

Efficiency of neural transmission as a function of synaptic noise, threshold, and source characteristics

Bartosz Paprocki^a, Janusz Szczepanski^{a,b,*}

^a Institute of Mechanics and Applied Computer Science, Kazimierz Wielki University, Bydgoszcz, Poland

^b Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw, Poland

ARTICLE INFO

Article history:

Received 16 July 2010

Received in revised form 11 March 2011

Accepted 14 March 2011

Neuronal computation

Entropy

Mutual Information

Estimators

Neuron

Quantal failure

Activation threshold

Neural network

ABSTRACT

There has been a growing interest in the estimation of *information* carried by a single neuron and multiple single units or population of neurons to specific stimuli. In this paper we analyze, inspired by article of [Levy and Baxter \(2002\)](#), the efficiency of a neuronal communication by considering dendrosomatic summation as a [Shannon-type channel \(1948\)](#) and by considering such uncertain synaptic transmission as part of the dendrosomatic computation. Specifically, we study Mutual Information between input and output signals for different types of neuronal network architectures by applying efficient entropy estimators. We analyze the influence of the following quantities affecting transmission abilities of neurons: synaptic failure, activation threshold, firing rate and type of the input source. We observed a number of surprising non-intuitive effects. It turns out that, especially for lower activation thresholds, significant synaptic noise can lead even to twofold increase of the transmission efficiency. Moreover, the efficiency turns out to be a non-monotonic function of the activation threshold. We find a universal value of threshold for which a local maximum of Mutual Information is achieved for most of the neuronal architectures, regardless of the type of the source (correlated and non-correlated). Additionally, to reach the global maximum the optimal firing rates must increase with the threshold. This effect is particularly visible for lower firing rates. For higher firing rates the influence of synaptic noise on the transmission efficiency is more advantageous.

Noise is an inherent component of communication in biological systems, hence, based on our analysis, we conjecture that the neuronal architecture was adjusted to make more effective use of this attribute.

© 2011 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The method and efficiency of information transmission by brain is one of the major issues that have been recently investigated, both through data analysis and theoretical modeling ([de Ruyter van Steveninck and Laughlin, 1996](#); [Levin and Miller, 1996](#); [Juusola and French, 1997](#); [Rieke et al., 1997](#); [Levy and Baxter, 2002](#); [Salinas and Bentley, 2007](#); [Lánský and Greenwood, 2007](#); [London et al., 2008](#)). Recent attempts to quantify information transmission have concentrated on treating neuronal communication process in the spirit of information theory. Information theory is the most rigorous way to analyze efficiency of transmission in aspect of probability theory. It was developed in the 1940s as a mathematical framework for quantifying information transmission in communication systems ([Shannon, 1948](#)). In general, communication system is represented

by: a source of messages, communication channel and message representation expressed in output alphabet ([Shannon, 1948](#); [Cover and Thomas, 1991](#); [Ash, 1965](#)). The fundamental concepts of this theory are entropy and Mutual Information. Mutual Information measures the reduction of uncertainty (defined as entropy) of the input (e.g. stimuli) having knowledge of the output. Both input and output are treated as random vectors. In defining what we mean by a source of information we try to capture the physical idea of the mechanism which emits a sequence of symbols in accordance with specified statistical rules. Thus, from mathematical point of view the source is modeled as a stochastic process. Under the ergodicity condition (which is commonly assumed) Mutual Information can be handled from single (input–output) trajectories by application of entropy estimators ([Strong et al., 1998](#); [Gao et al., 2008](#)). This practical idea is extensively exploited, although to get high accuracy of the numerical estimation it needs relatively long sequences.

Single neurons are the fundamental constituents of all nervous systems and have been the subject of many theoretical studies ([Abbot and Kepler, 1990](#); [Jolivet et al., 2008](#)). Two broad categories of spiking neuron models have been extensively analyzed and used: Hodgkin–Huxley type neuron models ([Hodgkin and Huxley, 1952](#))

* Corresponding author at: Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawlowskiego 5B, 02-106 Warsaw, Poland.

E-mail addresses: bartekp@ukw.edu.pl (B. Paprocki), jszczepa@ippt.gov.pl (J. Szczepanski).

and simplified phenomenological neuron models, among which the integrate-and-fire model is the most famous representative (Polsky et al., 2004). Computational models of the neuron are inspired just by the integrated-and-fire models (Victor, 2006; Abbot and Kepler, 1990). The most common of these computational models are those based on McCulloch and Pitts (1943) model. This model computes the weighted sum of the inputs and produces the output based on the activation function. In this paper we analyze, inspired by article of Levy and Baxter (2002), the efficiency of neuronal communication by considering dendrosomatic summation as a Shannon-type channel (1948) and by considering such uncertain synaptic transmission as part of the dendrosomatic computation. They analyzed optimization of information processing relative to energy use and they provided a general expression relating optimal failure rate to average rate. It turned out that appropriate uncertainty of synaptic transmission can even improve efficiency in this sense.

In this paper we study Mutual Information between input and output signals by applying extensively efficient entropy estimators. We also consider the Mutual Information optimization problem for different neural network architectures. There are a few neuron parameters that must be taken into account while analyzing the information transfer process. Among them the crucial role is played by synaptic failure rate and neuron activation threshold. The questions arise: how the values of these parameters affect the transmission efficiency (in the sense of Mutual Information) for single neuron and consequently what are the optimal values of these parameters for different architectures of neuronal networks.

We observed a number of surprising non-intuitive effects. It turns out that, especially for lower activation thresholds, significant synaptic noise can lead even to double the transmission efficiency. Moreover, the efficiency occurs to be a non-monotonic function of the activation threshold. We find a universal value of threshold for which a local maximum of Mutual Information is achieved for most of the neuronal architectures, regardless of the type of the source (correlated and non-correlated). Additionally, to reach the global maximum the optimal firing rates increases with the threshold. This effect is particularly visible for lower firing rates. For higher firing rates the influence of synaptic noise on the transmission efficiency is more advantageous.

Noise is inherent component of communication in biological systems, hence, based on our analysis, we conjecture that the neuronal architecture was adjusted to make more effective use of this attribute.

The paper is organized as follows. In Section 2 we briefly recall the basic information theory concepts to describe the Shannon communication channel idea. In Section 3 we present the model of neuron and network architectures which we have analyzed. In Section 4 results of our calculations are described in detail. The last section contains conclusions and discussion of the results obtained.

2. Theoretical Concepts

Information theory is concerned with the analysis of an entity called a *communication system*. In general, such system is represented by: a source of messages, a communication channel and messages representations expressed in an output alphabet (Shannon, 1948; Cover and Thomas, 1991; Ash, 1965). Recent attempts to quantify information transmission have concentrated on treating neurons already as communication channels (Levy and Baxter, 2002). From mathematical point of view messages can be understood as trajectories of stochastic processes being in fact sequences of symbols. It is assumed that the set of symbols (alphabet) is finite and the stochastic processes under consideration have stationary distributions (Cover and Thomas, 1991; Ash, 1965).

In this paper we consider two types of information sources: independent and correlated. Stochastic stationary process $X=(X_i)$

is said to be independent information source if its component variables X_i are identically distributed and independent. Such process is called the Bernoulli process if these variables are binary, that is the process takes only two values (states). We denote these two states by 1 and 0, indicating *spike* and *no-spike*, respectively. Due to stationarity, the probability that $X_i=1$ (spike) is the same for all i and we denote it by $p \in [0, 1]$. Since for long time intervals the fraction of spikes to the number of all symbols in the sequence for a typical Bernoulli trajectory is close to p , this parameter can be assumed as the *firing rate* f_r of neuronal activity.

The most common probabilistic model used in neurophysiology, the inhomogeneous Poisson process, often seems to be too crude. Recently a more general Markov like models (Berry and Meister, 1998; Kass and Ventura, 2001; Gerstner and Kistler, 2002) have been considered. They take into account both the current experimental time and the time from the last spike. Therefore, for the correlated sources considered in our paper we assume that the source is described by such a Markov process, i.e. by a set of conditional probabilities $P(X_{n+1} | X_n)$ describing changes of process' states. These changes are called *transitions* and the probabilities associated with them are called *transition probabilities*. These probabilities can be put together into a matrix $p_{ij} = P(X_{n+1}=j | X_n=i)$ called the *transition matrix*, which for the two-states-process is of the form:

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}. \quad (1)$$

This is stochastic matrix, that is each of its rows sum to 1, representing probability of transition from the current state to either of the two in the next step. The *stationary distribution* of such process is a (row) vector π that satisfies $\pi\mathbf{P}=\pi$ and for the above matrix \mathbf{P} can be easily determined analytically as:

$$\pi = \left[\frac{p_{10}}{p_{01} + p_{10}}, \frac{p_{01}}{p_{01} + p_{10}} \right]. \quad (2)$$

The stationary distribution of such Markov process is intuitively an attempt to approximate its, in fact, dependent distribution by an state-independent one. Therefore, the *firing rate* parameter can be obtained directly from its second entry:

$$f_r = \pi_2 = \frac{p_{01}}{p_{01} + p_{10}}. \quad (3)$$

2.1. Entropy and Information

Let X^L be a set of all blocks (or words) $x^L = x_1 x_2 \dots x_L \in X^L$ of length L , built of symbols (letters) coming from some finite alphabet. This set can be interpreted as messages generated by an information source X (or, equivalently, stationary stochastic process). If $P(x^L)$ denotes the probability of the word $x^L \in X^L$, then the Shannon *i* nformation carried by this word is defined as:

$$I(x^L) := -\log P(x^L). \quad (4)$$

Since logarithms to the base 2 are used, this quantity is measured in units of bits. In this sense, less probable events (words) carry more information. Average (expected) information contained in X^L , called Shannon *block entropy*, is defined as:

$$H(X^L) := \mathbb{E}(I(X^L)) = - \sum_{x^L \in X^L} P(x^L) \log P(x^L), \quad (5)$$

and is also measured in units of bits. The word length L can be chosen arbitrary, so the block entropy does not perfectly describe the process X . The *entropy rate* (or source entropy) is an invariant quantity characterizing the source X :

$$H(X) := \lim_{L \rightarrow \infty} \frac{H(X^L)}{L} \quad (6)$$

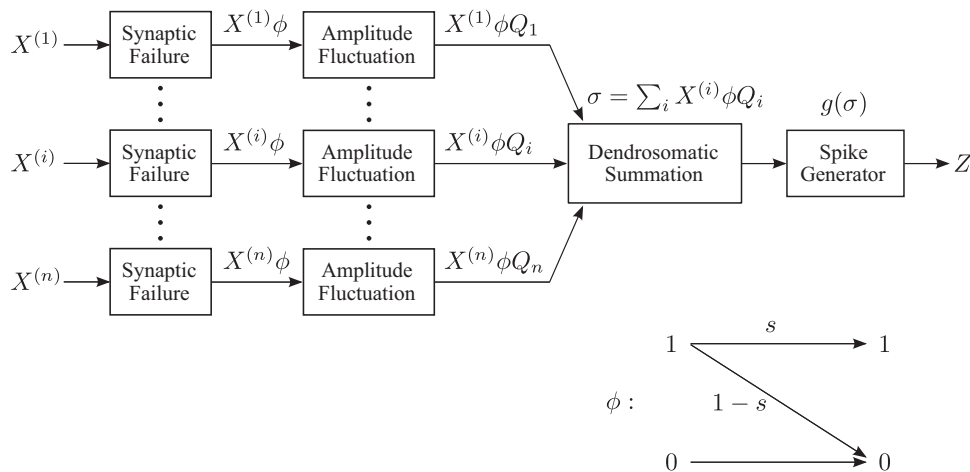


Fig. 1. The information flow for a single neuron. $X^{(i)}$ is a stochastic process representing presynaptic input to the i th dendrite of the postsynaptic neuron. Each input is subject to quantal failures ϕ and quantal amplitude modulation Q_i . Such transformed stimuli $X^{(i)}\phi Q_i$ are then all summed up together dendrosomatically, and this sum, $\sigma = \sum_i X^{(i)}\phi Q_i$, is the input to the spike generator, $g(\sigma)$. The spike generator output is a binary stochastic process Z . This figure is a modification of Fig. 1 published by Levy and Baxter (2002).

and this limit exists if and only if the process is stationary (Cover and Thomas, 1991). We see that entropy rate can be understood as the average information transmitted by a source per symbol.

Having two stochastic processes X and Z the *Mutual Information* between them is given as:

$$I(X; Z) := H(X) - H(X|Z) = H(X) + H(Z) - H(X, Z), \quad (7)$$

and can be intuitively understood as a measure of information of one process reflected in the realization of the other one. This quantity is especially important if Z is an outcome of some transformation of X (or vice versa), that is $X \rightarrow f(X) = Z$, for example: evolution of signal transmitted through neuron, which actually is exactly what we do in this paper. In other words $I(X; Z)$ measures the reduction of uncertainty concerning realization of X having knowledge about the realization of Z . This concept is essentially broader because it includes also higher correlations (DeWeese, 1996; Panzeri et al., 1999) and thus can be complementary to cross-correlations analysis.

2.2. Entropy Rate Estimators

Mutual Information constitutes the basic quantity analyzed in our paper. However, if distribution of a stochastic process, say X , is unknown, we cannot determine $H(X)$ analytically using Eq. (5). Hence, component entropies in Eq. (7) have to be estimated numerically. In this paper we use the estimator recommended by Strong et al. (1998) because of its low computational complexity and high accuracy. It is based on calculating block entropies using observed frequencies of words x^L for some few consecutive lengths L . Then a line h , best fit to the points $(1/L, (H(X^L))/L)$, is determined. Finally, with $1/L \rightarrow 0$, h is extrapolated to the point of $(0, H(X))$. Using this method we are able to get fast and reliable estimations. It is worth stressing that we applied this algorithm to the sequences of 2^{20} bits long, which was computationally expensive, but this assured us to get very high accuracy of estimation.

Other approaches to entropy rate estimators are also discussed in the literature (Kennel et al., 2005; Kontoyiannis et al., 1998). It was shown (Amigo et al., 2004) that Lempel–Ziv complexity as defined by Lempel and Ziv (1976) can be used successfully as such an estimator.

3. Neuron Model and Architecture of Network

The neuron model we use in our research is that proposed by Levy and Baxter (2002). The way the information flows through such neuron is depicted in Fig. 1. For i th dendrite, input process $X^{(i)}$ represents its presynaptic stimulus. This i th signal then undergoes the quantal release process (reflected by the information-losing transformation ϕ) and the quantal amplitude variation of the released spike (denoted by the random variable Q_i). Eventually, the dendrosomatic summation occurs followed by the spike generator transformation. The neuron responds with a spike only if the magnitude of its excitation exceeds threshold.

With n denoting the number of dendrites, the multivariate input stimuli $\mathbf{X} = (X^{(1)}, \dots, X^{(n)})$ is modeled with all $X^{(i)}$ being either Bernoulli or Markov stochastic processes (see Section 2). The ϕ transformation is in fact a Bernoulli distributed random variable with success parameter s . The probability of a successful quantal release is s , therefore the probability of quantal failure is $1 - s$. The random variable Q_i (denoting amplitude fluctuation of i th input signal) is in this paper considered continuous and distributed uniformly on the interval $[0, 1]$. The spike generator, which takes the dendrosomatic summation value σ as an argument, is simply a threshold function that determines if σ exceeds a given threshold (then it takes the value of 1 and the spike flows

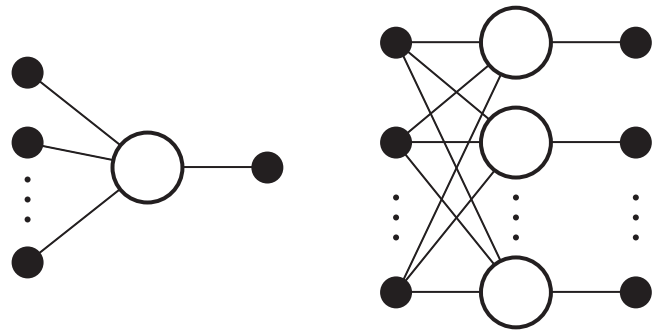


Fig. 2. Simplified schemes of discussed neural architectures. Small black circles represent input and output processes, bigger white—neurons. Single neurons scheme is depicted to the left, and neural network scheme to the right. The neurons do not interact but each input bit affects outputs from all neurons because of feedforward network architecture.

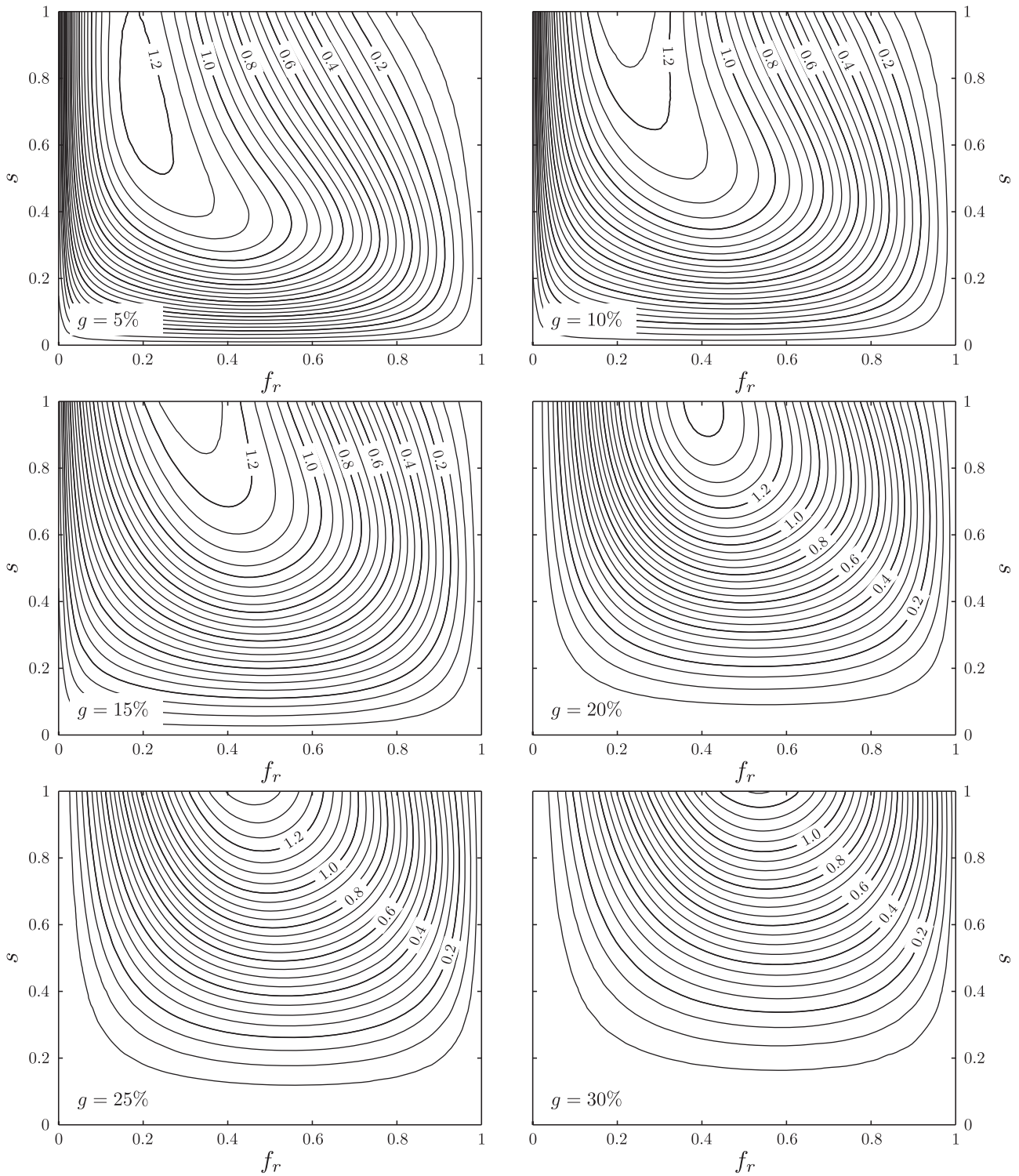


Fig. 3. Mutual Information $I(\mathbf{X};\mathbf{Z})=I(f_r, s, g)$, 5×5 neural network stimulated by a Bernoulli source. The lower the threshold, the synaptic failure effects are more wide and common. They are most spectacular for the lowest threshold $g=5\%$, while for $g=30\%$ they are hardly recognizable.

down the axon away from the neuron) or not (returned value of 0).

Under these assumptions, single neuron response is denoted by the implicit stochastic process Z . If a one-layer feedforward network of such neurons is built (also considered in our research) then the process $\mathbf{Z}=(Z^1, \dots, Z^k)$ is its output, with k denoting the number of neurons in the network.

4. Results

Given the neuron model and neural network architectures described in Section 3, we estimated Mutual Information, $I(\mathbf{X};\mathbf{Z})$, between stochastic process \mathbf{X} representing the source of stimuli, and process \mathbf{Z} , the neuron (or network) response. To comply with the multivariate form of input and output processes \mathbf{X} and \mathbf{Z} , Eq.

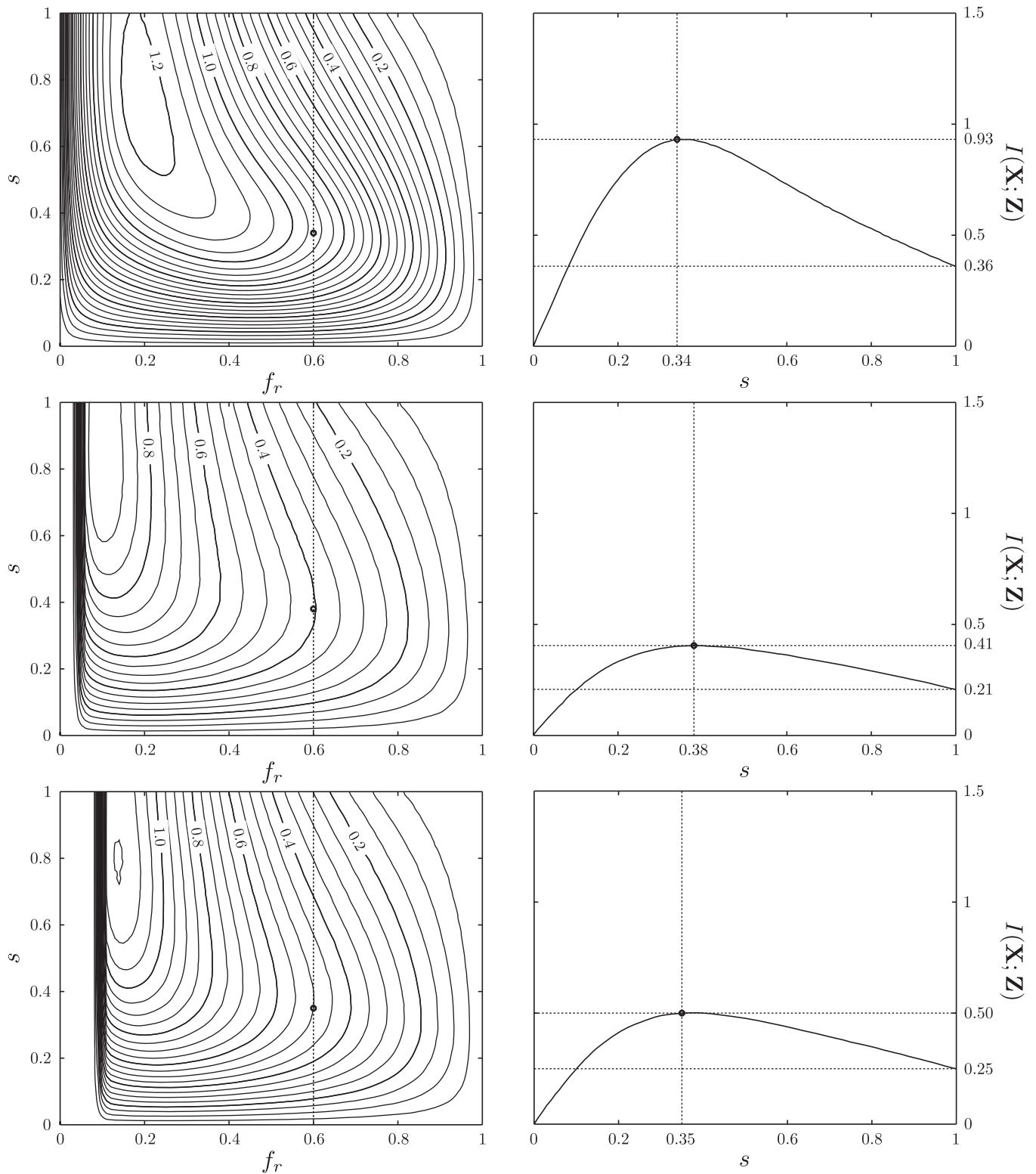


Fig. 4. Mutual Information dependencies on synaptic success, s , in neural network 5×5 architectures with threshold set to 5%. All plots (left column) are sliced at $f_r = 0.6$ and in all three cases Mutual Information values are maximized for synaptic successes less than 0.4, where they reach values about twice as large as for $s = 1$. Bernoulli source (top), Markov source with $p_{01} = 0.05$ (middle) and Markov source with $p_{01} = 0.10$ (bottom).

(7) is rewritten to:

$$I(\mathbf{X}; \mathbf{Z}) = H(\mathbf{X}) + H(\mathbf{Z}) - H(\mathbf{X}, \mathbf{Z}), \quad (8)$$

where $H(\mathbf{X}) = n \cdot H(X)$ can be calculated analytically, and the other two components must be estimated numerically. In estimation pro-

cedure we use Strong method (Strong et al., 1998), described briefly in Section 2.2.

Four different neural architectures were considered in this paper. Two single neurons with input size of 3 or 5, respectively (Fig. 2; left) and then two neural networks consisting of 3 or 5 such neurons (Fig. 2; right).

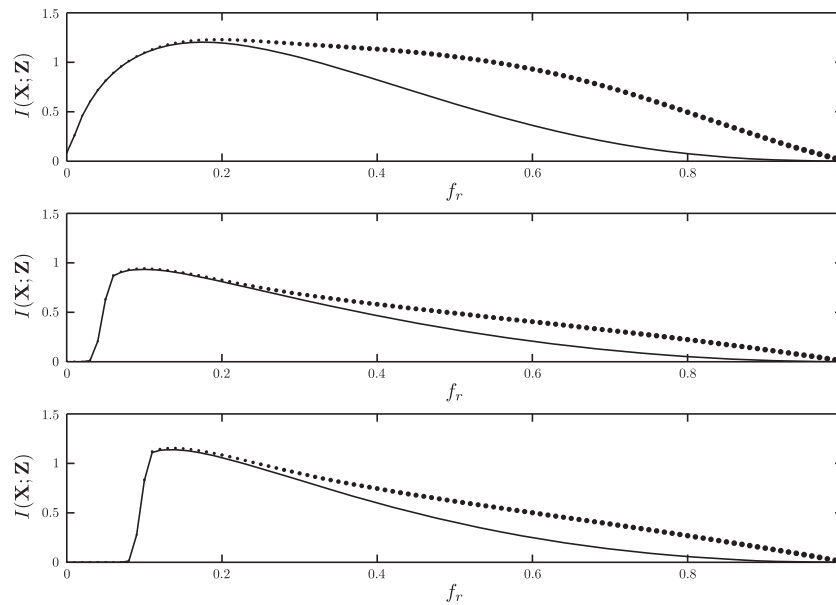


Fig. 5. Mutual Information dependencies on *synaptic success*, s , in neural network 5×5 architectures with threshold set to 5%. Comparison of maximal Mutual Information values (dotted line) with these achieved at $s=1$ (solid). Size of a dot is proportional to $1-s$, indicating the bigger the dot, the corresponding Mutual Information value is achieved at lower s . Bernoulli source (top), Markov source with $p_{01} = 0.05$ (middle) and Markov source with $p_{01} = 0.10$ (bottom).

To model stimuli we used three different types of information source. First we studied Bernoulli stochastic process as a model of independent input, then we used two Markov processes (modeling correlated inputs), with parameter p_{01} equal to 0.05 and 0.1, respectively. The firing rate parameter, f_r , directly connected with these processes, was gradually changed from its minimal value of 0 (for Bernoulli source) or $p_{01}/(p_{01} + 1)$ (for Markov model) to its maximal value of 1, in steps of 0.01.

Synaptic success parameter, s , was gradually changed from its minimal possible value of 0 (all spikes lost in synaptic noise) to 1 (all spikes allowed to pass, no synaptic noise), in steps of 0.01. We also assumed that all Q_i random variables were uniformly distributed on interval $[0, 1]$. The activation threshold parameter g , was set to 14 different values, varying from 5% to 90%.

The goal is to examine how a given neural structure behaves in response to various changes of these parameters described above. We focused mainly on synaptic success/failure effects, threshold effects and source type effects.

4.1. Synaptic Failure Effects

In this section we discuss how the *synaptic success* parameter, s , affects the efficiency of transmission in already described (Section 3) neural architectures, therefore how it affects the estimated Mutual Information.

Intuitively, synaptic success parameter helps neurons transmit information through the network, as it can be seen as a probability that a given spike would not vanish covered by any kind of synaptic noise. So, assuming neurons do act as information carriers, intuition suggests that the closer to 1 the value of s is (closer to its maximal value), the more information is transmitted and so the transmission itself is more efficient. Conversely, the lesser s is (closer to 0), the less number of spikes is permitted to pass to other cells and so neurons get less capable of transmitting information efficiently.

However, our research shows that in many cases lowering value of s results in neuron (and neural network) behaviour that seems to disagree with this intuition. This depicts an interesting feature having noisy (spikes-disappearing) synapses can make information transmission more efficient (in the sense of Mutual Information, $I(\mathbf{X}; \mathbf{Z})$, as a measure of this efficiency).

This effect is present for low thresholds and best visible for the lowest of values of g . In Table 1 we show maximal thresholds for which this synaptic effects appear.

As Table 1 shows it is clear that the size of the input process seems to have major impact on the range of thresholds for which the synaptic failure effects are visible while number of neurons is less influential. These observations make the 5×5 neural network (5 neurons with input size of 5 each) the most interesting case. Fig. 3 presents an example of this effect plotting Mutual Information values in such 5×5 network, stimulated by a Bernoulli source, for six lowest thresholds: 5, 10, 15, 20, 25 and 30 %.

Fig. 3 shows a strong dependency of appearance of this neural phenomenon on the value of threshold. The lower the threshold, the synaptic failure effect is common for wider range of firing rates and the differences, at given f_r , between maximal Mutual Information values and the ones achieved at $s=1$ (maximal s) are overall bigger. For example, in Fig. 4 we plot Mutual Information in function of *synaptic success*, s , with *firing rate* value fixed at 0.6 for all three kinds of input sources to show the surprisingly large scale of this difference when for the most complex of studied architectures, the 5×5 case, value of threshold is set to 5% (that is, the lowest value explored in our research).

In a 5×5 neural network with threshold set to 5% and with firing rate fixed at some value, Mutual Information can be maximized for rather non-intuitively low *synaptic success* values, for example for $s < 0.4$ as Fig. 4 shows. This literally means, that $I(\mathbf{X}; \mathbf{Z})$ can be maximized in a very noisy environment where less than a half of already generated spikes (or even only one third of their

Table 1

Maximal values of thresholds for given architecture and source types for which effects of synaptic failure appear. Thresholds shown in unit of percent [%].

Source type	Architecture type			
	3×1	5×1	3×3	5×5
Bernoulli	5	20	10	30
Markov, $p_{01} = 0.05$	5	20	10	30
Markov, $p_{01} = 0.10$	5	20	10	30

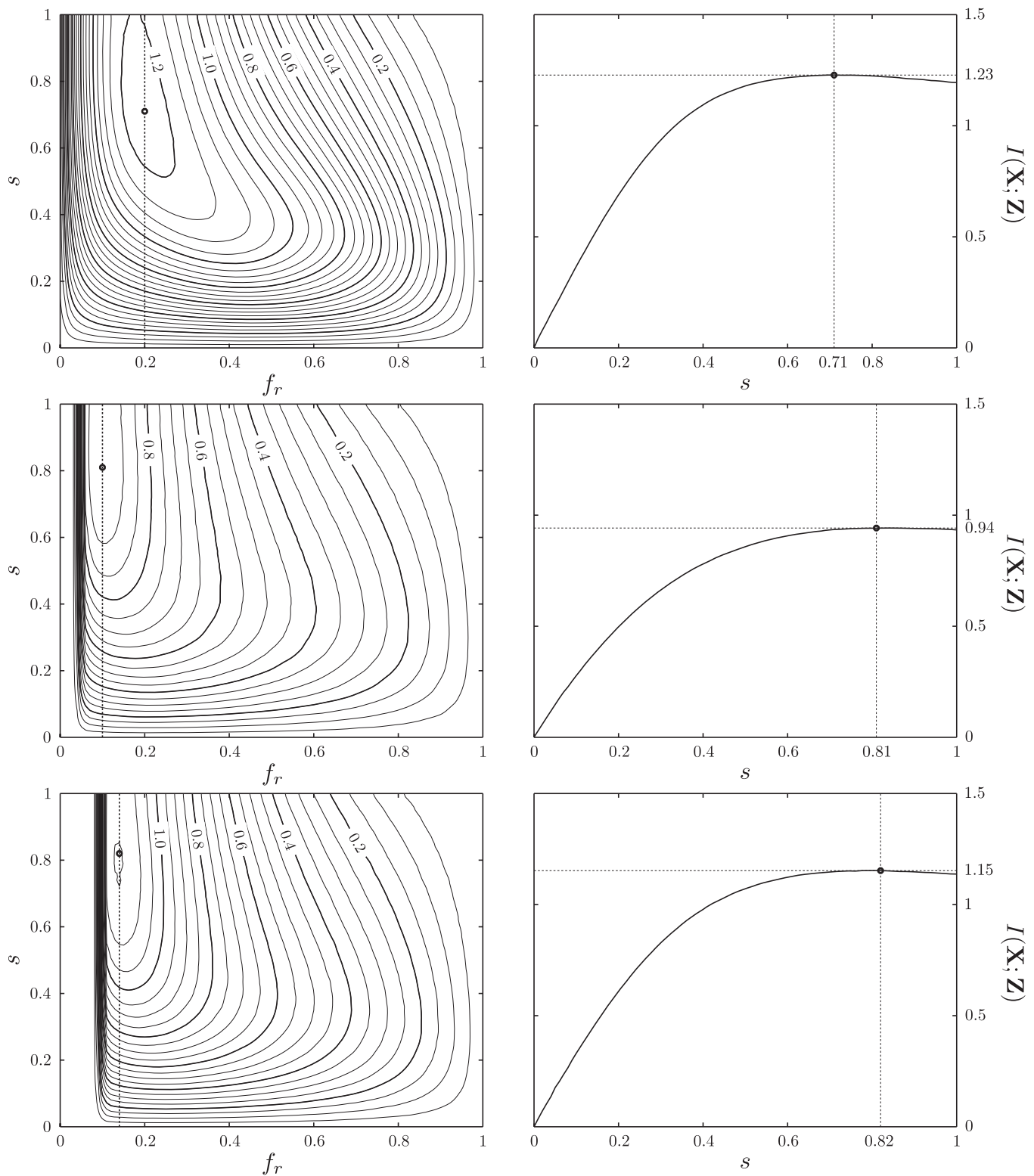


Fig. 6. Mutual Information dependencies on synaptic success, s , in neural network 5×5 architectures with threshold set to 5%. All plots (left column) are sliced at firing rates for which Mutual Information achieves its maximal value (equal to corresponding channel capacity). In all three cases Mutual Information values are maximized for synaptic successes less than 1. Bernoulli source (top), Markov source with $p_{01} = 0.05$ (middle) and Markov source with $p_{01} = 0.10$ (bottom).

total number) are allowed to pass further through the network. Moreover, value of this maximum can be more than two times greater than that achieved in noise-free conditions, that is where $s = 1$. In Fig. 5 we compare maximal Mutual Information values with these achieved at $s = 1$, for all firing rates and for all three input sources.

Another thing of major interest is that for discussed architecture (5×5 neural network with *threshold* parameter set to 5%) the maximal Mutual Information value over all firing rates is also achieved for $s < 1$ (regardless of stimuli source model). The evidence is plotted in Fig. 6. If we consider this network as a form of transmission channel, then this observation proves to be valuable in the examination

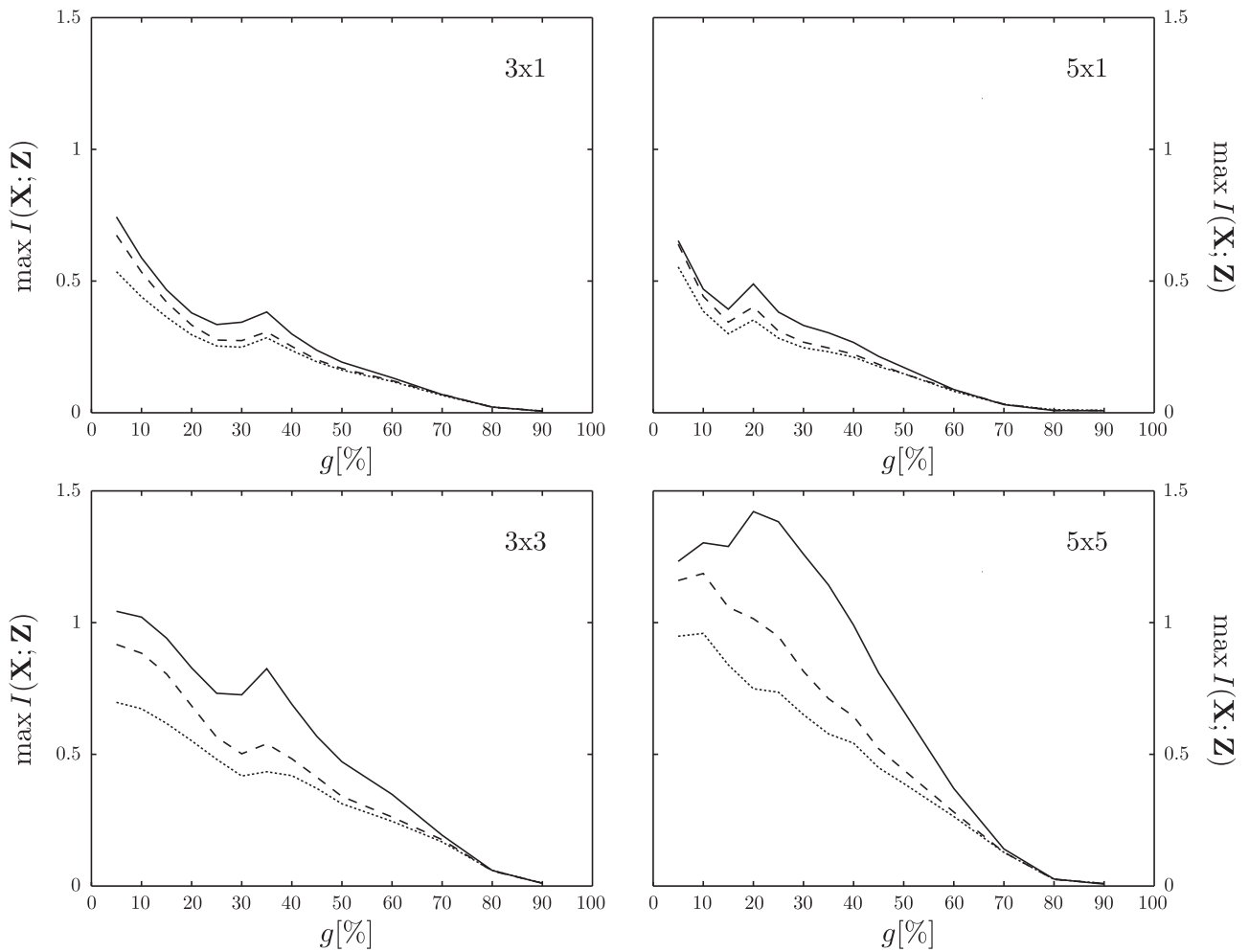


Fig. 7. Mutual Information's non-monotonic dependencies on *threshold* g , in all four discussed neural network architectures, indicating presence of local extrema. Position of these extrema seems to be independent of input stimuli. This is observed for all explored structures but the most complex one, the 5×5 network, where some of this regularity is lost. Bernoulli source (solid line), Markov with parameter $p_{01} = 0.05$ (dotted), Markov with $p_{01} = 0.10$ (dashed).

of the channel capacity concept, as it shows that the discussed neural structure is the most capable of transmitting information in condition of noisy synapses.

4.2. Threshold Effects

In this section we describe the way the *threshold* parameter affects estimated values of Mutual Information. Some of these properties are already shown in Section 4.1, where we noted that only for low thresholds the synaptic failure effects appear, precisely: the lower, the better as the synaptic failure effects are more common and greater in their values.

Threshold is a neural mechanism that prevents weak signals (in the sense of their cumulated amplitude) from being passed down through axon to other neurons. Since only these signals are permitted to fire that are strong enough to exceed the barrier, therefore intuition suggests that the maximal amount of information transmitted within given neural structure falls as the value of threshold grows. However, results of our research appear to counter this intuition. Fig. 7, plotting maximal values of Mutual Information in function of threshold for all discussed neural architectures and all stimuli, proves that this dependency is non-monotonic, what implies surprising existence of local extrema.

Fig. 7 shows that for all cases there always exists a local extremum. Moreover (excluding the 5×5 case), neither source type nor number of neurons appears relevant to the location of this max-

imum. For both structures with input size of 3, the extremum occurs at $g = 35\%$, while for the 5×1 case it is situated at $g = 20\%$.

Changing the unit of g from relative (%) to absolute reveals that the size of the input source is also irrelevant to the position of a given maximum. This change is reflected by scaling relative value of g by factor equal to the size of the input process. Therefore, for 5×1 network it is $5 \cdot 20/100 = 1$, and for both 3×1 and 3×3 it is $3 \cdot 35/100 = 1.05$. Both equations end with a unity or a number around one (unity would be achieved if the extremum was situated at relative $g \approx 33.3\%$).

Why unity? If threshold is lower than 1, only one signal is needed to reach (and exceed) it since this single signal carries 1 bit of information at most, precisely $X^{(i)} \Phi Q_i \in [0, 1]$ (see Section 3 for a reminder). Therefore a threshold of 1 is a barrier, below which only one strong enough signal is sufficient for the whole neuron to fire. But the probability of such event decreases as the value of threshold grows towards a unity. We have already mentioned that less probable events carry more information (4), so this could explain why curves in Fig. 7 grow in some left hand neighbourhood of (absolute) $g = 1$. Conversely, above the barrier of a unity-threshold, it is certain that if neuron fired, at least two signals participated in dendrosomatic summation. As soon as something becomes known for certain, the amount of information gained decreases, so the curves start to fall. This could be the explanation for the universal location of these local extrema but the phenomenon is still subject of question and deserves further research.

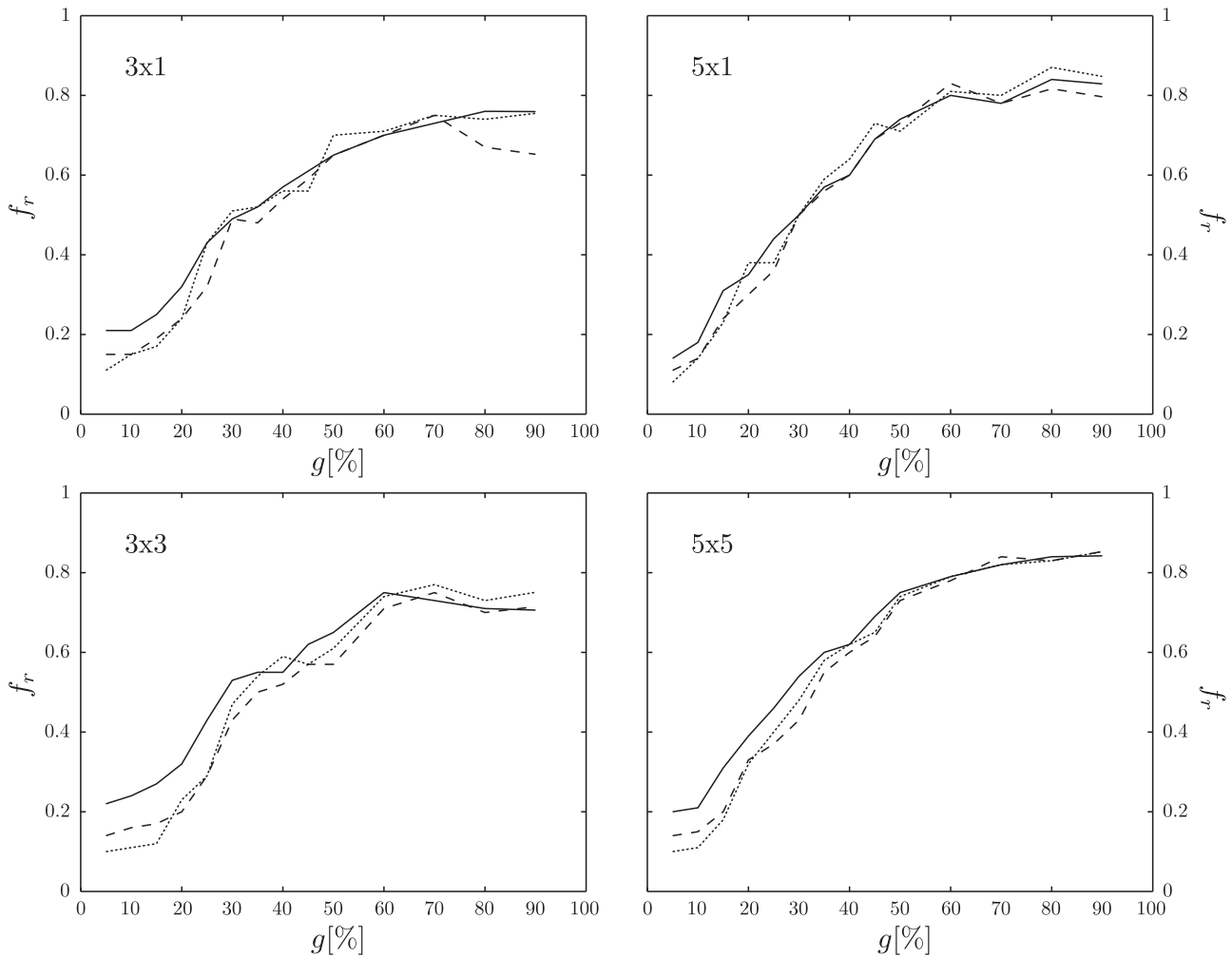


Fig. 8. The larger the threshold is, the larger firing rate is needed to maximize the Mutual Information. Bernoulli source (solid line), Markov with parameter $p_{01} = 0.05$ (dotted), Markov with $p_{01} = 0.10$ (dashed).

4.3. Source Type Effects

The last aspect of the problem we focused on in our research is the comparison of different stimuli source type models. Precisely, our goal is to specify which of the two: independent or correlated process cause neurons or neural networks to perform more efficient in the sense of Mutual Information values achieved. As mentioned at the beginning of the chapter, we used Bernoulli process to model independent input and two Markov processes to model correlated input (both of order 1, with parameter p_{01} equal to 0.05 and 0.1, respectively; see Section 2 for a reference). All three source processes were stationary.

Much in this area can be deduced from previous two Sections 4.1 and 4.2. Figs. 4–7 prove that explored neural structures behave better if independent source is applied to their input. This performance gain is best viewed in Fig. 7 where for all cases, best results are achieved by neurons stimulated with independent source. This assumption is true for all thresholds.

Synaptic failure effect (Section 4.1) is more spectacular for independent input than for its correlated counterparts. Fig. 6 shows not only that estimated maximal Mutual Information values are greater if neurons are stimulated with uncorrelated process, but also that these maxima are achieved for lower, less intuitive *synaptic success* parameter. Similar conclusion is reached after examination of Fig. 4.

Tables 2 and 3 describe the effect quantitatively. For a given architecture and input process, Table 2 lists all parameter's values at the point where Mutual Information achieves its maximal value over all thresholds. Table 3 is similar to Table 2 but concentrates on local maxima of Mutual Information in function of threshold.

Tables 2 and 3 prove that in all cases, values of both globally maximal Mutual Information (that is: neural chan-

Table 2

Global Mutual Information's maxima over all thresholds for all discussed neural structures and input source models.

Input size	No. of neurons	Source model	p_{01}	f_r	g [%]	s	$H(X)$	$I(X;Z)$
3	1	Bernoulli	–	0.21	5	1	2.22	0.74
		Markov	0.05	0.11	5	1	1.09	0.54
		Markov	0.10	0.15	5	1	1.64	0.67
5	1	Bernoulli	–	0.14	5	1	2.92	0.65
		Markov	0.05	0.08	5	1	1.71	0.55
		Markov	0.10	0.11	5	1	2.47	0.64
3	3	Bernoulli	–	0.22	5	1	2.28	1.04
		Markov	0.05	0.10	5	1	1.07	0.70
		Markov	0.10	0.14	5	1	1.61	0.92
5	5	Bernoulli	–	0.39	20	1	4.82	1.42
		Markov	0.05	0.11	10	1	1.81	0.96
		Markov	0.10	0.15	10	1	2.73	1.19

Table 3

Local Mutual Information's maxima for all discussed neural structures and input source models. For 5×5 network stimulated with correlated (Markov) inputs these local extrema are also global.

Input size	No. of neurons	Source model	p_{01}	f_r	g [%]	s	$H(X)$	$I(X;Z)$
3	1	Bernoulli	–	0.52	35	1	3.00	0.38
		Markov	0.05	0.52	35	1	0.83	0.28
		Markov	0.10	0.48	35	1	1.44	0.31
5	1	Bernoulli	–	0.35	20	1	4.67	0.49
		Markov	0.05	0.38	20	1	1.66	0.35
		Markov	0.10	0.30	20	1	2.82	0.40
3	3	Bernoulli	–	0.55	35	1	2.98	0.83
		Markov	0.05	0.54	35	1	0.81	0.43
		Markov	0.10	0.50	35	1	1.41	0.54
5	5	Bernoulli	–	0.21	10	1	3.71	1.30
		Markov	0.05	0.11	10	1	1.81	0.96
		Markov	0.10	0.15	10	1	2.73	1.19

nel capacities) and local extrema are greater for independent stimuli.

The process' firing rate parameter also affects the way the discussed networks perform. The larger the firing rate is, the stronger is the effect of positive influence of synaptic failure on transmission efficiency. This fact is particularly well exposed in Fig. 5, and can be also deduced from Figs. 4 and 6. Moreover, the larger the threshold is the larger the firing rate is needed to maximize the Mutual Information (Fig. 8).

The results presented in Figs. 3–6 suggest a relation between source parameter f_r and neuron's parameters s and n like $n \cdot f_r \cdot s \approx g$. This formula can be supported also theoretically. In fact, when we treat all the values X , ϕ , Q as average values of random variables and we assume that they are independent, we have

$$E \left(\sum_{i=1}^n X^{(i)} \phi Q_i \right) = n \cdot EX \cdot E\phi \cdot EQ = n \cdot f_r \cdot s \cdot EQ.$$

The expression $\sum_{i=1}^n X^{(i)} \phi Q_i$ in this formula is just the value which is subsequently compared with the assumed activation threshold g , thus this justifies the optimal relation $n \cdot f_r \cdot s \cdot EQ \approx g$.

5. Conclusions

The concept of information, as well as the related notion of neural coding and neural computation, is still a topic of intensive research in neuroscience. In this paper the information is understood in the Shannon sense and neurons are treated as communication channels. The fundamental quantity that links input and output signals in this theory is Mutual Information. It measures reduction of uncertainty concerning the input having knowledge about output, which is the essence of the transmission efficiency estimation in the sense of Information Theory. Mutual information can be expressed in terms of entropies (of the source, output source and joint of the both) and this fact, since we have well elaborated entropy estimators theory, leads to the possibility of computational treatment of the efficiency problem.

We analyze how the values of neuron parameters affect transmission efficiency of both single neuron as well as simple neuronal networks. In the model of neuron under consideration, these parameters are synaptic failure and activation threshold. The information processed by a channel clearly depends also on the properties of the source of inputs (stimuli). These characteristics of the source, important from the neuroscientific point of view, are firing rate and correlation range.

The question if the synaptic failures play any role in information processing is a subject of interest of many authors (de Ruyter van

Steveninck and Laughlin, 1996; Barbour, 2001; London et al., 2008). Intuitively, this in fact physiological phenomena should rather make the transmission less effective. However, our calculations show that for a large range of parameters the optimal transmission is reached when the effect of synaptic failure is relatively large. This is especially visible for lower thresholds and for more complex architectures of the neuronal network. We conjecture that the architecture of neuron is such that a kind of correcting code mechanism is exploited.

One expects that, because higher threshold of activation implies less spikes being transmitted, the Mutual Information should monotonically decrease with the increase of the threshold. But surprisingly we observed that there is a universal value of threshold for which Mutual Information has local maximum. This effect is probably due to specific form of dendrosomatic summation and it appears when the value of threshold goes through integer values.

Moreover, it turned out that this type of neuronal architecture prefers uncorrelated sources and the efficiency can increase even 50 percent for uncorrelated sources. In addition, the most efficient transmission (the global maximum) is for relatively small (0.1–0.2) firing rates values while the maximal local efficiency requires in general more active neurons (0.3–0.5). Our calculations also show that in many cases for a given fixed firing rate, the information process essentially depends on the values of neurons parameters (for example on synaptic success as it is seen in Figs. 3 and 4).

Our results are based on simplified model of neuron and the neuronal architecture in order to get insight into the phenomena that can occur when we study more advanced and realistic biological models. The similar type of model was already considered by Levy and Baxter (2002).

Finally, it is worth stressing that our estimations are of very high accuracy. This is because in our numerical simulations we used sequences of 1 Mibit long (about 10^6 bits). For experimental data they are significantly shorter due to biological restrictions. It is rather rare in biological modeling to be able to determine all the relevant parameters with the accuracy implied by the explicit calculations practiced on them.

Acknowledgements

This paper has been partially supported (BP) by Marshal Office of State Kujawsko-Pomorskie. Valuable discussion with Jan Karbowski is appreciated. The authors would also like to thank Alicja Chmielinska and Tomasz Terpilowski for careful reading of the manuscript.

References

- Abbot, L., Kepler, T., 1990. Model neurons: from Hodgkin–Huxley to Hopfield. In: Garrido, L. (Ed.), In: Statistical Mechanics of Neural Networks, 368. Springer-Verlag, Berlin, pp. 5–18.
- Amigo, J., Szczepanski, J., Wajnryb, E., Sanchez-Vives, M., 2004. Estimating the entropy rate of spike-trains via Lempel–Ziv complexity. Neural Computation 16, 717–736.
- Ash, R., 1965. Information Theory. Interscience.
- Barbour, B., 2001. An evaluation of synapse independence. Journal of Neuroscience 20 (21), 7969–7984.
- Berry, M.J., Meister, M., 1998. Refractoriness and neural precision. Journal of Neuroscience 18 (6), 2200–2211.
- Cover, T., Thomas, J., 1991. Elements of Information Theory. John Wiley and Sons, Inc.
- de Ruyter van Steveninck, R., Laughlin, S., 1996. The rate of information transfer at graded-potential synapses. Nature 379, 642–645.
- DeWeese, M., 1996. Optimization principles for the neural code. Network: Computation in Neural Systems 7, 325–331.
- Gao, Y., Kontoyiannis, I., Bienenstock, E., 2008. Estimating the entropy of binary time series: methodology, some theory and a simulation study. Entropy 10 (2), 71–99.
- Gerstner, W., Kistler, W.M., 2002. Spiking Neuron Models. Cambridge University Press, Cambridge.

- Hodgkin, A., Huxley, A., 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology* 117, 500–544.
- Jolivet, R., Roth, A., Schurmann, F., Gerstner, W., Senn, W., 2008. Special issue on quantitative neuron modeling. *Biological Cybernetics* 99 (4–5), 237–239.
- Juusola, M., French, A., 1997. The efficiency of sensory information coding by mechanoreceptor neurons. *Neuron* 18, 959–968.
- Kass, R.E., Ventura, V., 2001. A spike-train probability model. *Neural Computation* 13 (8), 1713–1720.
- Kennel, M., Shlens, J., Abarbanel, H., Chichilinsky, E., 2005. Estimating entropy rates with bayesian confidence intervals. *Neural Computation* 17, 1531–1576.
- Kontoyiannis, I., Algoet, P., Suhov, Y., Wyner, A., 1998. Nonparametric entropy estimation for stationary processes and random fields, with applications to english text. *IEEE Transactions on Information Theory* 44, 1319–1327.
- Lánský, P., Greenwood, P., 2007. Optimal signal in sensory neurons under an extended rate coding concept. *Biosystems* 89 (1–3), 10–15.
- Lempel, A., Ziv, J., 1976. On the complexity of finite sequences. *IEEE Transactions on Information Theory* 22, 75–88.
- Levin, J., Miller, J., 1996. Broadband neural encoding in the cricet cercal sensory system enhanced by stochastic resonance. *Nature* 380, 165–168.
- Levy, W., Baxter, R., 2002. Energy-efficient neuronal computation via quantal synaptic failures. *The Journal of Neuroscience* 22, 4746–4755.
- London, M., Larkum, M., Hausser, M., 2008. Predicting the synaptic information efficiency in cortical layer 5 pyramidal cells using a minimal integrate-and-fire model. *Biological Cybernetics* 99 (4–5), 393–401.
- McCulloch, W., Pitts, W., 1943. A logical calculus of the ideas imminent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115–133.
- Panzeri, S., Schultz, S.R., Treves, A., Rolls, E.T., 1999. Correlation and the encoding of information in the nervous system. *Proceedings of the Royal Society London B* 266, 1001–1012.
- Polsky, A., Mel, B., Schiller, J., 2004. Computational subunits in thin dendrites of pyramidal cells. *Nature Neuroscience* 7, 621–627.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R., Bialek, W., 1997. *Spikes. Exploring the Neural Code*. MIT Press, Cambridge, MA.
- Salinas, E., Bentley, N., 2007. A simple measure of the coding efficiency of a neuronal population. *Biosystems* 380 (1–3), 16–23.
- Shannon, C., 1948. A mathematical theory of communication. *Bell System Technical Journal* 27, 379–423, 623–656.
- Strong, S., Koberle, R., de Ruyter van Steveninck, R., Bialek, W., 1998. Entropy and information in neural spike trains. *Physical Review Letters* 80, 3197–3200.
- Victor, J., 2006. Approaches to information-theoretic analysis of neural activity. *Biological Theory* 1, 302–316.