THE EFFECTS OF PAIR-WISE AND HIGHER ORDER CORRELATIONS ON THE FIRING RATE OF A POST-SYNAPTIC NEURON


```
@Article{nc:bohte+spekreijse+roelfema:2000,
  author       = {Bohte, S.M. and Spekreijse, H. and Roelfsema, P.R.},
  title        = {The effects of pair-wise and higher-order
                  correlation on the firing rate of a
                  post-synaptic neuron},
  journal      = {Neural Computation},
  year         = {2000},
  volume       = {12},
  number       = {1},
  pages        = {153-179}
}
```


This re-print corresponds to the article "*The effects of pair-wise and higher order correlations on the firing rate of a post-synaptic neuron*", by Sander M. Bohte, Henk Spekreijse, and Pieter R. Roelfsema, and appeared in **Neural Computation**, 2000 (1), 153-179.

The effects of pair-wise and higher order correlations on the firing rate of a post-synaptic neuron

Sander M. Bohté [a] Henk Spekreijse[b]

Pieter R. Roelfsema [b]

[a] CWI, Kruislaan 413, 1090 GB Amsterdam, The Netherlands
[b] Graduate School Neurosciences Amsterdam, AMC, Dept. Medical Physics, University of Amsterdam, The Netherlands
Ophthalmic Research Institute, P.O. Box 12141, 1100 AC, Amsterdam.

Abstract

Coincident firing of neurons projecting to a common target cell is likely to raise the probability of firing of this post-synaptic cell. Therefore synchronized firing constitutes a significant event for post-synaptic neurons and is likely to play a role in neuronal information processing. Physiological data on synchronized firing in cortical networks is primarily based on paired recordings and cross-correlation analysis. However, pair-wise correlations among all inputs onto a post-synaptic neuron do not uniquely determine the distribution of simultaneous post-synaptic events. We develop a framework in order to calculate the amount of synchronous firing that, based on maximum entropy, should exist in a homogeneous neural network in which the neurons have known pair-wise correlations and higher order structure is absent. According to the distribution of maximal entropy, synchronous events in which a large proportion of the neurons participates should exist, even in the case of weak pair-wise correlations. Network simulations also exhibit these highly synchronous events in the case of weak pair-wise correlations. If such a group of neurons provides input to a common post-synaptic target, these network bursts may enhance the impact of this input, especially in the case of a high post-synaptic threshold. Unfortunately, the proportion of neurons participating in synchronous bursts can be approximated by our method under restricted conditions. When these conditions are not fulfilled, the spike trains have less than maximal entropy, which is indicative of the presence of higher order structure. In this situation, the degree of synchronicity cannot be derived from the pair-wise correlations.

1 Introduction

The occurrence of correlations in the spike-trains of neurons responding to the same object has raised considerable excitement during the last decade (review by [1]). Correlations between pairs of neurons are thought to reflect a high degree of synchronous firing within a larger assembly of neurons [2, 3] and can have a high temporal precision, in the range of a few milliseconds [4, 5, 6, 7, 8, 9, 5]. Von der Malsburg [10] suggested that assemblies of neurons might convey additional information by firing in synchrony, since synchrony could be instrumental in forming relationships between the members of such an assembly.

However, the possible relevance of fine temporal structure in spike-trains opposes another widespread belief. The irregular timing of cortical action potentials is quite often attributed to stochastic forces acting on the neuron [11, 12]. In such a stochastic model, the information is thought to be conveyed to the next processing stage (cortical layer) by pools of neurons using a noisy rate code. Each individual neuron is considered to be a slow, unreliable information processor, reflecting changes in its receptive field by modulating its average firing rate. Only by pooling the information from a larger number of neurons, a reliable rate code can be obtained. Obviously, this scheme does not need precise timing of the individual spikes to convey much information.

These two opposing views on the role of temporal structure of neuronal information processing are subject of considerable debate [13, 14, 15]. This debate has focussed on two important questions. First, is the cortical neuron a coincidence detector (on the millisecond time-scale) and second, how much coincident input is there?

The first question refers to the relevance of synchronous pre-synaptic spikes. It has been suggested that synchronous input induces a higher firing rate in the post-synaptic target cell. Does this assumption hold, especially on a millisecond time-scale? This question has been amply recognized, and several studies have attempted to answer it. Shadlen & Newsome [14] argue that, based on physiological considerations, a cortical neuron is not capable of detecting very tightly synchronized input. However, others have argued that cortical neurons might have a high sensitivity for the synchronicity in their input [16, 13]. Softky [16] pointed out that the biological data available leave too many parameters undetermined to draw any definite conclusions on biological properties that distinguish the various models. Two further studies on the impact of synchronized input on a post-synaptic target reinforce this observation. Using detailed models of groups of neurons, [17] and [18] studied the impact of coincident input on the firing rate of a post-synaptic neuron. Their conclusions are similar to Softky's: within the biologically plausible parameter ranges synchrony may either increase or decrease the firing rate of post-synaptic neurons.

In the present study we attempt to shed more light on the second question: how much synchrony is there? In general, it is implicitly assumed that pair-wise correlations provide a good estimate of the amount of synchrony in a pool of neurons from which recordings are obtained. However, to date there are no direct electrophysiological measurements of large synchronous pools of cortical

neurons. Most of the physiological data on neuronal synchronization so far have been obtained using cross-correlation techniques (with the notable exception of the work of [9]). These techniques merely provide information on the probability of events in which a pair of neurons fires at the same time (that is, within some time-window). Unfortunately, pair-wise correlations only provide an indirect estimate of the probability of higher-order events, like the coincident firing of, say, 5 or 50 neurons. Even when the pair-wise correlations between all neurons of a network are fixed, the probability of these higher-order events remains undetermined, as is illustrated in figure 1. Pair-wise correlation is defined as

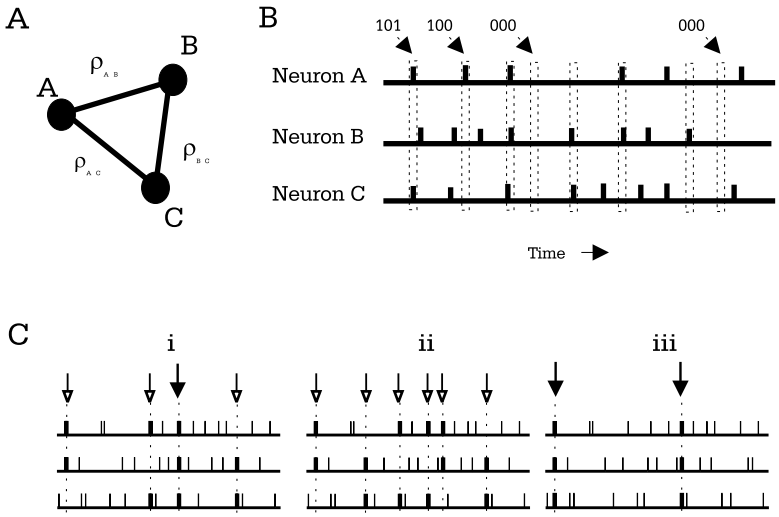

Figure 1: (A) Three neurons with correlated activity. The pair-wise correlation coefficients are ρ_{AB}, ρ_{AC} and ρ_{BC}. (B) Examples of different spike configurations in windows of the same size. The horizontal lines represent the spike trains, and each tick denotes a spike. Spikes occurring with a time-window (dotted box) are considered to be coincident. (C) Three examples of correlated spike trains. Pairs of neurons in the three panels have the same pair-wise correlation. Spike-doublets are shown as unfilled arrows, and triplets as filled arrows. It can be seen that the number of triplets differs from panel to panel.

the difference between the probability of two neurons firing simultaneously, and the product of their firing rates (the coincidence rate dictated by chance). This value is the same for any pair of neurons in the three panels of figure 1C. However, the number of spike-triplets differs considerably from panel to panel. Given this example, it is quite clear that the number of neurons that fire within some time-window (the measure of coherence) is not exclusively determined by the pair-wise correlation coefficient. And, although this is an artificial example,

it is already quite difficult to determine intuitively how many of these triplets can be attributed to higher-order correlation, and how many result from two "doublets" that happen to occur at the same time. In a first attempt to quantify the incidence of higher-order correlations, Martignon *et al.* [19] analyzed data from six cortical neurons. Unfortunately, we found that their methods cannot be used for the analysis of large numbers of neurons (as will be discussed). The main goal of the present article is to study the relationship between pair-wise correlations and the amount of synchronicity in a pool of neurons, and to determine the impact of the synchronous events on a post-synaptic target cell.

2 Mathematical Solution of the three-neuron problem

To illustrate the general methodology used for estimation of the probability of higher order events, we examine the three neuron network of figure 1. We wish to calculate the probability that a triplet (or an N-cluster, where N equals 3) occurs within a given time-window. The null hypothesis is that no structure is present in the spike-trains other than the pair-wise correlations. That is, all triplets should be due to the occurrence of two doublets at the same time, by chance forming a triplet. Csiszar [20] has proved the unique existence of a distribution with just this property. The basic approach for calculating the probability distribution involves maximizing the informational entropy of the data, while preserving the pair-wise correlation and the firing rate (see also [19]). This informational entropy is a measure of the "order" in the data: the more structured the data, the lower the entropy. By measuring the pair-wise correlations and the firing rates (the first order correlation), a certain degree of order is fixed. Taking these constraints into account, maximizing the entropy will minimize all higher-order correlations since higher-order correlations will add "structure" to the distribution, further lowering the entropy. Therefore, maximal entropy implies minimal higher-order correlations. Our aim is to obtain the distribution of N-clusters that has maximal entropy. We will illustrate the procedure by computing this distribution for three connected neurons.

The neurons are labeled A, B and C, their firing probabilities f_{1A}, f_{1B} and f_{1C} and the pair-wise correlation coefficients are denoted as ρ_{AB}, ρ_{BC} and ρ_{AC} (f_{1i} denotes the probability that neuron i fires in a particular time bin, and depends on the firing rate and the size of the time bins. Dividing the firing probability by the length of the time bin in seconds thus yields the firing rate. The rationale of the suffix 1 will become clear below). Given these coefficients as constraints, probabilities for all 2^n possible events can be calculated. For example G_{010} designates the probability of the event in which B is firing and A and C are silent (Fig. 1B).

There are 8 probabilities to solve for. Since firing probabilities and pair-wise correlation coefficients are fixed, 7 equations can easily be obtained: The sum

of the probabilities of all possible events equals one:

$$G_{000} + G_{001} + G_{010} + G_{011} + G_{100} + G_{101} + G_{110} + G_{111} = 1 \tag{1}$$

The firing-probability of a neuron A is:

$$G_{100} + G_{101} + G_{110} + G_{111} = f_{1A} \tag{2}$$

(The equations for f_{1B} and f_{1C} are derived analogously) Let f_{2AB} denote the probability that A and B fire simultaneously (the suffix now indicates a cluster of size 2). f_{2AB} is determined by the correlation coefficient ρ_{AB}, since

$$\rho_{AB} = \frac{f_{2AB} - f_{1A}f_{1B}}{\sqrt{(f_{1A} - f_{1A}^2)(f_{1B} - f_{1B}^2)}} \tag{3}$$

This implies that ρ_{AB} determines f_{2AB} and since f_{1A} and f_{1B} are fixed, this yields:

$$G_{110} + G_{111} = f_{2AB} \tag{4}$$

and corresponding equations are derived for ρ_{AC} and ρBC. These seven equations (1,2a-c and 4a-c) yield one free parameter (G_{111}, for example), which determines the entropy of the probability distribution. By calculating the value at which the distribution has maximal entropy, we fix this last free parameter. The entropy-function is defined as:

$$S \equiv -\sum_i G_i \ln(G_i) \qquad i \text{ over all possible spike configurations} \tag{5}$$

The entropy is maximized by solving the following equation, which has a unique solution (see appendix A)

$$\frac{\partial S}{\partial G_{111}} = 0 \tag{6}$$

Using this equation the probability distribution of configurations with the desired zero higher-order correlation can easily be calculated (see also [19]).

3 Calculating the Distribution with N Identical Neurons

For more than three neurons, the equations can no longer be solved easily by analytical means. For instance, for 4 neurons, the same calculations yield the entropy as a function of 5 free parameters. Calculating the maximum of this function is already quite complicated. Extending this to N neurons, only $1 + N(N+1)/2$ equations are determined by the pair-wise correlations and firing-probabilities, with 2^N parameters to solve for. To overcome this problem, an iterative algorithm for calculating the maximal entropy distribution has been provided by Gokhale & Kullback [21]. Unfortunately, this algorithm has a serious drawback: the number of calculations increases exponentially with

the number of neurons. If one is interested in the behavior of many correlated neurons, computational restrictions prohibit this method of calculating the distribution for more than 20 neurons on a workstation, thus putting the more serious number of neurons out of reach of even the fastest supercomputers.

To overcome the problem of increasing computation time we made the additional assumption that all neurons have identical properties, which implies a major degeneration of the probability-space. Since any two neurons are equal, the same will hold for the probability of all permutations of spike configurations. In addition, the N firing probabilities and the $N(N-1)/2$ pair-wise correlation coefficients will also be equal. The distribution of spike configurations is described by $N+1$ variables, D_0 through D_N, where D_i denotes the probability of a particular spike configuration in which exactly i neurons fire.

In the case of $N = 7$: $D_1 = G_{1000000} = G_{0100000} = \ldots = G_{0000001}; D_2 = G_{1100000} = G_{1010000} = G_{0010010}$, etc. Since all permutations of i spiking and $N-i$ silent neurons have the same probability D_i, the requirement that all probabilities add up to 1 now reads:

$$1 = \sum_{i=0}^{N} \left(\begin{array}{c} N \\ i \end{array} \right) \cdot D_i \tag{7}$$

The firing probability of a cell equals (see appendix A):

$$f_1 = \sum_{i=1}^{N} \left(\begin{array}{c} N-1 \\ i-1 \end{array} \right) \cdot D_i \tag{8}$$

Under the assumption of identical neurons, in (3) f_{1B} equals f_{1A}. Rewriting (5), replacing f_{1A} and f_{1B} by f_1, and f_{2AB} by f_2 yields:

$$\rho = \frac{f_2 - f_1^2}{f_1(1-f_1)} \tag{9}$$

Thus, f_2 can once again be calculated from the firing probability of a neuron and the correlation. f_2, the probability that any two particular neurons fire at the same time equals:

$$f_2 = \sum_{i=2}^{N} \left(\begin{array}{c} N-2 \\ i-2 \end{array} \right) \cdot D_i \tag{10}$$

By maximizing the entropy S, we derive the remaining $N-2$ equations: The entropy is defined as:

$$S = \sum_{i=0}^{N} \left(\begin{array}{c} N \\ i \end{array} \right) \cdot D_i \ln(D_i) \tag{11}$$

Maximization of entropy yields:

$$\frac{\partial S}{\partial D_i} = 0 \qquad \text{for } i = 3 \ldots N \tag{12}$$

(A proof of a unique solution of this set of equations is included in appendix A) After some calculations the $N-2$ equations turn out to have the form:

$$\begin{aligned} \ln(D_i) &= -(-\frac{1}{2}i+1)(i-1)\ln(D_0) - i(i-2)\ln(D_1) + \\ &\quad +\frac{1}{2}i(i-1)\ln(D_2) \qquad \text{for } i = 3\ldots N \end{aligned} \tag{13}$$

Inserting these $N-2$ equations into (7), (8) and (10), we can numerically solve for D_0, D_1, and D_2 using the Newton-Raphson method for nonlinear systems of equations [22].

The maximal entropy distribution, thus defined, depends on only two parameters, the firing probability f_1, and the pair-wise correlation ρ. Figure 2 illustrates the shape of the N-cluster distribution with maximal entropy for a network of 150 neurons, and its dependence on f_1 and ρ. Figure 2 shows the probability P_i that a particular *number* of neurons fire simultaneously, where P_i equals the sum of the probabilities of all configurations in which exactly i neurons fire:

$$P_i = \left(\begin{array}{c} N \\ i \end{array} \right) \cdot D_i \tag{14}$$

For small values of ρ, the distribution of N-clusters approaches a binomial distribution. This determines the first peak of the N-cluster distribution, corresponding to small cluster sizes. For larger values of ρ, a second peak appears in the distribution, the amplitude of which grows with increasing ρ. A second influence of increasing ρ is a divergence of the two peaks. This divergence can be seen most clearly in the contour plots of figure 2D-F. Indeed, in the limiting case of $\rho = 1$, the first peak approaches a cluster size of 0, and the second peak a cluster size of 150, since all neurons fire at exactly the same time.

An increase in f_1 shifts the first peak of the distribution to larger values, as is predicted by the binomial distribution (Fig. 2A-C). Remarkably, an increase in f_1 is also associated with a shift of the second peak, to *smaller* values. Thus, the maximal entropy distribution predicts that low firing probabilities are associated with sparse, but highly synchronized bursts (eg: the average size of the second peak in A at $\rho = 0.165$ equals 142 vs 110 in C, whereas the cumulative probability increases from 0.009 in A to 0.076 in C. The second peak is defined as all clusters with $p > 10^{-4}$, starting on the positive slope after the first peak). With a higher firing probability, synchronous bursts occur more frequently, but comprise fewer spikes.

4 An artificial neuron network

We now compare N-cluster distributions obtained from simulations of an artificial neural network to the maximal entropy distribution. Any difference between the two distributions can then be attributed to higher-order correlations. The network used is based on a network described in [23], which, in turn, is based on the work of [24]. This network was chosen since it could easily be adapted to consist of identical neurons.

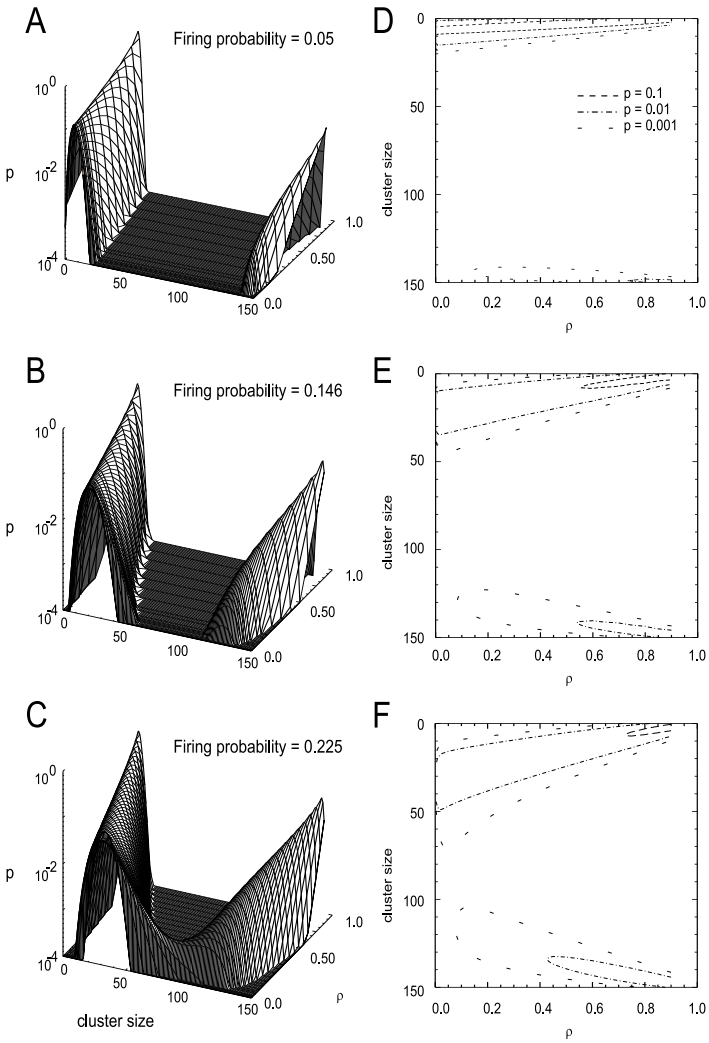

Figure 2: (A-C) Probability of N-clusters (p) as a function of the pair-wise correlation ρ, for three firing probabilities, $f_1 = 0.05$ (A), 0.146 (B) and 0.225 (C). Probabilities are clipped at a value of 10^{-4}. (D-F) Contour plots of the same data show that the separation between the two peaks becomes larger with increasing ρ. Contour levels indicate probabilities of 0.001, 0.01 and 0.1.

4.1 The neuron model

The neuron model involves four variables: the membrane potential $E(t)$, a potassium current $g(t)$, the spiking threshold $\Theta(t)$ and the neuronal output $o(t)$. The

dynamics of neuron i are described by four coupled equations:

$$\begin{aligned}
\tau_E \frac{dE_i(t)}{dt} &= -(E_i(t) - E^0) - (g_i(t) - g^0)(E_i(t) - E^k) + \eta_i(t) + \\
&\quad - \left[\sum_j w_{ij} \cdot o_j(t - \tau_i j) \right] (E_i(t) - E^{ex}) \qquad (15) \\
\tau_\Theta \frac{d\Theta_i(t)}{dt} &= -(\Theta_i(t) - \Theta^0) + c(E_i(t) - E^0) 0 \ \leq \ c \ \leq \ 1 \qquad (16) \\
\tau_g \frac{dg_i(t)}{dt} &= -(g_i(t) - g^0) + \tau_g b o_i(t) \qquad (17) \\
o_i(t) &= \begin{cases} 1 & \text{if } E_i(t) > \Theta_i(t) \\ 0 & \text{otherwise} \end{cases} \qquad (18)
\end{aligned}$$

Without input, the membrane potential $E_i(t)$ is driven towards its resting value E^0 with a time constant τ_E. An influx of potassium drives the potential towards the potassium equilibrium potential E^k. Excitatory input moves the potential towards the ionic equilibrium value E^{ex}. In the simulations, the values used in [23] were adopted: $E^0 = 0$, $\tau_E = 2.5$ msec, $E^k = -1$, $E^{ex} = 7$. We remark that the membrane time constant τ_E may appear to be rather small, although some have argued that it may be within the biologically plausible range [13, 25]. In the discussion we will address the dependence of our results on this particular choice. Input to the cells was strictly excitatory, with synaptic delay $\tau_{ij} = 0.5$ msec. External input is provided by external stochastically spiking units and is treated equivalent to internal input. Internal noise is added by the additive noise term $\eta_i(t)$ in (15), with a standard deviation of $0.06E_i(t)$. Equation (16) describes the slow adaptation of the threshold Θ_i to the membrane potential, modeling an adaptation of the neuron to excitation ($\Theta_0 = 1$, $\tau_\Theta = 10$ msec, $c = 0.3$). Equation (17) governs the potassium current decay dynamics, which is driven towards $g_0 = 0$ with time constant τ_g (5 msec). In the case of a spike, the potassium current rises by an amount $b(4.0)$ corresponding to the outward potassium current. A spike is generated each time the membrane potential exceeds the threshold (18), after which the neuron is in a refractory state for another 1.5 msec. In the actual numerical implementation of the neuron model, a discrete time analogue of (15)-(18) was used. These equations were iterated with time-steps corresponding to 0.5 msec of real time.

Each neuron was connected to every other neuron with a fixed homogeneous synaptic weight $w_{ij} = w_{int}/N > 0$. Here, w_{int} denotes the sum of the strengths of all synapses from within the network impinging on a single cell. External input consisted of N independent stochastic elements generating spikes with $p = 0.05$ per msec (50 Hz), connected in a 1 to 1 fashion with the network neurons with a constant weight $w_{ext} = 0.8$.

The parameter w_{int} was varied during the simulation. Increasing the value of the weights drives the network from a stochastic mode into an oscillatory mode (Deppisch *et al.*, 1993). Somewhere within this range of synaptic weights the network alternates between the stochastic and oscillatory mode.

All results are based on simulations with a network with $N = 150$ neurons, with the exception of the results in section 4.5. This was a compromise between a realistic number of neurons, and a network consideration: the absence of inhibitory neurons makes a network very quickly prone to saturation, a state where the constant excitation induces a very sharp oscillation. A further increase in the number of neurons also narrows the weight-range in which the network is in the alternating state. The network output was obtained by recording the output of all 150 neurons. A sample of the output of 50 neurons is shown in figure 3.

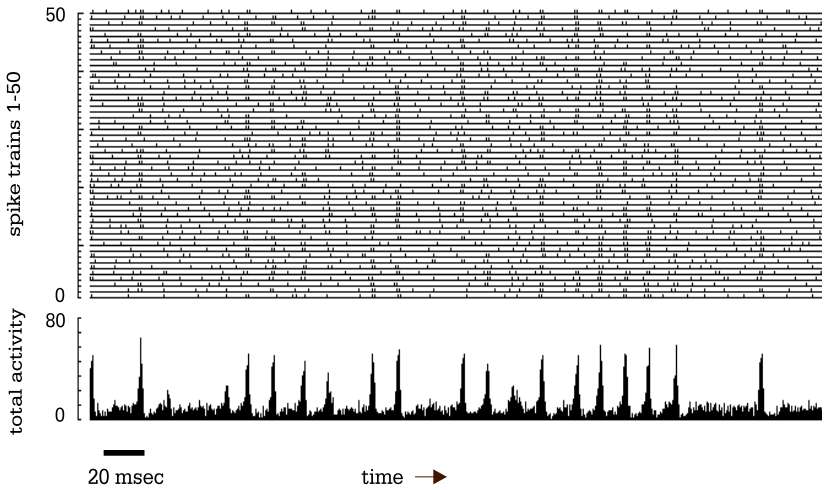

Figure 3: Upper panel, activity of 50 of the 150 neurons. On the horizontal axis time is displayed (sampling-time, 0.5 msec). The vertical ticks indicate action potentials. Oscillatory episodes are alternated by periods of more random activity. In the lower panel the summed activity of all 150 neurons is shown.

4.2 Network Simulations

In order to vary the average pair-wise correlation, simulations were performed with different values of the synaptic weight w_{int}. As was noted in [23], the network exhibits three distinct modes, which depend on the value of the synaptic weight. The first, at low values of the synaptic weights, is the stochastic mode in which the average correlation between pairs of neurons is near zero. At the other end of the scale is the mode with high values of internal weight. In this mode the network activity is highly oscillatory. In between these extremes

are synaptic weight values for which the network exhibits episodes in which many neurons fire synchronously, alternated by episodes in which neurons are synchronized to a lesser degree. Typical activity of the neurons in these three network modes is shown in figures 4A-C. Also plotted are the cross-correlations between a pair of neurons in the network (Fig. 4D-F). As observed in [23], these cross-correlograms show qualitative resemblance to data obtained from electrophysiological recordings (e.g [26]). Remarkably, the network with intermediate synaptic strength ($w = 3.875$, $\rho = 0.03$, for 2 msec bin-size) exhibits occasional population bursts, which hardly show up in the correlation function (Fig. 4B,E). Thus, a network state in which a number of highly synchronous population bursts occur, can be associated with correlation functions indicative of weak pair-wise coupling. When investigating the relationship between occurrences of N-clusters and the pair-wise correlation, an additional assumption has to be made with regard to the maximal time-difference between two spikes that are considered to be synchronous. As a first approach we used a time-window of 2 msec. The maximal firing rate of the neurons is determined by the refractory period (1.5 msec) and the duration of an action potential (0.5 msec). As all experiments, with the exception of those described in section 4.4, have a bin-width of 2 msec, the firing rate (in Hz) in these experiments can be obtained by multiplying the firing probability with a factor of 500. During network bursts, neurons reach this maximal firing rate, and a single spike occurs in each 2 msec bin. The effect of changing the bin-width on the distribution of cluster-sizes will be investigated in section 4.4.

Figure 5 shows the relationship between the occurrence of N-clusters and the pair-wise correlation. Plotted is the probability P_i (see also eq. (14)) of a particular number of coincident spikes, for simulations with different pair-wise correlations (different values of w_{int}). As can be seen in figure 5 A, most 2 msec bins are occupied by N-clusters containing a fairly low number of spikes. This corresponds to the stochastic activity between bursts and the probability of these events is approximated by a binomial distribution. Synchronous bursts are represented by the second peak in the probability distribution, as can be seen more clearly in the logarithmic plot of figure 5 B. Between these two peaks is a plateau of time bins in which an intermediate number of neurons fire. These events are caused by time bins that are aligned on the onset or end of a burst. Increases in the pair-wise correlation are associated with a slightly smaller first peak, and an enhanced second peak. For very large, and biologically implausible, pair-wise correlations a third peak containing intermediate cluster sizes is visible which can be attributed to the onset and offset of bursts. Let us now consider the impact of this distribution on a post-synaptic cell, receiving input from all the network neurons. For such a cell, the exact number of synchronous spikes (the quantity plotted in figure 5 A,B) is not important. Rather, for a post-synaptic neuron with a firing-threshold of m EPSP's the incidence of m or *more* coincident spikes is a much more significant quantity, since this determines its firing rate to a large extent. Therefore, we computed a *cumulative probability distribution*: the probability that m or *more* neurons fire synchronously. The cumulative probability distributions (the cumulatives of Fig. 5B) are plotted

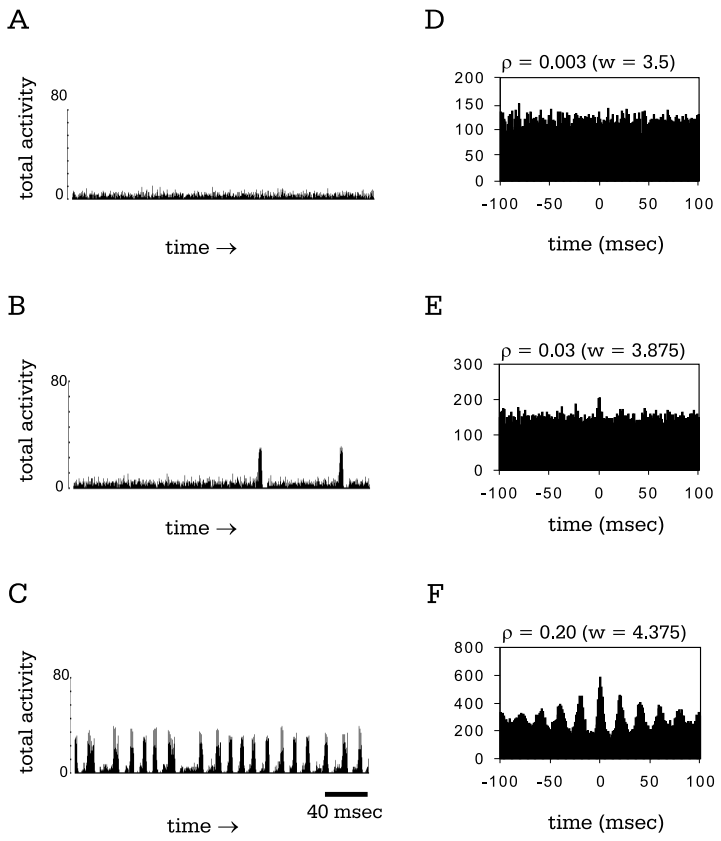

Figure 4: Network activity for 3 values of the synaptic weight, w_{int}. (A-C) Summed activity of the 150 neurons. The synaptic weight was 3.5 (A, $\rho = 0.003$), 3.875 (B, $\rho = 0.03$) and 4.375 (C, $\rho = 0.20$) Pair-wise correlations were determined for coincidences within a window of 2 msec. (D-F) Cross-correlation functions of 2 randomly selected neurons for the same synaptic weights as are plotted on the left. Even for a low pair-wise correlation ($\rho = 0.03$) highly synchronous network bursts occur, but these hardly show up in the cross-correlogram.

in figure 5C. Suppose that the membrane-potential of a post-synaptic neuron, which receives input from all cells in our network, equals the resting potential at time t. The probability that this neuron fires at time $t + 1$ is equal to the probability of Θ' or more coincident spikes, where Θ' is the ratio between the post-synaptic threshold (Θ) and the EPSP amplitude. In other words, the curves in figure 5C illustrate the relation between the threshold and firing

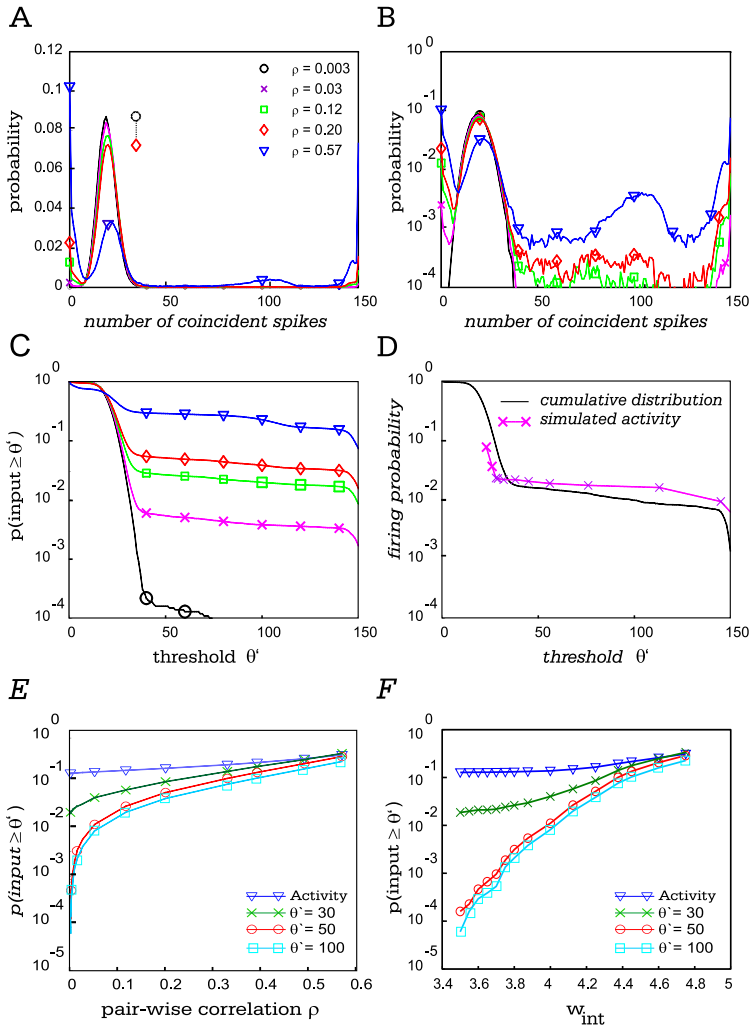

Figure 5: (A,B) Probability distribution of N-clusters, plotted on a normal scale (A) and a logarithmic scale (B). (C) Cumulative probability of N-clusters. Abscissa, cluster-size (Θ'). Ordinate, probability of a cluster with a size that is equal to, or larger than Θ'(D) Comparison of the cumulative N-cluster distribution (solid line) to the firing-probability of a post-synaptic neuron, which receives input from all 150 neurons in the network. The threshold Θ' is defined as $\Theta/w_{i,151}$. (E) Dependence of the firing probability of a neuron receiving input from all 150 neurons on its relative threshold Θ' and the pair-wise correlation in the network. Different curves correspond to different values of Θ'. Triangles, dependence of the average firing probability f_1 on the pair-wise correlation. (F) Same as (E), but the firing probability of the post-synaptic cell and f_1 are plotted as a function of w_{int}.

probability of a neuron, which receives input from all neurons in the network, in the case of a very short membrane time constant. The curve for a very low pair-wise correlation coefficient ($\rho = 0.003$) shows the quickly declining probability of higher order events. The curves for network activity with a larger pair-wise correlation exhibit a plateau before their decline at very high numbers of synchronized spikes. This plateau results from the second peak in the N-cluster probability distribution. It is remarkable that even small changes in the strength of the pair-wise correlations, which are not physiologically implausible (e.g [5, 27]), exhibit a strong effect on the firing probability in case of an intermediate threshold Θ'. Indeed, the firing probability of a neuron with a threshold of, say, 50 EPSP's is raised by more than an order of magnitude by an increase in the pair-wise correlation as small as 0.08 (compare the distributions for $\rho = 0.03$ and $\rho = 0.11$).

The question remains how closely the firing-rate of a post-synaptic neuron with non-zero membrane time-constant is approximated by the cumulative probability distribution. There would obviously be a one-to-one relationship for a post-synaptic neuron that has "zero" memory, i.e. a neuron which is only influenced by the input in the previous time-step (the bin-size for which spikes are considered synchronous, as discussed above).

However, typical neurons have parameters with larger time constants, including the refractory period, and the time-constant of the membrane. In order to assess the impact of the "non-zero memory", an additional 151th neuron was included in the network. This cell was identical to the other 150 neurons from which it received input, but it did not project back to them. The pair-wise correlation was fixed at 0.12 and the membrane-time constant was fixed to the same value as the other 150 neurons, thus realizing a non-zero memory post-synaptic neuron. Simulations were performed with different values of the post-synaptic threshold Θ', by varying $w_{i,151}$ ($1 < i < 150$), the strength of the synapses projecting onto neuron 151. The dependence of the firing probability on Θ' ($\Theta/w_{i,151}$) is shown in figure 5D. Superimposed on this graph is the distribution of N-clusters. As expected, some minor differences between the firing probability and the cumulative probability distribution are observed. Most of these differences can be explained, given the fact that the non-zero average activity in a pool of neurons keeps the membrane potential of individual neurons at a somewhat higher level than the resting potential. This lowers the average number of coincident spikes the neuron would require to cross threshold. On the whole however, the firing probability of the post-synaptic neuron with a non-zero membrane constant is predicted with good accuracy by the cumulative probability distribution. A remarkable feature of figure 5D is that the firing probability of a neuron with a threshold of for example 40 does not differ much from that of a neuron with a much higher threshold (e.g. 130). Most of the spikes of a post-synaptic cell with a threshold larger than 40 are triggered by synchronous bursts in which the majority of neurons participate. As these bursts are separated in time by about 15 msec, a relatively large number compared to the membrane time-constant, this result shows that regarding the firing rate of a post-synaptic neuron in our model, the actual time-structure

in the cluster-distribution is of far less importance than the sheer number of bursts.

Using this interpretation of the cumulative probability distribution, the effect of an increase in the pair-wise correlation in the network on post-synaptic neurons was investigated by varying the synaptic weight (w_{int}). Figure 5E shows the relationship between the average pair-wise correlation and the firing probability of a hypothetical post-synaptic neuron with threshold Θ'. Calculated were the firing-probabilities for thresholds Θ' = 30, 50 and 100. It can be seen that in the biologically relevant range of pair-wise correlations ($\rho = 0 - 0.2$), there is a monotonic relationship between the post-synaptic firing probability and the pair-wise correlation. For values of ρ below 0.2, changes in the pair-wise correlation are associated with a relatively large increase in the firing probability of the post-synaptic neuron. Importantly, higher values of ρ are also associated with an enhanced firing rate of the pre-synaptic network neurons (Fig. 5E). This increase in activity undoubtedly contributes to the enhanced probability of higher order events. However, it should be noted that the probability of higher order events exhibits a steeper dependence on ρ than the firing probability of the network neurons (Fig. 5E). Figure 5F illustrates the dependency of the firing probability of the post-synaptic neuron on w_{int}, the parameter that was actually varied during the simulations.

4.3 Estimation of the N-cluster distribution by entropy maximization

In most electrophysiological experiments only data is available on firing probabilities and pair-wise correlations. Using the mathematical framework developed in section 3, we investigated whether the observed relationship between the probability of N-clusters and the correlation coefficient could have been predicted from these two types of measurements. For the network behavior at different values of w_{int} (3.5, 3.875 and 4.125), distributions of N-clusters were calculated which maximized entropy under the constraints of the observed firing probability and pair-wise correlations. A comparison between distributions based on the maximal entropy calculation and the experimental distribution is shown in figure 6A-C. The distribution that maximized the entropy deviates somewhat from the observed distribution for larger values of ρ.The maximal entropy calculation underestimates the incidence of clusters between 60 and 100 spikes, which occur during start and end of population bursts, as was discussed above. In addition, the second peak in the maximal entropy distribution is located at a cluster size of 130, whereas the actual location of this peak is 150. Typically, all neurons fire within a 2 msec window during a network burst. In order to estimate the effect of these discrepancies on the firing probability of a neuron receiving input from the network, the cumulative probability distributions are plotted in figures 6D-F. It can be seen that the underestimation of the incidence of intermediate cluster-sizes by the maximal entropy calculation is compensated by the overestimation of the incidence of clusters with a size between 100 and 140. For values of Θ' smaller than 130, the largest deviation

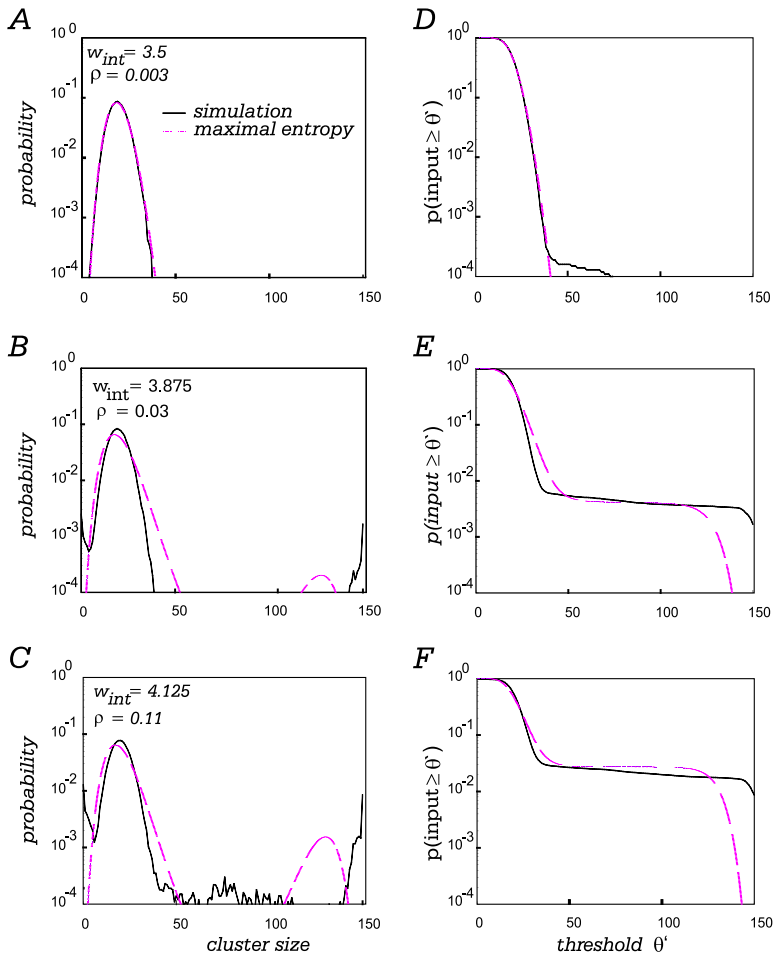

Figure 6: Comparison of the probability of N-clusters in the network simulation, and their probability based on maximal entropy. (A-C) Probabilities of N-clusters in the simulation are shown as solid curves for three values of the synaptic weights. Dashed line: the corresponding maximal entropy probability distributions with the same firing probability and pair-wise correlation. (D-F) Relationship between firing probability (ordinate) and threshold (abscissa) of a neuron with an integration window of 2 msec that receives input from all 150 neurons of the network (cumulatives of A-C).

is about a factor of 2. This approximation is reasonable, since the maximal entropy calculation depends on only 2 parameters, which are estimated from the network activity: the firing probability and the pair-wise correlation. Nevertheless, large deviations occur for values of Θ' larger than 130. However, these high threshold values are physiologically implausible, because they would imply that a post-synaptic neuron would only fire when almost every input is active

within a narrow time-window.

4.4 The effects of varying bin-width on the maximal entropy distribution

The results of the maximal entropy calculation described so far, were obtained with a bin-width of 2 msec, which equals the minimal inter-spike interval. Figure 7 shows the dependence of the N-cluster distribution and the maximal entropy estimation on the bin-width. Spike trains obtained with a synaptic weight (w_{ij}) of 4.125 were re-binned, using bin-widths of $0.5, 1.0, 2.0$ and 4.0 msec. The rebinning process influenced the pair-wise correlation, and f_1, the probability of firing in a time-bin (eq. 8). Smaller bin-widths reduce f_1. This results in a leftward shift in the location of the first peak in the distribution of cluster-sizes, which represents stochastic activity between bursts (compare Fig. 7A, B to Fig. 7C). A second effect of reducing the binwidth is a disappearance of clusters with sizes larger than 110. This disappearance is related to the refractory period, which prohibits neurons from firing in consecutive bins during population bursts. Spikes fired by different neurons during these bursts are therefore divided between successive time-bins, and no bins remain in which all neurons fire simultaneously. This modification of the distribution of cluster-sizes is not captured by the maximal entropy estimation, which underestimates the incidence of clusters with sizes between 20 and 110, and grossly overestimates the incidence of clusters with a size larger than 130 (Fig. 7A-B). In other words, the refractory period adds structure to the spike trains, which therefore have less than maximal entropy. When a bin-width is used that is longer than the refractory period, this additional structure is lost, and the maximal entropy calculation may provide a reasonable estimate of the distribution of cluster sizes.

For bin-sizes that are longer than the minimal inter-spike interval, there are time-windows, during which individual cells fire more than a single spike. It is possible, in principle, to adapt the maximal entropy estimation to this situation. One approach, in the case of a bin-size that may include 2 spikes, is to compute the probability that N neurons fire once, and M neurons fire twice in a bin, for each combination of N and M. Unfortunately, this increases the number of variables that should be calculated from 151 (D_0 to $D_1 50$) to more than 10.000. The computational requirements increase further if 3 or more spikes can occur in a single bin. Therefore, we took an alternative approach in which the state of a neuron was labeled "off", in the case of no spike, and labeled "on" in the case of one or more spikes within a time bin. This keeps the computational requirements within bounds, but at the cost of losing spikes in the rebinning process. Figure 7D compares the distribution of cluster sizes to the maximal entropy distribution for a bin-width of 4 msec. The experimental distribution is largely described by a broad first peak, which is shifted to the right, and a very narrow second peak at a cluster size of 150. An increase in f_1 also shifts the first peak of maximal entropy distribution to the right. However, larger values of f_1 are also associated with a leftward shift of the second peak in the distribution of maximal entropy, as was discussed in section 3. This causes large deviations

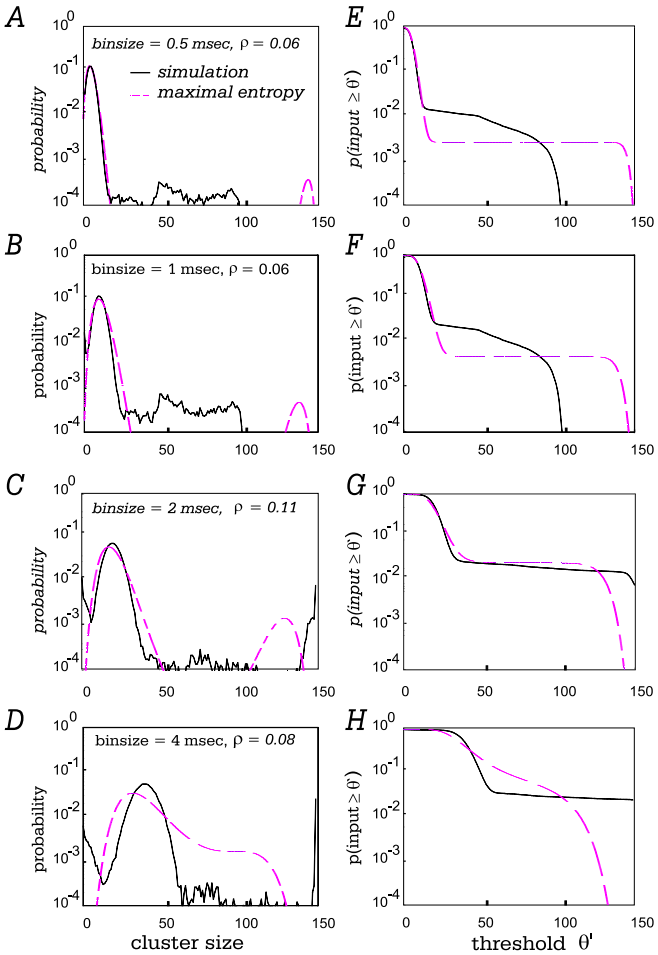

Figure 7: (A-D) Comparison of the maximal entropy distribution to the distribution of N-clusters in the simulation, for bin-sizes ranging from 0.5 msec to 4 msec. The quality of the maximal entropy approximation depends on bin-size, i.e. on what is considered coincident. In our simulations, a bin-size of 2 msec yields the best approximation (C). For smaller binsizes (A: 0,5 msec, B: 1 msec), the actual and predicted distributions deviate considerably. (E-H) The impact of these deviations on a post-synaptic neuron with threshold Θ'.

for cluster sizes between 60 and 120, the incidence of which is overestimated by the maximal entropy estimation. Moreover, the narrow peak at a cluster size of 150 is absent in the distribution of maximal entropy. In summary, reasonable estimates of the N-cluster distribution are only obtained for a bin width, that is equal to the minimal interspike interval. In the discussion we will address the question whether these results can be generalized to other network architectures.

4.5 The effects of network scaling

In order to investigate how the network behavior and the goodness of fit of the maximal entropy distribution depends on network size, simulations were run with networks composed of 75, 150 and 200 neurons. Figure 8A shows the cumulative distribution of N-clusters for a network of 75 neurons with a w_{int} of 3.55, which exhibited a pair-wise correlation of 0.07. It should be noted that w_{int} represents the sum of the synaptic weights w_{ij} of all inputs converging onto a single neuron (section 4.1). When the network size is doubled, each cell receives input from twice as many neurons, and the strength of the individual synapses w_{ij} was reduced accordingly, in order to maintain a constant value of w_{int}. Nevertheless, a doubling of the network size resulted in a clear leftward shift of the distribution of N-clusters, and a reduction of the pair-wise correlation to 0.003. This indicates that a constant value of w_{int} is not sufficient to guarantee a qualitatively similar network behavior when the network size is increased. Indeed, a larger number of synapses with reduced weight impinging on a network unit results in a reduction of the fluctuations in the input, as long as the network is in the stochastic mode. This reduces the probability of bursts in the network [28]. Tsodyks & Sejnowski [28] have suggested that the variance in the input to network units may be kept approximately constant, by reducing the release probability of the synapses, rather than reducing their strength w_{ij}, when network size is increased. Figure 8C shows the cumulative distribution of N-clusters for a network of 150 neurons, in which the effective input strength was kept constant by reducing the release probability to 50%. The pair-wise correlation was 0.12, and the cumulative distribution of N-clusters was qualitatively similar to that of the smaller network, in accordance with [28]. Figure 8D shows a similar result for a network composed of 200 neurons. In all cases the estimate based on maximal entropy was reasonable (dashed lines in Fig. 8A-D), which indicates that the quality of the maximal entropy calculation does not depend critically on the size of the network.

5 Discussion

In most physiological studies on the synchronization behavior of cortical neurons, recordings are obtained from pairs of neurons, or pairs of cell clusters. The present results illustrate that these data can only supply limited information about the probability of higher order events, like the probability that, for example, 30 or more neurons project that to a target neuron fire simultaneously. As an approximation for the probability of higher order events, we used the distribution of maximal entropy. This method provides the most unstructured distribution, given the constraints supplied by the firing probabilities and the pair-wise correlations. The N-cluster distribution with maximal entropy for a homogeneous network without higher-order correlations exhibits two peaks. The first peak represents stochastic activity, and resembles a binomial distribution. This peak also occurs in the absence of correlations. If the firing probability

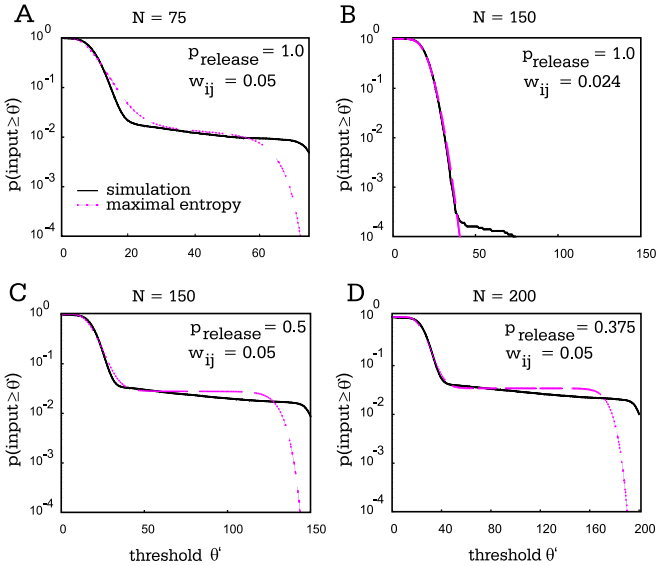

Figure 8: Effect of scaling the network size. (A) Continuous line, probability of N-clusters with a minimal size of Θ', as a function of Θ', for a network with 75 neurons. Dashed line, distribution of maximal entropy with the same firing probability and pair-wise correlation. Parameters of the simulation: $w_{ij} = 4.73?10^{-2}$, $\rho = 0.07$. (B) N-cluster distribution for a larger network with 150 neurons. The total input converging on a neuron, w_{int}, was kept constant by reducing w_{ij} to $2.37?10^{-2}$. Nevertheless, the distribution of N-clusters was shifted to the left, and the pair-wise correlation was reduced to 0.003. (C) Same as B, but the effective w_{int} was kept constant by reducing the release probability to 50% rather than by reducing w_{ij}. The synaptic strength w_{ij} was $4.73?10^{-2}$, and $\rho = 0.12$. (D) N-cluster distribution for a network of 200 neurons, with a ρ?of 0.15. The synaptic weight w_{ij} was identical to that in (A), but release probability was reduced to 37.5%. Note the similarity of the distributions in (C) and (D) to the distribution in (A).

is moderate (< 0.25), a second peak occurs at a relatively large cluster size. The magnitude of this second peak depends on the strength of the pair-wise correlation. Thus, an absence of *higher order* correlations dictates that the network should generate coincidences in a number of tightly synchronized network bursts. The N-cluster distribution observed in the simulations also exhibited two peaks, although the position of the second peak was different from the second peak in the maximal entropy distribution. Questions about the generality of these results, and in particular, about their dependence on the details of the network implementation will have to await further experimentation.

Previous studies on the impact of correlations among neurons that project to a common post-synaptic target have used N-cluster distributions with a

drastically different shape (e.g. [17, 18]). In these studies correlations were introduced in the input by forcing a subset of the pre-synaptic neurons to fire in perfect synchrony, but independently of the other pre-synaptic neurons. Since the resulting N-cluster distribution is relatively devoid of highly synchronous events, and has less than maximal entropy, the generality of the results obtained in these earlier studies may also be limited.

It is obvious that network bursts in which a large number of neurons participate are rather effective in driving a post-synaptic neuron, especially in the case of a high post-synaptic threshold. Indeed, a recent study [8] uncovered tight correlations, with a peak-width in the cross-corellogram of less than 1 msec, among pairs of neurons in the LGN. In cases in which the LGN neurons projected to a common cortical neuron, the impact of synchronous events was stronger than that of asynchronous spikes. We observed an orderly relationship between the incidence of highly synchronous network bursts and the pair-wise correlation. Relatively small increases in the pair-wise correlation, which are not physiologically implausible, may raise the incidence of highly synchronous events, and thereby the firing rate of a post-synaptic cell, by more than an order of magnitude (Fig. 5C,E).

We will now discuss the limitations of the maximal entropy estimation, and the way in which these limitations may be overcome by future studies. First, the quality of the approximation by the distribution of maximal entropy exhibited a strong dependence on the choice of the binsize. Reasonable results were only obtained for binsizes larger than the refractory period of the neurons. If the binsize was smaller than the refractory period, structure was added to the N-cluster distribution, which therefore had less than maximal entropy. This limitation is a direct consequence of the way in which the entropy was defined. The entropy was determined by the distribution of N-clusters within individual time-bins, and was independent of the order of N-clusters in successive bins. The equivalent situation for a cross-correlation study would be to only calculate the center bin of the cross-correlation functions, i.e. only the probability that two neurons fire at exactly the same time. A possible extension of the method would be to reformulate entropy in order to include 2nd order correlations with a time delay (i.e. the non-central bins in the auto- and cross-correlation functions) in the calculation. We wish to remark, however, that such an extension is likely to result in a substantial increase in the number of variables and equations.

Second, the quality of approximation by the maximal entropy distribution was also degraded if a binwidth was chosen that was larger than the minimal interspike interval (larger than 2 msec in our simulations). In this situation, multiple spikes of a single neuron occurred within a single time bin. The membrane time constant of the postsynaptic neuron provides a natural temporal window over which input is integrated, i.e. the natural coincidence window. At present, the value of the effective membrane time constant in cortical neurons is a topic of considerable debate (e.g. [25, 13, 14]). A coincidence window of 2 msec, as we used, is presumably at the lower end of the physiologically plausible range. However, it seems likely that an increase in the width of the bin in which spikes are considered to be coincident to 5 or even 10 msec may be unproblematic in

the case of cortical neurons. Typical peak firing rates of cortical neurons are around 100 Hz, which implies that the probability of multiple spikes within a 5 or 10 msec bin will still be rather low. Without multiple spikes, it is easy to see that widening the time-window than merely corresponds to a simple re-scaling of the time-axis.

Alternatively, the calculation of the maximal entropy may be adapted to allow for multiple spikes in a single bin, as was discussed in section 4.4. If a very large binsize is chosen, the distribution of the number of spikes fired by a single neuron in a bin may approach a normal distribution. In this case it is relatively easy to calculate the distribution of N-clusters if the only correlations are of second order.

Third, the method according to which we derived the distribution of maximal entropy is only valid for a homogeneous network. For a non-homogeneous network, the number of variables grows exponentially with the number of neurons (section 3). We have, for example, incorporated a single inhibitory neuron in our network, which received input from all excitatory cells and provided strong inhibitory feedback (data not shown). The inhibitory neuron added structure to the spike trains by curtailing population bursts. We did not explicitly incorporate the firing pattern of the inhibitory neuron in our calculations, which caused a considerable discrepancy between the maximal entropy distribution and the actual distribution of N-clusters.

These limitations, taken together, imply that it will be rather problematic to predict the probability of highly synchronous events from firing rates and pair-wise correlations in physiological experiments. In a physiological study on higher order events among neurons of the frontal cortex, Martignon *et al.* [19] obtained discrepancies between the probability of actually occurring events and their probability predicted by entropy maximization. Unfortunately, even in this study, in which simultaneous recordings from 6 neurons were studied, sampling problems occurred which were caused by the exponentially growing number of spike configurations and the limited recording time available during such experiments. Another, presumably fruitful way in which the probability of highly synchronous events can be estimated is by direct, in vivo measurements of the distribution of the post-synaptic potential. A comparison of the post-synaptic potential to the size of individual EPSPs, with and without blockade of inhibition, could provide valuable insight in the degree of synchronicity among neurons that project to a common target cell.

6 Acknowledgements

We thank Holger Bosch for helpful comments and a critical reading of the manuscript. The research of P.R.R. was funded by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

References

[1] Wolf Singer and Charles M. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.*, 18:555–586, 1995.

[2] W. Singer. Development and plasticity of cortical processing architectures. *Science*, 270:758–764, 1995.

[3] A.K. Engel, P. König, A.K. Kreiter, T.B. Schillen, and W. Singer. Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends Neurosci.*, 15:218–226, 1992.

[4] R. Eckhorn, R. Bauer, W. Jordan, M. Broch, W. Kruse, M. Munk, and H.J. Reitboeck. Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biol. Cybern.*, 60:121–130, 1988.

[5] C.M. Gray, P. König, A.K. Engel, and W. Singer. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338:334–337, 1989.

[6] M. Konishi. Deciphering the brain's codes. *Neural Computation*, 3(1):1–18, 1991.

[7] P.R. Roelfsema, A.K. Engel, P. König, and W. Singer. Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature*, 385:157–161, 1997.

[8] J-M. Alonso, W.M. Usrey, and R.C. Reid. Precisely correlated firing in cells of the lateral geniculate nucleus. *Nature*, 383:815–819, 1996.

[9] M. Abeles, H. Bergman, E. Margalit, and E. Vaadia. Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. Neurophysiol.*, 70:1629–1658, 1993.

[10] C. von der Malsburg. The correlation theory of brain function. *Internal Report 81-2, Max-Planck-Institute for Biophysical Chemistry, G ttingen, Germany*, 1981.

[11] W. Bair, Ch. Koch, W. Newsome, and K. Britten. Power spectrum analysis of bursting cells in area mt in the behaving monkey. *J. Neurosci.*, 14:2870–2892, 1994.

[12] M.N. Shadlen and W.T. Newsome. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.*, 4:569–579, 1994.

[13] P. König, A.K. Engel, and W. Singer. Integrator or coincidence detector? the role of the cortical neuron revisited. *Trends NeuroSci.*, 19:130–137, 1996.

[14] M.N. Shadlen and W.T. Newsome. Is there a signal in the noise? *Curr. Opin. Neurobiol.*, 5:248–250, 1995.

[15] W.R. Softky and Ch. Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsp's. *J. Neurosci.*, 13:334–350, 1993.

[16] W.R. Softky. Simple codes versus efficient codes. *Curr. Opin. Neurobiol.*, 5:239–247, 1995.

[17] Öjvind Bernander, Christof Koch, and Marius Usher. The effect of synchronized inputs at the single neuron level. *Neural Computation*, 6(4):622–641, 1994.

[18] Venkatesh N. Murthy and Eberhard E. Fetz. Effects of input synchrony on the firing rate of a three-conductance cortical neuron model. *Neural Computation*, 6(6):1111–1126, 1994.

[19] L. Martignon, H. von Hasseln, S. Gr n, A. Aertsen, and G. Palm. Detecting higher-order interactions among the spiking events in a group of neurons. *Biol. Cybern.*, 73:69–81, 1995.

[20] I. Csiszar. I-divergence geometry of probability distributions and minimization problems. *Ann. Probab.*, 3:146–158, 1975.

[21] D.V. Gokhale and S. Kullback, editors. *The information in contingency tables.* Dekker, New York, 1978.

[22] W.H. Press, B.P. Flanney, S.A. Teukolsky, and W.T. Vetterling, editors. *Numerical Recipes: the art of scientific computing.* Cambridge University Press, Cambridge, 1986.

[23] J. Deppisch, H-U. Bauer, T. Schillen, P. König, K. Pawelzik, and T. Geisel. Alternating oscillatory and stochastic states in a network of spiking neurons. *Network: Computation in Neural Systems*, 4:243–257, 1993.

[24] R.J. MacGregor and R.M. Oliver. A model for repetitive firing in neurons. *Kybernetik*, 16:53–64, 1974.

[25] Ö. Bernander, R.J. Douglas, K.A.C. Martin, and C. Koch. Synaptic background activity influences spatiotemporal integration in single pyramidal cells. *Proc. Natl. Acad. Sci. USA*, 88:11569–11573, 1991.

[26] Peter König, Andreas K. Engel, Pieter R. Roelfsema, and Wolf Singer. How precise is neuronal synchronization? *Neural Computation*, 7(3):469–485, 1995.

[27] M.S. Livingston. Oscillatory firing and interneuronal correlations in squirrel monkey striate cortex. *J. Neurophysiol.*, 75:2467–2485, 1996.

[28] M. V. Tsodyks and T.J. Sejnowski. Rapid state switching in balanced cortical network models. *Network: Computation in Neural Systems*, 6:111–124, 1995.

A Maximizing the entropy in a homogenous network

For N neurons three equations are derived from the pair-wise correlation, the firing probability and the fact that all probabilities should add up to 1. First, the probabilities should add up to 1:

$$1 = \sum_{i=0}^{N} \binom{N}{i} \cdot D_i \tag{19}$$

Second, the firing probability f_1 is fixed. For example, in the case that $N = 7$ it is easy to see that f_1 equals:

$$f_1 = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1}\sum_{l=0}^{1}\sum_{m=0}^{1}\sum_{n=0}^{1} G_{1ijklmn} = \sum_{i=1}^{7} \binom{6}{i-1} \cdot D_i \tag{20}$$

In the general case of $N = n$ this reads:

$$f_1 = \sum_{i=1}^{N} \binom{N-1}{i-1} \cdot D_i \tag{21}$$

Third, the pair-wise correlations are fixed, and this implies that the probability that two neurons fire simultaneously is also fixed. Analogously to (8) we find for f_2:

$$f_2 = \sum_{i=2}^{N} \left(\begin{array}{c} N-2 \\ i-2 \end{array} \right) \cdot D_i \tag{22}$$

These equations can be rewritten as:

$$\begin{aligned} D_0 &= 1 - \sum_{i=1}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) D_i \\ &= 1 - nD_1 - n \left(\begin{array}{c} n \\ 2 \end{array} \right) D_2 - \sum_{i=3}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) D_i \end{aligned} \tag{23}$$

$$\begin{aligned} D_1 &= f_1 - \sum_{i=2}^{n} \left(\begin{array}{c} n-1 \\ i-1 \end{array} \right) D_i \\ &= f_1 - [n-1]\, D_2 - \sum_{i=3}^{n} \left(\begin{array}{c} n-1 \\ i-1 \end{array} \right) D_i \end{aligned} \tag{24}$$

$$D_2 = f_2 - \sum_{i=3}^{n} \left(\begin{array}{c} n-2 \\ i-2 \end{array} \right) D_i \tag{25}$$

Solving for D_0 , D_1 and D_2:

$$\begin{aligned} D_0 &= 1 - nf_1 + n(n-1)f_2 - \left(\begin{array}{c} n \\ 2 \end{array} \right) f_2 + \\ &\quad + \left[\left(\begin{array}{c} n \\ 2 \end{array} \right) - n(n-1) \right] \cdot \sum_{i=3}^{n} \left(\begin{array}{c} n-2 \\ i-2 \end{array} \right) D_i + \\ &\quad + n \sum_{i=3}^{n} \left(\begin{array}{c} n-1 \\ i-1 \end{array} \right) D_i - \sum_{i=3}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) D_i \end{aligned} \tag{26}$$

$$D_1 = f_1 - (n-1)f_2 + (n-1) \sum_{i=3}^{n} \left(\begin{array}{c} n-2 \\ i-2 \end{array} \right) D_i - \sum_{i=3}^{n} \left(\begin{array}{c} n-1 \\ i-1 \end{array} \right) D_i \tag{27}$$

$$D_2 = f_2 - \sum_{i=3}^{n} \left(\begin{array}{c} n-2 \\ i-2 \end{array} \right) D_i \tag{28}$$

With the entropy defined as:

$$S = \sum_{i=0}^{N} \left(\begin{array}{c} N \\ i \end{array} \right) \cdot D_i \ln(D_i) \tag{29}$$

This is a concave function (which is easily verified), and therefore it has a single, unique maximum. To obtain the maximum, we calculate the derivative to D_i:

$$\frac{\partial S}{\partial D_i} = 0 \qquad \text{for } i = 3 \ldots N \tag{30}$$

$$
\begin{aligned}
\frac{\partial S}{\partial D_i} &= -\left\{\left[\binom{n}{2} - n(n-1)\right]\binom{n-2}{i-2} + n\binom{n-1}{i-1} - \binom{n}{i}\right\}\ln(D_0) + \\
&\quad -n\left\{(n-1)\binom{n-2}{i-2} - \binom{n-1}{i-1}\right\}\ln(D_1) + \\
&\quad + \binom{n}{2}\binom{n-2}{i-2}\ln\ (D_2) + \\
&\quad - \binom{n}{i}\ln\ (D_i) = 0
\end{aligned} \tag{31}
$$

We note that the equations (26-28) are linear in D_i, and that a concave function remains concave on a linear subspace. In other words: equations (26-28) and (30) taken together, also have a unique solution. Equation (31) can be rewritten as:

$$
\begin{aligned}
\ln(D_i) &= -(-\frac{1}{2}i+1)(i-1)\ln(D_0) - i(i-2)\ln(D_1) + \\
&\quad +\frac{1}{2}i(i-1)\ln(D_2) \qquad \text{for } i = 3\ldots N
\end{aligned} \tag{32}
$$

Inserting these values for D_i into equations (26-28) yields three equations with three unknowns and a unique solution which is found using the Newton-Raphson method as described in [22]. Use of this algorithm in Maple V v4.0 allowed us to solve the equations for up to 150 neurons in about 2 hours on a Pentium 150Mhz.