Stimulus segmentation in a stochastic neural network with exogenous signals

Sybert Stroeve, Bert Kappen, Stan Gielen

RWCP, Theoretical Foundation SNN, University of Nijmegen
6525 EZ Nijmegen, The Netherlands
E-mail: sybert@mbfys.kun.nl, bert@mbfys.kun.nl, stan@mbfys.kun.nl

Abstract

Segmentation by synchrony of firing is investigated in stochastic neural networks with binary neurons. The network is trained by presenting sparsely coded patterns with a Hebbian-type learning rule. Retrieval of these patterns and synchrony of firing is investigated by presenting one or multiple patterns simultaneously to the network. For stimuli consisting of several superimposed patterns the model correctly predicts a high covariance of firing (indicating synchrony) for neurons which are excited by the same pattern in the stimulus. The model gives a negative covariance for neurons, which are stimulated by different patterns (indicating absence of synchrony). To obtain useful covariance levels, noise levels should neither be too large, since this would induce complete random firing, nor be too small, since this would minimize the chance to switch between firing assemblies. Furthermore, the strength of the multi-pattern input may neither be too large, as this would overrule the memory function of the network, nor so small that the desired firing assemblies are never attained.

1 Introduction

In a series of experiments [8],[10] synchronization of spike patterns of neurons in the cortex was observed. In contrast to the traditional view, that emphasizes the use of firing rate as an information carrier in the central nervous system, these results show that temporal coding mechanisms should also be taken into account. Recently, other studies [6] have suggested that these results are better understood as the result of fixed lateral connectivity than as a mechanism for binding of local features for texture segregation.

Theoretical studies dealing with synchronous firing mostly considered oscillator models or integrate-and-fire model neurons [2],[3],[7]. In particular, Ritz et al. [7] showed that for suitable delays, pattern segmentation and collective oscillations could be achieved with a spiking neuron model. As an alternative for the integrate-and-fire models Kappen [4] extended the traditional mean firing rate analysis [1] of stochastic neural networks (SNN) by applying linear response theory to obtain a first-order approximation of the correlation between the firing patterns of binary, stochastic neurons. In this work this approach was used for the analysis of long range correlation between neurons with short-range interactions as a function of the coherence of the external (sensory) input and the lateral coupling strength.

In this paper we will analyse whether and under which conditions the correlation of the firing of stochastic neurons in a network, which is trained with sparsely coded patterns, can be used to segment external stimuli which are a superposition of several of these patterns.

2 Methods

2.1 Dynamics of the stochastic neural network

We use binary stochastic neurons with states $s_i \in \{0,1\}$ and a probability of firing $f(v_i)$ which depends on the local field v_i

and a noise parameter β

$$s_i = \begin{cases} 1 \text{ with prob. } f(v_i) \\ 0 \text{ with prob. } 1 - f(v_i) \end{cases} \quad (1)$$

$$f(v_i) = 1/(1 + \exp(-\beta v_i)) \quad (2)$$

If N neurons are arranged in a network with connection strengths w_{ij}, thresholds θ_i and exogenous input signals h_i, then the local potential is

$$v_i = \sum_{j=1}^{N} w_{ij} s_j + \theta_i + h_i \quad (3)$$

The exogenous signal h_i is a real number which is linked to an external stimulus $u_i \in \{0, 1\}$ by a stimulus gain g

$$h_i = g u_i \quad (4)$$

The states of the neurons are updated sequentially following Eqs. 1 - 4.

The firing behaviour of the networks is characterized by the mean firing $m_i = \langle s_i \rangle$ and the covariance $C_{ij} = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$. The firing statistics used in this paper are the ensemble averages over 100 trials, each of which is a time average over 200 complete sequential update cycles. Each time average is calculated starting from a stationary state, reached after an initial period of 50 update cycles. For example, estimated mean firing is

$$\hat{m}_i = \frac{1}{100} \sum_{k=1}^{100} \frac{1}{200} \sum_{t=51}^{250} s_i^k(t)$$

The initial state $\mathbf{s}^k(0)$ of a trial follows from Eq. 1 using the local fields $v_i = \theta_i$. Since the thresholds inhibit firing, $\mathbf{s}(0) \to \mathbf{0}$ for low noise levels ($\beta \to \infty$).

2.2 Choice of weights and thresholds

Consider L sparse patterns with bits $\xi_i^\mu \in \{0, 1\}$ chosen from a Bernoulli probability distribution with mean $a < 0.5$. The weights are chosen according to the covariance learning rule

$$w_{ij} = \begin{cases} \frac{1}{N} \sum_{\mu=1}^{L} (\xi_i^\mu - a)(\xi_j^\mu - a) & i \neq j \\ 0 & i = j \end{cases} \quad (5)$$

which is a straightforward extension of the Hebb rule to correct for the mean activity a. Tsodyks and Feigelman [9] showed that for $\{0, 1\}$-coding the storage capacity diverges as $\alpha_c \sim 1/(a \ln a)$ for $a \downarrow 0$ which is equal to the optimal storage capacity. This statement was generalized by Vicente and Amit [11] who adapted the coding scheme to achieve optimal storage for general a. In this work we will use the more natural $\{0, 1\}$-coding and adapt the thresholds.

The thresholds are chosen according to

$$\theta_i = -b \sum_{j=1}^{N} w_{ij} - c \quad (6)$$

where the parameters b and c have to be chosen appropriately (see below). Using Eqs. 5 and 6 to encode the weights and thresholds, the local potential v_i due to a state $\mathbf{s} = \xi^\nu$ can be divided in a signal part S, which favors the present state ξ_i^ν, and a noise part R due to the storage of the other patterns. The noise R due to all patterns except for pattern ν is

$$R = \frac{1}{N} \sum_{j \neq i}^{N} \sum_{\mu \neq \nu}^{L} (\xi_i^\mu - a)(\xi_j^\mu - a)(\xi_j^\nu - b)$$

Since the patterns are uncorrelated and $\langle \xi_j^\mu - a \rangle = 0$, the noise term R has a zero mean. Straightforward analysis shows that for large N the variance of the noise is

$$\langle R^2 \rangle \approx \frac{L}{N} a^2 (1 - a)^2 (a - 2ab + b^2)$$

The noise is minimal if $b = a$, which gives

$$\langle R^2 \rangle \approx \frac{L}{N} a^3 (1 - a)^3$$

The signal S for large values of N is

$$S \approx (\xi_i^\nu - a) a (1 - a) - c + h_i$$

The global threshold parameter c has to be adjusted for a desired weighting between the probability of attaining a positive potential for a high bit $p(v_i > 0 | \xi_i^\nu = 1)$ and the probability of attaining a negative potential for a low bit $p(v_i < 0 | \xi_i^\nu = 0)$. If these probabilities are weighted equally and when an external stimulus $u_i = \xi_i^\nu$ is present, c should be chosen

$$c = a^3 - 1.5a^2 + 0.5a + 0.5g$$

This choice implies equal expected signal amplitudes $|S|$ for both $\xi_i^\nu = 0$ and $\xi_i^\nu = 1$

$$|S| = a(1 - a)/2 + g/2$$

The noise R is unaffected by the stimulus. Therefore, the signal-to-noise-ratio (SNR) can be adjusted at will by changing g. However, the memory function of the network and correlated firing between neurons are lost for large stimulus gains, since the effect of the external stimulus now overrules the contributions of neurons in the network. In this paper we use the gain

$$g = g_c a(1-a)$$

with g_c a gain coefficient. If $g_c = 1$ the contribution of all other neurons equals the contribution by the external stimulus.

The calculations presented in this paper refer to a network with $N = 100$ neurons and $L = 10$ patterns. All patterns consist of 10 high bits, 90 low bits and have an overlap of 1 high bit with each other pattern, such that $\langle \xi_i^\mu \rangle = a = 0.1$, $\langle \xi_i^\mu \xi_j^\nu \rangle = 0.01$ and all higher order correlations are zero.

2.3 Mean field analysis

If the SNN with dynamics specified by Eqs. 1-3 and weights specified by Eq. 5 has reached equilibrium, the probability distribution $p(\mathbf{s})$ of the states of the SNN is given by the Boltzmann distribution:

$$\begin{aligned} p(\mathbf{s}) &= \frac{1}{Z}\exp(-\beta E(\mathbf{s})) \\ Z &= \sum_{\{\mathbf{s}\}} \exp(-\beta E(\mathbf{s})) \\ E(\mathbf{s}) &= -\tfrac{1}{2}\sum_{i,j} w_{ij} s_i s_j - \sum_i (\theta_i + h_i) s_i \end{aligned}$$

where Z is the partition function over all possible states, and $E(\mathbf{s})$ is the energy associated with state $\mathbf{s}$.

Since exact calculation of the partition function is often not computationally feasible, one has to resort to approximations of the probability distribution. In this paper we use the mean field (MF) formalism [5]. It assumes a decoupled probability distribution and optimizes the mean fields of the decoupled model such that the free energy of the decoupled model best approximates the free energy of the original uncoupled distribution. The MF equation for a $\{0,1\}$-neuron is

$$m_i = f\Big(\sum_{j=1}^{N} w_{ij} m_j + \theta_i + h_i\Big) \qquad (7)$$

with $m_i = \langle s_i \rangle$. Under the mean field formalism an estimate of the covariance can be obtained by the linear response theorem, which states that

$$C_{ij} = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle \approx \frac{1}{\beta}\frac{\partial m_i}{\partial \theta_j}$$

Differentiation of the MF equations leads to

$$(\mathbf{C}^{-1})_{ij} = \frac{\delta_{ij}}{m_i(1-m_i)} - \beta w_{ij}$$

This paper uses a first-order approximation of the covariance matrix which avoids the matrix inversion

$$C_{ij} = \begin{cases} m_i(1-m_i) & i = j \\ \beta w_{ij} m_i(1-m_i) m_j(1-m_j) & i \neq j \end{cases}$$

3 Results

3.1 Single-pattern stimulus

As a first step towards the analysis of the segmentation of a multi-pattern stimulus, the neural activity in the network with external input by one pattern $\mathbf{u} = \xi^1$ is considered. Figure 1 shows characteristics of the firing behaviour as a function of the noise parameter β for different groups of neurons. First we consider the simulated Glauber dynamics (solid lines). For high noise levels (small β) the mean firing m of all neurons tends to 0.5. For smaller noise levels the mean firing of all non-stimulated neurons rapidly decreases to zero and the mean firing of the externally stimulated neurons increases. The peak in the covariance C_e indicates that the stimulated neurons have an increased probability to fire synchronously.

The dashed lines in Fig. 1 indicate the solutions of the mean field (MF) equations for the same stimulus. The mean firing rates m and variances σ are in reasonable agreement with the simulation results. The covariance levels C are underestimated, but the shapes of the graphs are similar.

Using the MF approximation, the effect of gain and noise variation on the firing statistics is depicted in Fig. 2. For small stimulus gains no enhanced firing rate is attained and a sharp transition exists for gains which induce a high mean firing rate. Highest covariance levels are attained for medium noise levels ($\beta \approx 50$) and gains near the transition point.

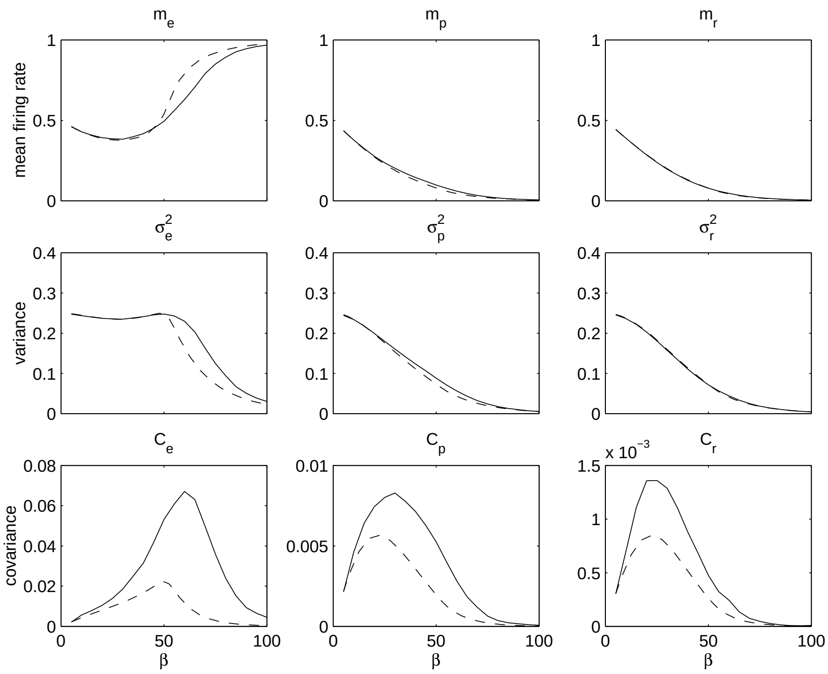

Figure 1: *Mean firing m (upper row), variance σ^2 (middle row) and covariance C (bottom row) as function of the noise parameter β for an external stimulus $\mathbf{u} = \xi^1$ and a stimulus gain $g_c = 0.2$. Left column: neurons i excited in pattern ξ^1 ($\xi_i^1 = 1$). Middle column: $\xi_i^\mu = 1$ for $\mu \neq 1$. Right column: $\xi_i^\mu = 0$ for all μ. Solid lines: simulation results of the Glauber dynamics. Dashed lines: solution of the MF analysis.*

3.2 Multi-pattern stimulus

The statistics of the converged Glauber dynamics for a stimulus consisting of two patterns are depicted in Fig. 3. For a range of noise levels the neurons with $\xi_i^\mu = 1$ in one of the two patterns (but not in both) exhibit a mean firing of about a half. Further, in the middle of this range there exists a large positive covariance between neurons in one pattern and a clear negative covariance between neurons in the two different patterns. This signifies that for these noise levels the neurons in one pattern have a high probability to fire synchronously, while the neurons in two patterns mostly fire asynchronously.

In contrast to the MF solutions of single-pattern stimuli no reasonable results could be obtained with the MF analysis for multi-pattern stimuli.

Simulation results of the Glauber dynamics similar to those shown in Fig. 3 can easily be obtained for stimuli consisting of the superposition of three or more patterns. An insightful example is shown for a 3-pattern stimulus in Fig. 4. It is clear that binding of neurons in one encoded pattern and segmentation of neurons in different encoded patterns is achieved using the appropriate noise levels and stimulus gain. It is worth mentioning that the network can respond in two ways. Most of the time, one pattern dominates such that only neurons, which are encoded in that pattern, are firing in synchrony (such as in the time interval 10-40 in Fig. 4). Occasionally, multiple patterns are represented simultaneously (like in the time interval around 60 in Fig. 4). On average, the covariance between the firing of neurons in different patterns is negative.

4 Discussion

The stochastic neural network (SNN) with an external stimulus is an extension of the traditional autoassociative SNN. The external stimulus offers a good mechanism to change the state of the network. It was shown in this paper that for suitable noise levels an external stimulus induces a sub-maximum average firing rate and facilitates segmentation of a stimulus by analysis of the correlation of the firing events. In particu-

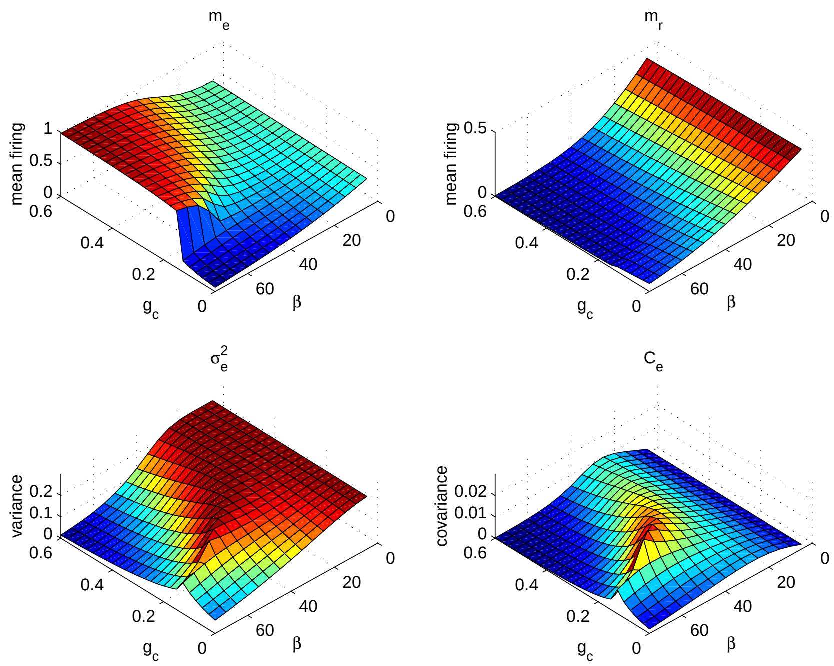

Figure 2: *Mean firing m_e, variance σ_e and covariance C_e of stimulated neurons and mean firing m_r of non-stimulated neurons as function of the noise parameter β and stimulus gain g_c.*

lar, a positive covariance between the neurons in one pattern and a negative covariance between neurons with $\xi_i^\mu = 1$ for different patterns μ was found. Noise is required in the SNN with external stimulus such that the network can jump into and between firing assemblies. To attain useful covariance levels the stimulus gain should neither be too large, since this would override the memory function of the network, nor too small, since this would minimize the probability to attain the desired firing assembly.

A reasonable correspondence between the MF analysis and the simulation of the Glauber dynamics could be attained for stimuli with one pattern. Often poor results were attained in the mean field approximation for stimuli with multiple patterns. This indicates a violation of the assumptions which underlie the MF approximation. More powerful approximation methods therefore seem to be required to efficiently calculate the statistics of the dynamics of the firing behaviour in these cases.

References

[1] Amit DJ, Gutfreund H, Sompolinsky H (1985). Physical Review A 32:1007

[2] Ernst U, Pawelzik K, Geisel T (1998). Physical Review E 57:2150

[3] Gerstner W, Hemmen JL van, Cowan JD (1996). Neural Computation 8:1653

[4] Kappen HJ (1997). Physical Review E 55:5849

[5] Kappen HJ, Rodriguez (1998). Neural Computation 10:1137

[6] Lamme VAF, Spekreijse H (1998). Nature 396:362

[7] Ritz R, Gerstner W, Fuentes U, Hemmen JL van (1994). Biological Cybernetics 71:349

[8] Singer W, Gray CM (1995). Annual Review of Neuroscience 18:555

[9] Tsodyks MV and Feigelman MV (1988). Europhysics Letters 6(2):101

[10] Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, Aertsen A (1995). Nature 373(9):515

[11] Vicente CJP and Amit DJ (1989). Journal of Physics A 22:559

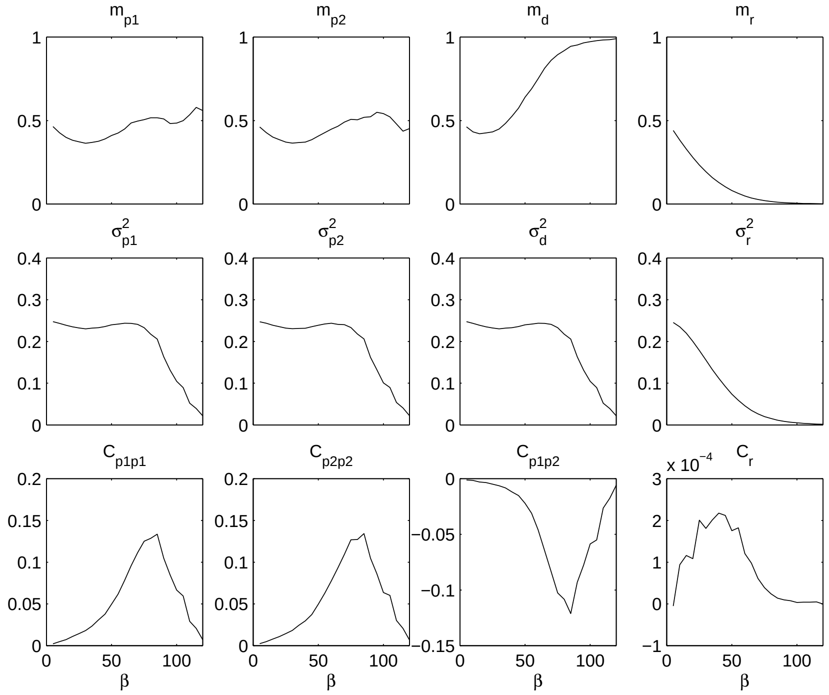

Figure 3: *Mean firing m (upper row), variance σ^2 (middle row) and covariance C (bottom row) as a function of the noise parameter β for simulations of the Glauber dynamics. The external stimulus is the logical "or" of two patterns $\mathbf{u} = \xi^1 \vee \xi^2$. The stimulus gain is $g_c = 0.2$. First column: neurons i, which are excited in pattern ξ^1 but not in pattern ξ^2 ($\xi_i^1 = 1$ and $\xi_i^2 = 0$) (first column). Second column: $\xi_i^1 = 0$ and $\xi_i^2 = 1$. Third column: $\xi_i^1 = \xi_i^2 = 1$. Fourth column: $\xi_i^1 = \xi_i^2 = 0$.*

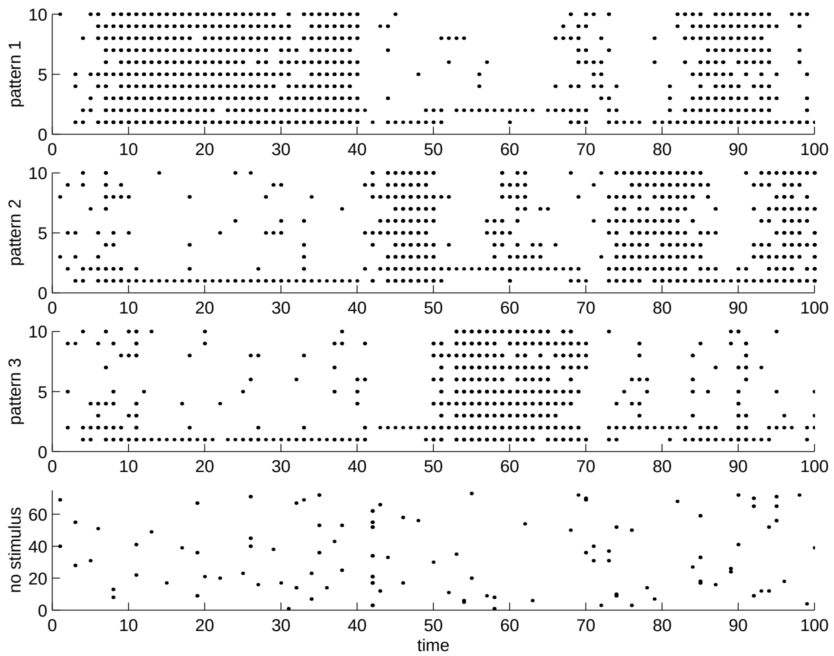

Figure 4: *Spike diagram of the SNN with a stimulus consisting of three superimposed patterns. The upper three panels show the action potentials for the ten neurons with $\xi_i^1 = 1$, $\xi_i^2 = 1$, and $\xi_i^3 = 1$ (panels 1 to 3, respectively). The bottom panel shows the action potentials for neurons with $\xi_i^\mu = 0$ for $\mu \in \{1, 2, 3\}$.*