global synchronous (a-periodic) oscillations in our model. Low CV (<0.5) signifies high firing rates. Low CV will become apparent in the next section, where the mild oscillations in the background activity during wave propagation will be addressed. In Figure 3, the CV of the population rate is plotted as a function of network load, P/N_E, for various network sizes. We present statistics for two cases: pre- and post- pattern ignition. In the former, the time window is taken prior to pattern ignition (time 200 to 500ms) and in the latter, the time window is taken post ignition (500 to 800ms).

All curves exhibit transitions near critical points. We define the critical point of the *network load* by P_c, and we define the *capacity* of the network as $\alpha_c \equiv P_c / N_E$. The capacity obtained here is 0.1 and 0.065 for HCA and SFC respectively.



(a) (b)

Figure 3: BN population CV as a function of α, for various values of *K*. Left: HCAs are embedded. Right: A SFC is embedded. Post- and pre- ignition curves are indicated by 'o' and '+' markers respectively. In both cases statistics was gathered from the 300ms pre- and post- ignition. Curves for three values of *K* (500 (dotted curve), 1000 (dashed curve) and 1500 (full line)) are superimposed.

The transitions to the oscillatory mode occur near the same α value, α_c, (0.1 for HCAs, 0.07 for SFCs pre-ignition and 0.65 for SFC post-ignition) for all *K* values.

One should note an important difference between HCAs and SFCs. Whereas when HCAs are embedded, the pre- and post-ignition curves are indistinguishable, the SFCs shows a significantly higher network CV after their ignition.

4.2 Doubly balanced network

In the background activity of Figure 1, oscillation can be seen during wave propagation. In this section, we provide a mechanism that gets rid of this problem. The solution we arrived at involves architectural modifications to the original SFC structure.

Let us attach a pool of randomly chosen inhibitory neurons, a *shadow pool*, to each excitatory pool in a synfire chain. A neuron in an excitatory pool projects its output, not only to the next pool in the chain, but also to all neurons in its shadow pool. A neuron in the shadow pool, does not project its output in any ordered manner, but diffuses its output randomly to the rest of the network, as in a completely random network. Similar connectivity, but for different reasons, was suggested in (Hayon, 2002; Mehring, Hehl, Kubo, Diesmann, & Aertsen, 2003). A sketch of the connectivity scheme of a modified synfire chain is shown in Figure 4.

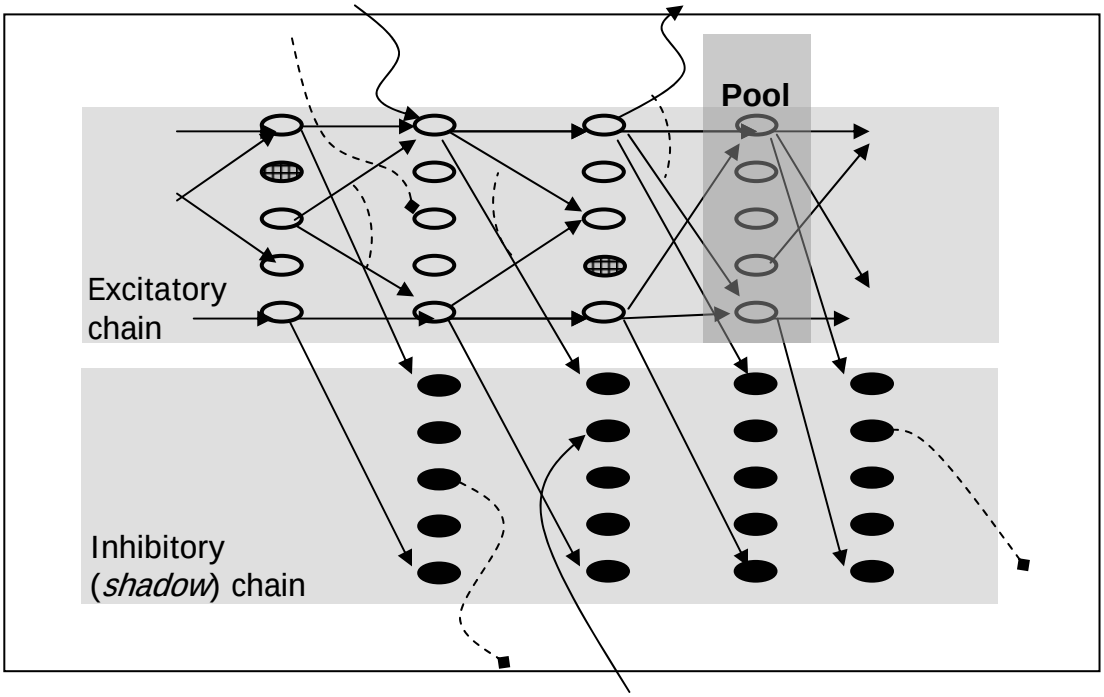


Figure 4: A modified synfire chain. Excitatory neurons (open circles) construct the "traditional" excitatory chain, characterized by the full feed-forward connections. Inhibitory neurons (filled circles) form the "shadow" pools. Each shadow pool receives excitation from its associated excitatory pool. A neuron can participate more than once in the chain. Additional excitatory (solid line) and inhibitory (dashed lines) inputs and outputs are allowed (marked with curved lines).

These inhibitory pools do not carry specific information down the chain, as is the case for the excitatory pools, but rather echo a synchronized activity of their attached excitatory pool. The role of the 'shadow pools' is to guarantee a correct amount of inhibition during the propagation of the wave. The excess of converging connections in the network, due to the embedded chain, induces excess of correlated excitation, which may in turn, lead to global excitation (Aviel, Mehring et al., 2003). The sole

purpose of the shadow pools is to cancel that excess of excitation. An analogous shadow pattern is associated with every HCA in an HCA model.

The size of a shadow pattern is defined as $w_I \equiv \tilde{d} w_E$. This leads to the factor *d*, representing the relative strength of inhibitory to excitatory currents, due to a pattern or pool, affecting a neuron that is connected to both:

$$d \equiv \frac{-J_I w_I}{J w_E} = \frac{g J_0 \sqrt{K} \tilde{d}}{J_0 \sqrt{K_I}} = \frac{g \tilde{d}}{\sqrt{\gamma}}, \quad x \in \{E, I\}.$$

In other words $w_I = d\left(\sqrt{\gamma}/g\right) w_E$. In the simulations reported below we use *d*=2 for HCAs and *d*=1 for SFCs, i.e. $w_I = 0.2 \cdot w_E$ and $w_I = 0.1 \cdot w_E$ respectively. These values were chosen after a search for optimal *d* values that lead to the smoothest background behavior after pattern ignition.

We refer to this new type of network as a *Doubly Balanced Network* (DBN). To avoid confusion we will refer to the previous case, *d*=0, as a *Singly Balanced Network* (SBN).

In Figure 5, we repeat the simulations of Figure 3, only this time we simulate DBNs. Since the pre- and post- ignition curves overlap in the DBN, we present only the pre-ignition curve. The two cases are plotted on the same scale, so that the difference in capacity can be appreciated.

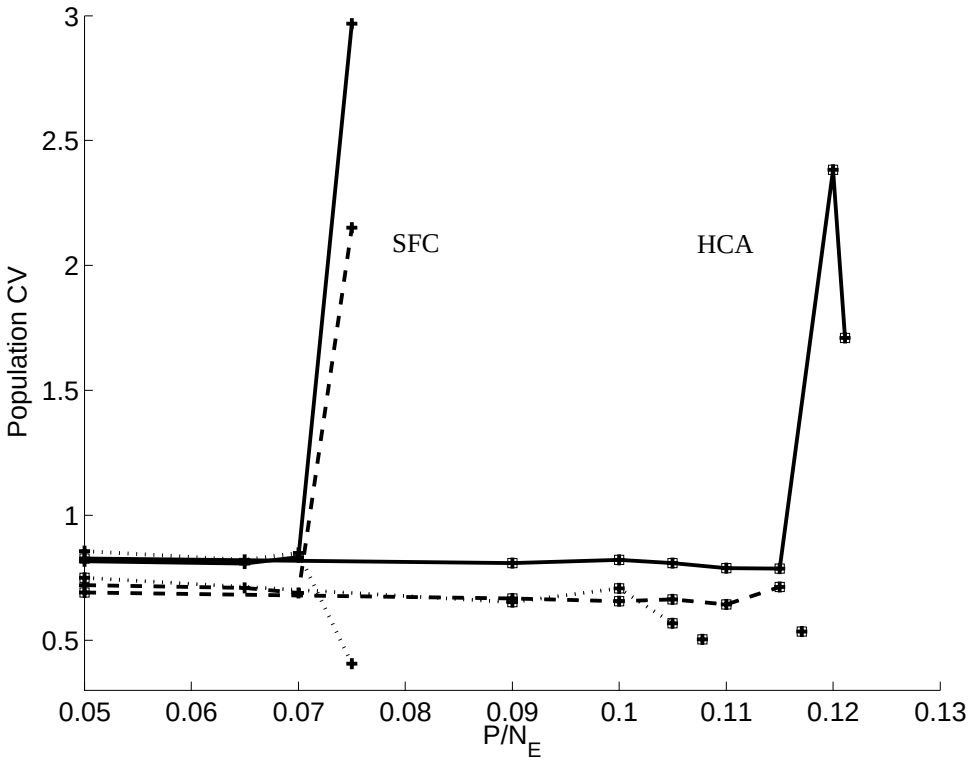


Figure 5: DBN pre-ignition population CV as a function of α , for three values of *K* (500 (dotted curve), 1000 (dashed curve) and 1500 (full line)). The transition in SFC ('+') occurs before that of HCA (square marks), and is independent of *K*. Capacity is 0.115 and 0.07 for HCA and SFC respectively.

As evident, the capacity of the DBN is superior to that of the SBN in both cases.

4.3 Capacity

In this section we show that the maximal number of patterns that can be embedded in the network is limited by combinatorial considerations of synaptic resources, and is proportional to *N*. That such a limit has to exist follows from simple counting of all excitatory synapses. On one hand we know there are *KN* such synapses available. On

the other we know that each HCA, or pair of consecutive pools in a SFC, use up w^2 of them. Hence $Pw^2 < KN$, or $\alpha = \frac{P}{N} < \frac{K}{w^2} = \frac{1}{C_w^2}$.

A more careful analysis can take care of the accounting of synapses in the way they are assigned within our model. We divide the K excitatory synapses of each neuron into chunks of L synapses. A neuron of population x (E or I) can participate in at most $m \equiv \lfloor K/L \rfloor$ patterns due to synaptic constraints. The total number of chunks, $N_x m$, sets an upper bound on the number of patterns, since embedding P patterns requires $w_x P$ chunks.

Hence we get:

(10) $$w_x P_{\max} \le m \cdot N_x$$

where P_{max} is the maximal number of patterns or pools. Next, defining $\alpha_x \equiv \frac{P_{\max}}{N_x}$, we find that

(11) $$\alpha_x \le \frac{m}{w_x} = \frac{\lfloor K/C_L\sqrt{K} \rfloor}{w_x}$$

To leading order in N_E this turns into

(12) $$\alpha_x N_x = \frac{\lfloor K/C_L\sqrt{K} \rfloor}{D_x C_w \sqrt{K}} N_E = (C_w C_L D_x)^{-1} N_E - O\left(\sqrt{N_E}\right)$$

where $D_x \equiv d/\left(g\sqrt{\gamma}\right)$ if x=I, or 1 for x=E.

Thus we conclude that synaptic combinatorial considerations lead to a maximal number of patterns P_{max}. If D_I>1, then $\gamma\alpha_I < \alpha_E$ and the inhibitory neurons set the maximum value of P. Otherwise, if D_I<=1, the excitatory neurons set P_{max}.

(13) $$P_{\max} = \alpha_{\max} N_E$$

$$\alpha_{\max} \equiv \min\left(\left(C_w C_L\right)^{-1}, \left(C_w C_L D_I\right)^{-1}\right)$$

In our DBN, as well as the SBN where D_I<1, hence the excitatory neurons determine the limit to be $P_{\max} = \left(C_w C_L\right)^{-1} N_E$. Substituting the parameters of our HCA in Eq. (11), we get $P_{\max} \cong 0.12 N_E$.

In

Figure 6, P_c of HCAs is plotted as a function of N_E. A linear relation can be observed up to values of N_E as high as 45,000. Furthermore, the proximity of the dynamical capacity to the combinatorial one is evident. Results for cases of SBN are also plotted for comparison. As already noted before, their dynamical capacity is lower than that of DBNs.

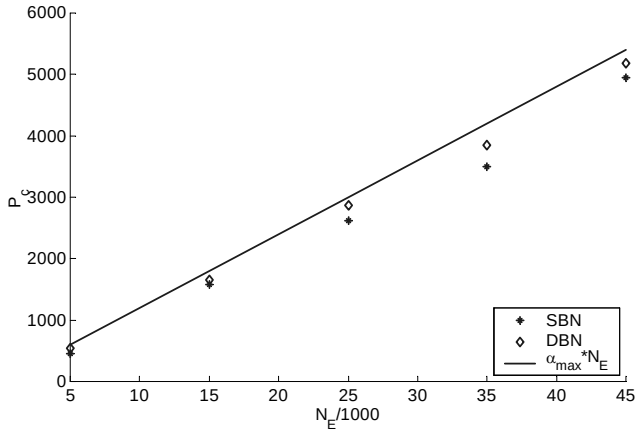


Figure 6: Combinatorial and dynamical capacities of HCAs as function of N_E for DBN ('o') and SBN ('*'). The solid line is $\alpha_{\max} \cdot N_E$.

4.4 Allowed phase space

We are now in a position to return to the issues of scaling and inquire about the dependence of C_b of Eq (9) on network load. We fix α and progressively increase w_E in an SBN until global oscillation appears. As in Figure 3 and Figure 5, the curves display a critical point, w_c. This criticality was discussed in (Aviel, Mehring et al., 2003), where we found $w_c = C_b\sqrt{K}$. In Figure 7, we repeat these simulations for various values of α and K. Each simulation (C_w, α, K) is repeated five times to give an indication of the trial error. We also plot a crude estimate, based on simulations, of C_a and the combinatorial capacity upper limit, $\alpha_{\max}$, according to Eq (13).

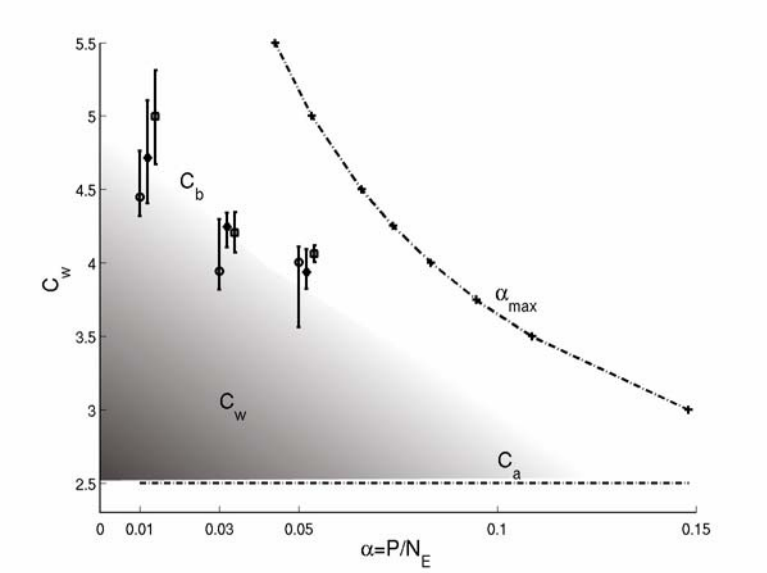


Figure 7: The allowed region of C_w, the shaded gray area, is limited by C_a from below and C_b from above. The latter, represented by markers with error bars, is estimated from numerical simulations for three different values of K. The markers are slightly shifted with respect to each other (K=500 (circles) to the left and 1500 (squares) to the right) to allow easy interpretation. These values lie below the constraint $C_W C_L \alpha < 1$ that led us (Eq. (13)) to the combinatorial upper bound $\alpha_{\max}$.

In the phase portrait illustrated in Figure 7, values of C_w and α that are above the combinatorial upper bound (dashed-dotted line) are forbidden due to the synaptic constraints of our model. Combinations of C_w and α that are below the combinatorial upper bound but outside the shaded area lead to instability of the AS. Below C_a pattern activities dissolve into the background, hence they are unstable. The only regime that allows stability of both the AS and the pattern activity is the shaded region in the figure.

As the network load increases, C_b and C_a approach each other, shrinking the allowed regime of C_w, in which embedding is possible.

In a DBN, the C_b values are higher, closer to the combinatorial upper bound. This enlarges the stability regime of the assemblies in a BN to the extent that it enables realizing the combinatorial upper bound in some cases (Aviel, Horn, & Abeles, 2003a).

When embedding SFCs in the BN, the phase portrait stays qualitatively similar but C_b values are lower, leading to lower dynamical capacity than in the case of HCAs.

5 Discussion

In this paper we studied embedding of memory patterns in a balanced network of IAF neurons. Two types of memories were used, Hebbian cell assemblies (HCAs) (Hebb, 1949) and synfire chains (SFCs) (Abeles, 1991). We propose a scaling behavior of synaptic weights and other parameters that render our model invariant to changes in K, the synaptic size of the network. This square-root-scaling allows for high capacity of both HCA and SFC.

We emphasized the scaling of variables with K, but it is usually N_E that one varies. Using the relation $K = \varepsilon N_E$ these two variables are linearly related. Note however that K may also be assumed to be constant, or vary at a different rate than N_E, and our results will still be valid as long as $K/N_E << 1$. In particular, the increase of the maximal number of patterns P_{max} with N_E is valid even if K is constant.

We distinguished between the "traditional" balanced network, which we termed *Singly Balanced Network (SBN)*, and our new *Doubly Balanced Network (DBN)*. In a DBN, memory patterns embedded in the excitatory-to-excitatory part of the connectivity matrix project also onto their *shadow patterns*. The shadow patterns are merely random pools of inhibitory neurons that receive inputs from their associated excitatory patterns. Counteracting emerged excitatory correlations by inhibitory ones imposes another type of balance. Hence it is a doubly balanced network. We have shown in (Aviel, Horn, & Abeles, 2003b) that an optimal choice of d, the ratio of inhibitory to excitatory correlation currents, achieves the desired effect. If d is too small, background oscillations appear, as was the case of the SBN. If d is too large, the induced inhibition kills the synfire waves. The proper value, around d=1, allows for optimal performance. The DBNs have the advantage that their background activity during memory recall stays asynchronous even for high memory loads.

Another indication of the stabilizing affect of the shadow patterns is given in (Aviel, Horn et al., 2003b), where it is shown that if a strong synchronized input is used to ignite a wave in a SBN, rather than a step current, then global oscillations are inevitable. With a DBN on the other hand, waves are possible on top of asynchronous background activity.

Introducing double-balance is only one way of stabilizing the AS. Other ways involve more sophisticated neuronal models (Wang, 1999), or introducing variability through non-homogeneous neuronal population or through different connectivity schemes. For example, models, such as (Amit & Brunel, 1997; van Vreeswijk & Sompolinsky, 1998), used variable number of synapses on each neuron. This leads to a broad

distribution of firing rates across the population, which in turn, introduces variability that helps to stabilize the AS. In our work, however, each neuron receives exactly the same number of synapses (Brunel, 2000). Here we suggest a stabilizing mechanism that is directly related to the cause of the instability. Clearly, additional stabilizing mechanisms, such as variable number of synapses, variable transmission delays, or more sophisticated neuronal models will only help to increase the AS's stability.

A capacity limit $\alpha_c = 0.12$ of HCAs calls for comparison with analytic bounds obtained for binary models, like $\alpha_c = 0.14$ in the Hopfield model (Amit et al., 1985). The comparison must be done with care, as the two types of neuronal models are qualitatively different. Hertz (Hertz, 1999) has argued that a capacity limit obtained in a network of integrate-and-fire neurons should be multiplied by $\tau/2$ to compare it with a network of binary neurons. Hence the $\alpha_c = 0.115$ obtained here, is equivalent to $\alpha_c = 0.57$ in a binary model. It is not surprising that the last number is higher than 0.14, since our model's memory patterns are sparse, as, e.g. in the Tsodyks-Feigelman (Tsodyks & Feigelman, 1988) model, where larger capacities were achieved.

An example of an IAF network with high capacity was demonstrated by Sommers & Wennekers (Sommer & Wennekers, 2001). However, their network was not balanced, and memories retrieved from oscillatory modes. To the best of our knowledge, the model presented here is the first to obtain such a high capacity in an asynchronous mode of a network of spiking neurons.

Finally let us touch some of the interesting open issues.

The firings of the neurons in the ignited pool are of much higher rate and are higher regularity than those reported by delayed match-to-sample experiments (Miyashita & Chang, 1988). In our model, neurons in the ignited pool receive much more excitatory than inhibitory input and therefore do not operate in the balanced-input regime anymore. Ways to reduce the ignited state's firing rate were suggested before (Amit & Brunel, 1997; Wang, 1999). It will be interesting to incorporate these studies with the DBN approach in order to better fit experimental data.

In this work we recall only one pattern at a time. We require the ignited pattern to be stable (i.e. sustained high firing rate for HCA or sustained propagation of Synfire wave) for at least 100ms without provoking global oscillations. The next step is to ask how many concurrent patterns can be successfully recalled. This is yet another type of capacity. Preliminary results show that three HCAs can be recalled in a K=1500 model, one of which will survive the competition and exhibit sustained activity for hundreds of milliseconds.

Our model uses binary synapses; a synapse either exists, with strength of one synaptic unit, or is absent. We strengthen bonds if the two neurons participate together in more memories. This construction leads to the combinatorial maximum capacity. One may speculate that different synaptic arrangements, e.g. along the spirit of the Willshaw model (Willshaw, Buneman, & Longuet-Higgins, 1969), will lead to even higher dynamical capacity.

Acknowledgements

This work was supported in part by grants from GIF and DIP.

References

Abeles, M. (1982). *Local cortical circuits - An electrophysiological study*. Berlin: Springer-Verlag.

Abeles, M. (1991). *Corticonics*: Cambridge University Press.

Amit, D., J. (1989). *Modeling Brain Function: The world of attractor neural networks*. New York: Cambridge University Press.

Amit, D., J, & Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex, 7*, 237-252.

Amit, D., J, Gutfreund, H., & Sompolinsky, H. (1985). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters, 55*, 1530-1533.

Aviel, Y., Horn, D., & Abeles, M. (2003a). The doubly balanced network of spiking neurons: a memory model with high capacity. *NIPS 2003*.

Aviel, Y., Horn, D., & Abeles, M. (2003b). Synfire waves in small balanced networks. *Neurocomputing*, Volumes 58-60, Pages 123-127.

Aviel, Y., Mehring, C., Horn, D., & Abeles, M. (2003). On embedding synfire chains in a balanced network. *Neural Computation, 15*(6), 1321-1340.

Bienenstock, E. (1995). A model of the neocortex. *Network: Computation in Neural Systems, 6*, 179-224.

Braitenberg, V. (1978). Cell assemblies in the cerebral cortex. In: Heim, R., Palm, G. (eds.): *Theoretical Approaches to Complex Systems. Lecture Notes in Biomathematics*, Vol. 21, pp. 171-188. Berlin, Springer.

Braitenberg, V., & Schuz, A. (1991). *The anatomy of the cortex. Statistics and geometry.* Berlin: Springer.

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J Comput Neurosci, 8*(3), 183-208.

Brunel, N., & Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience, 11*, 63-85.

Compte, A., Brunel, N., Goldman-Rakic, P., S., & Wang, X.-J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex, 10*, 910-923.

Diesmann, M., Gewaltig, M. O., & Aertsen, A. (1995). *SYNOD: An environment for neural systems simulations*. Rehovot, Israel:Weizmann Institue of Science.

Diesmann, M., Gewaltig, M. O., & Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature, 402*(6761), 529-533.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P., S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophsiology, 61*, 331-349.

Gerstein, G. L., & Mandelbrot, B. (1964). Random walk models for the spike activity of a single neuron. *Biophysical Journal, 4*, 41-68.

Gerstner, W., & van Hemmen, L. (1992). Associative memory in a network of 'spiking' neurons. *Network, 3*, 139-164.

Hansel, D., Mato, G., & Munier, C. (1995). Synchronization in excitatory neural networks. *Neural Computation, 7*, 307.

Hayon, G. (2002). *Modeling compositionality in biologival neural networks by dynamic binding of synfire chains.* Unpublished PhD dissertation, Hebrew University, Jerusalem.

Hebb, D. O. (1949). *The organization of behaviour*. New York: Wiley.

Herrmann, M., Hertz, J. A. & Prugel-Bennett, A. (1995). Analysis of synfire chains *Network: Comput. Neural Syst*. *6* No 3, 403-414

Hertz, J. A. (1999). Modeling synfire networks. In G. Burdet & P. Combe & O. Parodi (Eds.), *NEURONAL INFORMATION PROCESSING - From Biological Data to Modelling and Application*.

Hopfield, J. J. (1982). Neural networks and physical systems woth emergant collective computational abilities. *PNAS, 79*, 2554-2558.

Horn, D., Sagi, D., & Usher, M. (1991). Segmentation, Binding and Illusory Conjunctions. *Neural Computation, 3*, 510-525.

Huang, P. E., & Stevens, F. C. (1997). Estimating the distribution of Synaptic Reliabilities. *Journal of Neurophsiology, 78*(6), 2870-2880.

Izhikevich, E. M., Gally, J., A, & Edelman, G., M. (2004). Spike-timing dynamics of neuronal groups. *Cerebral Cortex*.

Levy, N., Horn, D., Meilijson, I., & Ruppin, E. (2001). Distributed synchrony in cell assembly of spiking neurons. *Neural Networks, 14*, 815-824.

Mehring, C., Hehl, U., Kubo, M., Diesmann, M., & Aertsen, A. (2003). Activity dynamics and propagation of synchronous spiking in loacally connected random networks. *Biological Cybernetics*.

Miller, R. (1996). Neural assemblies and laminar interactions in the cerebral cortex. *Bilogical Cybernetics, 75*(3), 253-261.

Miyashita, Y., & Chang, H., S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331, 68-70.

Morrison, A., & Diesmann, M. Distributed simulation of large biological neural networks. *In preparation*.

Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Curr Opin Neurobiol, 4*(4), 569-579.

Sommer, F., T, & Wennekers, T. (2000). Modeling studies on the computaional function of fast temporal structure in cortical circuit activity. *Journal of Physiology Paris, 94*(5-6), 473-488.

Sommer, F., T, & Wennekers, T. (2001). Associative memory in networks of spiking neurons. *Neural Networks, 14*(6-7), 825-834.

Tetzlaff, T., Buschermöhle, M., Geisel, T. & Diesmann, M. (2003) The spread of rate and correlation in stationary cortical networks, *Neurocomputing,* Volumes 52-54, Pages 949-954.

Treves, A. (1990). Greaded-response neurons and information encodings in autoassociative memories. *Physical Review A, 42*(4), 2418.

Tsodyks, M. V., & Feigelman, M. V. (1988). The enhanced storage capacity in neural networks with low activity level. *Europhysics Letters, 6*(2), 101-105.

van Vreeswijk, C., & Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Comput, 10*(6), 1321-1371.

von der Malsburg, C., & Schneider, W. (1986). A neural cocktail party processor. *Biological Cybernetics, 54*, 29-40.

Wang, X.-J. (1999). Synaptic basis of cortical persistent activity: The importance of NMDA receptors to working memory. *The Journal of Neuroscience, 19*(21), 9587-9607.

Willshaw, D. J., Buneman, O. P., & Longuet-Higgins, H. C. (1969). Non-holographic associative memory. *Nature (London), 222*, 960-962.