Learning Short Synfire Chains by Self-Organization

John Hertz and Adam Prügel-Bennett
NORDITA, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark.

Abstract. A model of cortical neurons capable of sustaining a low level of spontaneous activity is investigated. Without learning the activity of the network is chaotic. We report on attempts to learn synfire chains in this type of network by introducing a Hebbian learning mechanism and exciting a small set of neurons at random intervals. We discuss the types of instabilities that can arise and prevent the formation of long synfire chains and also discuss various biologically plausible mechanisms which to some extent cure these instabilities.

1. Introduction

Synfire chains have been proposed as a mechanism for neural information processing in the cortex [1]. They have received considerable recent interest as they provide an explanation for spatio-temporal patterns of firing discovered in data from multi-cellular recordings [2, 3]. Synfire chains are made up of pools consisting of a small number of neurons linked together through feed-forward connections so that a volley of activity can propagate from one pool to the next to form a chain of activity. Because the neurons in each pool receive a similar set of inputs they tend to fire together, thus providing a mechanism for the precise timing observed in the recording data. The pools are defined only through the couplings between the neurons so that one neuron may participate in many different pools. Furthermore, as the pools consist of only a small fraction of the total number of neurons, several chains can be active simultaneously without unduly interfering with each other.

The static analysis of synfire chains, that is, with fixed coupling weights, has been carried out by various authors [1, 4, 5]. They show that if the couplings are appropriately chosen, volleys of activity will propagate stably along the chain with precise temporal timing and that the number of pools that can be stored is realistic for computational purposes. In this paper we investigate whether synfire chains can be formed through a biologically plausible self-organizing mechanism. Suggestions that this maybe possible have been reported in reference [4] following the Ph. D. thesis of R. Doursat. In this paper we develop the more realistic models of reference [5] where integrate-and-fire neurons are considered. However, we consider here only parallel updating. In references [4] and [5] it was argued that near synchronous volleys of activity arise naturally because of the feed-in structure of the synaptic couplings and therefore it was justified to consider parallel updating in the static analysis of the capacity. However, whether this arguments holds for the much more complicated task of learning synfire chains requires further investigation.

The paper is divided into two main sections. In the first part we investigate the self-sustained spontaneous activity and the condition under which this activity is stable. This we investigate both analytically and numerically. In the second section we report some preliminary results when we "switch on" Hebbian learning. We briefly discuss the types of instabilities we have observed and various mechanisms that can be introduced to overcome them. Our philosophy is to see whether it is possible to construct a biologically plausible model capable of learning synfire chains and to see what elements appear necessary for this to be achieved.

2. Self-Sustained Spontaneous Activity

We investigate a network capable of sustained spontaneous random activity as a basis from which to learn synfire chains. A similar state has been analyzed in detail by Amit and Brunel [6], who have call this a "vacuum" state. They give a detailed discussion of the plausibility of such a state in cortex. Since our main concern is with learning which we will investigate predominantly in the context of simulation we will consider a cruder analysis than that given in [6] which can be directly applied to our simulation. However, the conclusions are very similar to those found in reference [6].

We consider a set of neurons which are updated in parallel. The time step can be regarded as a crude model for the time delay between a neuron firing and the other neurons receiving the spike — for orientation this time can be taken as 1ms. The cells fire if the post-synaptic potential, $V_i(t)$ — where $i = 1, \ldots, N$ label the neurons —, exceeds a threshold θ

$$S_i(t) = \Theta\left(V_i(t) - \theta\right). \tag{1}$$

Once the neuron fires its post-synaptic potential is reset to some resting potential and may remain there for some refractory period before it starts integrating its inputs again. We assume that the neuron acts as a leaky integrator which integrates the instantaneous fields

$$h_i(t) = \sum_j J_{ij} S_j(t), \tag{2}$$

where J_{ij} is an excitatory (i. e. non-negative) coupling from neuron j to neuron i. There may not be a complete connectivity, in which case some of the couplings will be zero. In the simulations we have performed we usually take the average connectivity to be 0.25, although the final results are not very sensitive to what value we chose. Initially the couplings are assumed to be drawn from some random distribution with mean 1 and variance σ_J^2. For example, they may be distributed according to

$$P(J_{ij}) = \frac{2}{\pi} \mathrm{e}^{-J_{ij}^2/\pi}\, \Theta(J_{ij}), \tag{3}$$

in which case $\sigma_J^2 = \pi/2 - 1$, or they may be uniformly distributed between 0 and 2, in which case $\sigma_J^2 = 1/3$.

To maintain a stable background activity it is necessary to include some form of inhibition which grows "faster" than the average activity. Both subtractive inhibition and shunting (multiplicative) inhibition can stabilize the spontaneous activity, provided they are suitably non-linear. For simplicity we consider shunting inhibition of the form

$$V_i(t+1) = \big(1 - \Gamma_0 - \eta\, Q(t)\big)\ \big(V_i(t) + h_i(t)\big)\,, \tag{4}$$

where $Q(t)$ is the total activity of the network

$$Q(t) = \sum_j S_j(t). \tag{5}$$

Physiologically the $1-\Gamma_0$ term can be regarded as the usual passive decay term while the term $-\eta\, Q(t)$ arises from inhibitory synapses acting at the soma and shunting the post-synaptic potential. The inhibitory synapses would receive inputs from inter-neurons which, in this simplified model, would measure the average activity at the previous time step, thus they would have to act faster than the excitatory neurons. As the total activity rises, the post-synaptic potential decays faster, making it less likely for each neuron to fire at the next step. This simplified model of shunting becomes pathological as the average activity exceeds $1-\Gamma_0/\eta$. In a more realistic (and more conventional) model the inhibitory part $\eta\, Q(t)$ would appear in the denominator, however, in this case a further non-linear term is require for the inhibition to dominate the excitation at high activity levels. Thus to reduce the number of parameters we consider the present model and assume that the activity does not grow too large.

To study the stability of the spontaneous activity state we consider the function

$$\langle Q(t+1)\rangle = NF\left(Q(t)\right), \tag{6}$$

that is the average activity at time $t+1$ given the activity at time t. Assuming that the probability of each neuron firing is $\langle Q(t+1)\rangle/N$ then, if $\langle Q(t+1)\rangle$ is sufficiently small, $Q(t+1)$ will be a Poisson distributed random variable with mean and variance $\langle Q(t+1)\rangle$. Thus the magnitude of the fluctuations will be of order $\sqrt{\langle Q(t+1)\rangle}$. In the cortex the size of the fluctuations will depend on the average activity within a local highly connected region. Assuming that such a region consists of 10^5 neurons firing on average every 200ms then on average 500 neurons would be active in 1ms and the size of the average fluctuations would be of the order of 5%. To obtain an estimate of $F(Q)$ we estimate the probability distribution for the instantaneous fields, $h_i(t)$, and the post-synaptic potentials, $V_i(t)$, (we assume that the average and variance for $h_i(t)$ depends on the activity $Q(t)$ while for $V_i(t)$ they depend only on the time averaged activity $\bar{Q}$ since the integration time of the neurons is assumed to be long). In the following calculation we make the ansatz that the pattern of firing, $S_i(t)$, is random and we will assume that the number of neurons, N, is sufficiently large that we can neglect terms of order $1/\sqrt{N}$. In this case the fields $h_i(t)$ will be independent at each time step and consequently $h_i(t)$ will be independent of $V_i(t)$. The probability for a neuron to fire is then

$$F(Q) = \int_{-\infty}^{\infty} \mathrm{d}h\, P(h) \int_{-\infty}^{\infty} \mathrm{d}V\, P(V)\, \Theta\left(\left(1-\Gamma(Q)\right)(V+h)-\theta\right). \tag{7}$$

where $\Gamma(Q) = \Gamma_0 + \eta\, Q$. The (population) average instantaneous field is

$$h_{av}(t) = \frac{1}{N}\sum_{i=0}^{N} h_i(t) = \langle J_{ij}\rangle Q(t) = cQ(t), \tag{8}$$

where c is the connectivity of the network. The variance in the fluctuation of the instantaneous field is

$$\sigma_h(t)^2 = \frac{1}{N}\sum_{i=0}^{N}\left(h_i(t)-h_{av}(t)\right)^2 = c\sigma_J^2 Q(t). \tag{9}$$

As a useful first approximation we will assume that the instantaneous fields are Gaussian distributed

$$P\left(h_i(t)\right) = \frac{\mathrm{e}^{-\left(h_i(t) - h_{av}(t)\right)^2 / 2\sigma_h^2(t)}}{\sqrt{2\pi}\sigma_h(t)}. \tag{10}$$

The average post-synaptic potential, $V_{av}(t) = N^{-1}\sum_i V_i(t)$, will be given by

$$V_{av}(t+1) = \left(1 - \Gamma(Q(t))\right)\left(V_{av}(t) + h_{av}(t)\right), \tag{11}$$

assuming this remains constant and ignoring correlations in the fluctuations

$$V_{av}(t) = \bar{V}_{av}(\bar{Q}) = \left(1 - \Gamma(\bar{Q})\right)\frac{\bar{h}_{av}}{\Gamma(\bar{Q})} \approx \frac{\bar{h}_{av}}{\Gamma(\bar{Q})}. \tag{12}$$

where $\bar{Q}$ is the time averaged activity and we assume that $\Gamma(\bar{Q})$ is small. The fluctuations in the average post-synaptic potential are given by

$$\sigma_V^2(t) = \frac{1}{N}\sum_{i=0}^{N}\left(V_i(t) - V_{av}(t)\right)^2 = \left(1 - \Gamma(Q)\right)^2\left(\sigma_V^2 + \sigma_h^2\right) \tag{13}$$

the right hand side of equation (13) being evaluated at time $t-1$. Again assuming this does not change in time (and ignoring correlations in the fluctuations), we find

$$\bar{\sigma}_V^2 = \frac{\left(1 - \Gamma(\bar{Q})\right)^2\bar{\sigma}_h^2}{1 - \left(1 - \Gamma(\bar{Q})\right)^2} \approx \frac{c\sigma_J^2\bar{Q}}{2\Gamma(\bar{Q})} \tag{14}$$

The simplest approximation is to assume that the post-synaptic potentials are Gaussian distributed with a mean value $\bar{V}_{av}$ and variance $\bar{\sigma}_V^2$ except in the tail $V > \theta$ since we assume that all neurons that exceed the threshold will fire. Using these approximations for $P(h)$ and $P(V)$ we can substitute them into equation (7) to obtain an estimate of $\langle Q(t+1)\rangle$. In figure 1 we show a plot of $\langle Q(t+1)\rangle$ versus $Q(t)$ for a typical set of parameters. The inhibition and threshold have been chosen so there exists three "fixed points": the stable fixed point at $Q = 0$; an unstable fixed point around $Q = 8$ and a second stable fixed at $Q = 25$ corresponding to the spontaneous background activity state. Here the parameters have been chosen so that the fixed point is maximally stable (i. e. $F'(Q) = 0$). Since the actual activity $Q(t+1)$ is random there is no real fixed point and the activity will fluctuate around the fixed point. (In fact, the average fixed point is shifted due to the fluctuations — this shift can be calculated but is a rather small effect. The fluctuations in $Q(t)$ will be slightly larger than that for a Poisson distribution because $F(Q)$ is not flat.)

The calculation presented is clearly very crude, but it can be refined, for example, by taking into account when the neuron last fired. It can also readily be extended to more complicated forms of inhibition. Nevertheless, even the form presented here provides a useful guide to selecting the parameters to obtain a stable spontaneous background activity. The simulations show several finite size effects not present in the analysis. For systems of the order of 1000 neurons the fluctuations in the activity can be sufficiently large that it falls below the unstable threshold. Thus to maintain a spontaneous activity the network needs to be externally excited. Also in finite systems the neurons will have slightly different average synaptic weights and consequently they have different firing rates. This is more pronounced because of the non-zero resting potential. To remove this bias we can normalize the input weights or introduce a modifiable threshold. With normalized weights the pattern of activity appears to satisfy all tests of randomness, lending support to our initial ansatz.

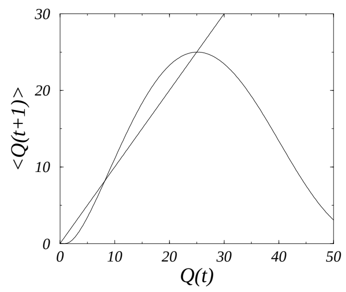

Figure 1. Plot of $\langle Q(t+1)\rangle$ versus $Q(t)$ for a typical set of parameters. The threshold and inhibition has been chosen so that the fixed point $\langle Q(t+1)\rangle = Q(t)$ is maximally stable.

3. Learning Synfire Chains

Having set up a network with a spontaneous background activity we have examined whether synfire chains could be learned by incorporating Hebbian learning into the model and by repeatedly stimulating a small set of neurons in the network. Initially we consider the learning rule

$$J_{ij}(t+1) = J_{ij}(t) + rS_i(t+1)\left(c_{ij}S_j(t) - \frac{J_{ij}(t)}{N}\sum_{k=1}^{N} c_{ik}S_k(t)\right), \qquad (15)$$

and where $c_{ij} = 1$ if the coupling from j to i exists and 0 otherwise. This learning rule has the advantage that it keeps the weights normalized (in the sense defined above) without them going negative. The network was kept in a state of spontaneous activity by externally exciting a random set of neurons whenever the activity dropped below a critical threshold. At the same time fixed sets of neurons were excited at random intervals in an attempt to "seed" a synfire chain.

To see whether synfire chains were formed we looked at density plots of the average auto-correlation function between different time steps

$$q(t,t') = \frac{\sum_{i=1}^{N} S_i(t)S_i(t')}{\sqrt{\sum_{i=1}^{N} S_i(t)\sum_{i=1}^{N} S_i(t')}} \qquad (16)$$

Any time the network has a similar set of active neurons there will be a peak in $q(t,t')$. Before learning, the correlations are Poisson distributed (provided the synaptic weights are normalized), again indicating that the background activity is chaotic. When learning is switched on and an external stimulus is applied, the input patterns start to recruit neurons to form the next pool in a synfire chain. This shows up in the density plot of $q(t,t')$ as a peak when the two times t and t' are one step after different presentations of the same input stimulus. However, each time the stimulus is applied the resting potentials are different because the past history of the neurons are very different, so that a new set of neurons can be recruited. After many presentations of a stimulus a very large number of neurons have learned to respond to the stimulus so that the activity becomes very large. In the following step the inhibition cuts in and prevents any of the neurons firing. Consequently no synfire chains is formed.

There are a number of ways to alleviate this instability. Rather than exciting a large set of neurons at one time step a smaller number of neurons can be stimulated sequentially. Also rather than using an abrupt learning a more non-linear learning rule can be used [7, 8]. Furthermore the inhibition can be made to adapt so that it acts more strongly in response to patterns of excitation it has seen before. Finally we can normalize the synaptic weights coming *from* a neuron — this would be the case if, for example, the modification of a synapse relied on some restricted resource of the efferent cell.

Using different combinations of these strategies it is possible to learn synfire chains consisting of around five patterns. However, this type of network is also unstable against the formation of short cycles. Once these begin to form they are very rapidly learned and prevent any other activity. This prevents longer chains from being learned. This instability is present even in n-winner-takes-all networks where only the neurons with the n highest post-synaptic potential fire. As the learning proceeds eventually a short cycle is found.

This instability can be alleviated by using a modifiable threshold, particularly when this threshold increases supra-linearly with its average firing rate. Such a fatigue mechanism appears very plausible and would seem to be necessary in any recurrent network to prevent short cycles from taking over. Having an adaptive threshold is also important because the variance in the coupling typically increases with time so that the average activity would otherwise increase as the network learns.

Using the combinations of the various strategies discussed above we can "grow" synfire chains of up to 20 pools starting from stimulating 5 pools in sequence. A typical example of a such a synfire chain is illustrated in figure 2. As the learning develops the initial parameters which created a stable background of activity change so that it becomes harder to maintain the spontaneous activity. Because of these difficulties we were unable to learn longer synfire chains.

4. Discussion

Learning synfire chains from a background of chaotic spontaneous activity poses many problems of control. The conditions for the spontaneous activity will slowly change over time and in addition the network can learn parasitic states rather than synfire chains. Nevertheless these instabilities can often be remedied by including plausible control mechanisms. Unfortunately in doing so the space of parameters grows and it becomes increasingly difficult to understand the underlying mechanism. The recording data of Abeles *et al.* [2, 3] suggests that in cortex there may exist synfire chains of up to several hundred pools long. The question of whether such long synfire chains can plausible be learned through a self-organizing mechanism based on a spontaneous background activity remains, at present, unanswered by this investigation. Certainly there are instabilities which makes the problem non-trivial. On the other hand, there would appear to be plenty of plausible control mechanisms which, if appropriately adjusted, might allow long synfire chains to be formed.

References

[1] M. Abeles. *Corticonics: Neuronal Circuits of the Cerebral Cortex.* Cambridge University Press, Cambridge, UK, 1991.

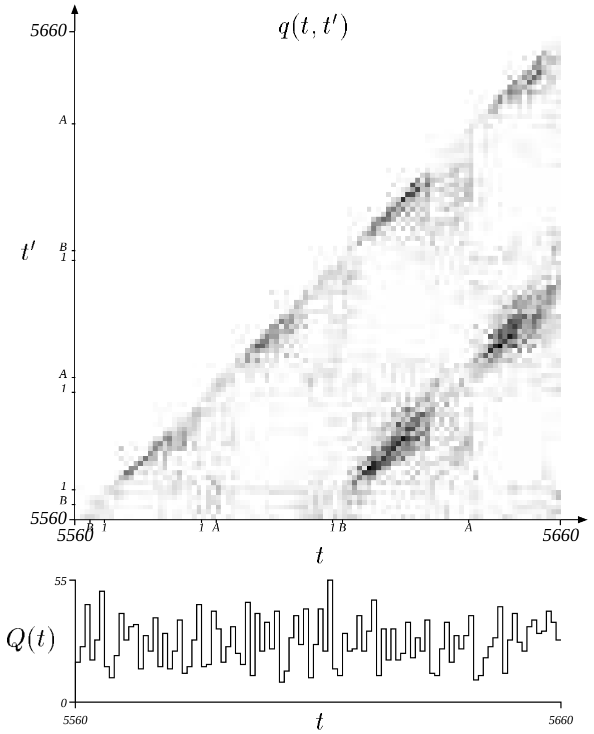

Figure 2. Density plot $q(t, t')$ is shown for $5560 \geq t, t' \geq 5660$ together with the activity $Q(t)$ over the same time interval. The letters A and B mark where two different sets of input stimuli are applied. The letter I marks where a random external stimulus was applied to prevent the activity from decaying to zero. The simulation consists of 1000 neurons in this case with normalization of both the input and output coupling weight vectors. We have also used an adaptive threshold and learning inhibition.

[2] M. Abeles, E. Vaadia, H. Bergman, Y. Prut, I. Haalman, and H. Slovin. Dynamics of Neural Interactions in the Frotal Cortex of Behaving Monkeys. *Concepts in Neuro-Science*, 4(2):131–158, 1993.

[3] M. Abeles, H. Bergman, E. Margalit, and E. Vaadia. Spatiotemporal Firing Patterns in the Frontal Cortex of Behaving Monkeys. *J. Neurophys.*, 70(4):1629–1638, 1993.

[4] E. Bienenstock. A Model of Neocortex. *Network: Computation in Neural Systems*, 6:179, 1995.

[5] Herrman M., J. Hertz, and A. Prügel-Bennett. Analysis of Synfire Chains. *Network: Computation in Neural Systems*, 6(3):403–414, 1995.

[6] Amit D. J. and N Brunel. Global spontaneous activity and local structured (learned) delay activity in cortex. Dipartimento di Fisica Universitá di Roma La Sapienza Preprint 1085., 1995.

[7] Prügel-Bennett A. and J. L. Shapiro. Statistical mechanics of unsupervised Hebbian learning. *J. Phys. A: Math. Gen.*, 26:2343–2369, 1993.

[8] Prügel-Bennett A. and J. L. Shapiro. The partitioning problem in unsupervised learning for non-linear neurons. *J. Phys. A: Math. and Gen.*, 26:7417–7426, 1993.