Article

Signal fluctuations and the Information Transmission Rates in binary communication channels

Agnieszka Pregowska [1],[†] (ID)

[1] Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawinskiego 5B, 02–106 Warsaw, Poland; aprego@ippt.pan.pl

* Correspondence: aprego@ippt.pan.pl; Tel.:https://www.ippt.pan.pl/staff.html?idj5=aprego

Abstract: (1) Background: In nervous system information is conveyed by a sequence of action potentials, called spikes-trains. As MacKay and McCulloch proposed, spike-trains can be represented as bits sequences coming from Information Sources. Previously, we studied relations between Information Transmission Rates (ITR) carried out by the spikes, their correlations, and frequencies. Here, we concentrate on the problem of how spikes fluctuations affect ITR. (2) Methods: The Information Theory Method developed by Shannon is applied. Information Sources are modeled as stationary stochastic processes. We assume such sources as two states Markov processes. As a spike-trains' fluctuation measure, we consider the Standard Deviation σ, which, in fact, measures average fluctuation of spikes around the average spike frequency. (3) Results: We found that character of ITR and signal fluctuations relation strongly depends on the parameter s which is a sum of transitions probabilities from no spike state to spike state and vice versa. It turned out that for smaller s ($s < 1$) the quotient $\frac{ITR}{\sigma}$ has a maximum and can tend to zero depending on transition probabilities. While for s large enough ($1 < s$) the $\frac{ITR}{\sigma}$ is separated from 0 for each s. Similar behavior was observed also when we replaced Shannon entropy terms in Markov entropy formula by their approximation with polynomials. We also show that the ITR quotient by Variance behaves in a completely different way. (4) Conclusions: Our results show that for large transition parameter s the Information Transmission Rate by σ will never decrease to zero. Specifically, for $1 < s < 1.7$ the ITR will be always, independently on transition probabilities which form this s, above the level of fluctuations, i.e. in this case we have $\sigma < ITR$. Thus, we conclude that in a more noisy environment, to get appropriate reliability and efficiency of transmission, Information Sources with higher tendency of transition from the state no spike to spike state and vice versa should be applied.

Keywords: information source; information transmission rate; fluctuations; entropy; standard deviation; spike-trains

---

1. Introduction

Information transmission processes in natural environments are usually affected by signals fluctuations due to the presence of noise-generating factors [1,2]. It is especially visible in biological systems, in particular in signal processing in the brain [3–5]. The physical information carriers in the brain are small electrical currents [6]. Specifically, the information is carried by sequences of action potentials also called spikes-trains. Assuming some time resolution MacKay and McCulloch proposed a natural encoding method that associates to each spike-train a binary sequence [7]. Thus, the information is represented by a sequence of bits which, from a mathematical point of view, can be treated as a trajectory of some stochastic process [8,9].

In 1948 C. Shannon developed his famous Communication Theory where he introduced the concept of information and its quantitative measure [10]. The occurrences of both inputs transmitted through a

communications channel and output symbols are described by sequences of random variables that define already stochastic processes and form some Information Sources [8,11]. Following this line, to characterize the amount of information transmitted per symbol the Information Transmission Rate (ITR) is applied.

Spike-trains Information Sources are often modeled as Poisson point processes [12,13]. On the other hand, it is known that such processes exhibit Markov properties [14,15]. This is because in these processes when describing spikes arrival times, current time and the time from the last spike is primarily taken into account [16].

Complex systems dynamics, from financial markets [17] to the neural networks of living beings[18], can be characterized, mostly due to the presence of noise, by using fluctuations, variations, or other statistical tools [19]. The natural measure of fluctuations should, in general, reflect oscillations around the mean average value of the signal. Therefore, in most systems in physics, economics, fluid mechanics, fluctuations are most often quantifying using the Standard Deviation [20–22].

In this paper, we analyze the relationship between the Information Transmission Rate of signals coming from time-discrete two states Markov Information Source and these signals fluctuations. As a spike-trains' fluctuation measure, we consider already the Standard Deviation of encoded spikes. Moreover, to get the better insight we have also analyzed the case when the ITR is referred to the signals Variance V instead to the Standard Deviation σ.

Our previous research, when we studied the properties of neural coding, shows that neural binary coding cannot be captured by straightforward correlations between input and output signals [23]. In [24,25] it was found that a key role in assessing the information sent by Markov type Information Sources in dependence on both Firing Rate and signals correlations plays the jumping (transition) parameter s, which is the sum of transition probabilities from the no-spike state to the spike state and *vice versa*. Here, we found that also the character of the relation between ITR and signal fluctuations strongly depends on the parameter s. It turned out that for small s ($s < 1$) the quotient $\frac{ITR}{\sigma}$ has a maximum and tends to zero when the probability of transition from no spike state to spike state never reaches 0. While for large enough s the quotient $\frac{ITR}{\sigma}$ is limited from below. We observed that similar behavior of $\frac{ITR}{\sigma}$ is also when we replaced (approximate) Shannon entropy formula by appropriate polynomials.

On the other hand, we found that when we refer the quotient $\frac{ITR}{\sigma}$ to σ, i.e. when we consider, in fact, the quotient $\frac{ITR}{V}$ this quotient behaves in a completely different way. This behavior is not regular. Specifically, we observed that for $1 < s$ there is some range of parameter s for which $\frac{ITR}{V}$ has a few local extremas, in opposition to the case $\frac{ITR}{\sigma}$.

The paper is organized as follows. In Section 2, we briefly recall Shannon Information Theory concepts (entropy, information, binary Information Sources, Information Transmission Rate), and fluctuation measure (Standard Deviation and Root Mean Square). In Section 3 we analyzed the quotients $\frac{ITR}{\sigma}$ and $\frac{ITR}{V}$. Section 4 contains the discussion and final conclusions.

2. Theoretical Background and Methods

To introduce the necessary notation, we briefly recall Shannon Information Theory's basic concepts [8,10,11] i.e. Information, Entropy, Information Source, and Information Transmission Rate.

2.1. Shannon's Entropy and Information Transmission Rate

Let Z^L be a set of all words of length L, built of symbols (letters) from some finite alphabet Z. Each word $w \in Z^L$ can be treated as an encoded message sent by Information Source **Z** being a stationary

stochastic process. If $P(w)$ denotes the probability that the word $w \in Z^L$ already occurs, then the information in the Shannon sense carried by this word is defined as

$$I(w) := -\log_2 P(w). \tag{1}$$

This means that less probable events carry more information. Thus, the average information of the random variable $\mathbf{Z}^L$ associated with the words of length L is called the Shannon block entropy and is given by

$$H(\mathbf{Z}^L) := -\sum_{w \in Z^L} P(w) \log_2 P(w). \tag{2}$$

The appropriate measure for estimation of transmission efficiency of an Information Source $\mathbf{Z}$ is the information transmitted on average by a single symbol, i.e. ITR [8,11]

$$ITR^{(L)}(\mathbf{Z}) := \frac{1}{L} H(\mathbf{Z}^L) \tag{3}$$

$$ITR(\mathbf{Z}) = \lim_{L \to \infty} \frac{1}{L} H(\mathbf{Z}^L). \tag{4}$$

This limit exists if and only if the stochastic process $\mathbf{Z}$ is stationary [8].

In the special case of a two-letters alphabet $Z = \{0, 1\}$ and the length of words $L = 1$ we introduce the following notation

$$H_2(p) := H(\mathbf{Z}^1) = -p \log_2 p - (1 - p) \log_2 (1 - p). \tag{5}$$

where $P(1) = p, P(0) = 1 - p$ are associated probabilities. This is, in fact, the formula for the entropy rate of a Bernoulli source [11]. Index 2 in (5) indicates that we consider logarithm wit base 2 what means that we consider the information expressed in bits.

2.2. Information Sources

In general, Information Sources are modeled as stationary stochastic processes [8,11]. The information is represented by trajectories of such processes. Here, to study the relation between Information Transmission Rate (ITR) and trajectories fluctuations, we consider Information Sources which are modeled as two-states Markov processes. The trajectories of these processes can be treated as encoded spike-trains [3,9,26]. The commonly accepted natural encoding procedure leads to binary sequences [9,26]. Spike-trains are, in fact, the main objects to carry information [3,6]. We additionally consider among the Markov processes as a special case the Bernoulli processes.

2.2.1. Information Sources – Markov Processes

We consider time-discrete, two-states Markov process $\mathbf{M}$, which is defined by a set of conditional probabilities $p_{j|i}$ which describe the transition from state i to state j, where $i, j = 0, 1$, and by the initial probabilities $P_0(0), P_0(1)$. The Markov transition probability matrix $\mathbf{P}$ can be written as

$$\mathbf{P} := \begin{bmatrix} p_{0|0} & p_{0|1} \\ p_{1|0} & p_{1|1} \end{bmatrix} = \begin{bmatrix} 1 - p_{1|0} & p_{0|1} \\ p_{1|0} & 1 - p_{0|1} \end{bmatrix}. \tag{6}$$

Each of the columns of the transition probability matrix $\mathbf{P}$ has to sum to 1 (i.e. it is a stochastic matrix [8]).

The time evolution of the states probabilities is governed by the Master Equation [27]

$$
\begin{bmatrix} P_{n+1}(0) \\ P_{n+1}(1) \end{bmatrix} = \begin{bmatrix} 1 - p_{1|0} & p_{0|1} \\ p_{1|0} & 1 - p_{0|1} \end{bmatrix} \cdot \begin{bmatrix} P_n(0) \\ P_n(1) \end{bmatrix}
\tag{7}
$$

where n stands for time, $P_n(0)$, $P_n(1)$ are probabilities of finding states "0" and "1" at time n, respectively. The stationary solution of (7) is given by

$$
\begin{bmatrix} P_{eq}(0) \\ P_{eq}(1) \end{bmatrix} = \begin{bmatrix} \frac{p_{0|1}}{(p_{0|1}+p_{1|0})} \\ \frac{p_{1|0}}{(p_{0|1}+p_{1|0})} \end{bmatrix}.
\tag{8}
$$

It is known [8,11] that for Markov process **M** the Information Transmission Rate as defined by (4) is of the following form

$$
H^{\mathbf{M}} = P_{eq}(0) \cdot H(p_{1|0}) + P_{eq}(1) \cdot H(p_{0|1})
\tag{9}
$$

In previous papers [23–25], when we studied the relation between *ITRs* and firing rates and when we compared *ITR* for Markov processes and for corresponding Bernoulli processes we have introduced a parameter s, which can be interpreted as the tendency of a transition from the no-spike state ("0") to the spike state ("1") and *vice versa*:

$$
s := p_{0|1} + p_{1|0}
\tag{10}
$$

It turned out that this parameter plays an essential role in our considerations also in this paper. Note that $s = 2 - tr\mathbf{M}$ and $0 \le s \le 2$. One can observe that two-states Markov processes are Bernoulli processes if and only if $s = 1$.

2.2.2. Information Sources – Bernoulli Process case

The Bernoulli processes play a special role among the Markov processes. Bernoulli process is a stochastic stationary process $\mathbf{Z} = (Z_i)$, $i = 1, 2, \ldots$ formed by binary identically distributed and independent random variables Z_i. In the case of the encoded spike-trains, we assume that the corresponding process (to be more precise its trajectories) takes successively the values 1 (when spike has arrived in the bin) or 0 (when spike has not arrived). We assume that for a given size of time-bin applied (this depends in turn on the time resolution assumed), spike trains are encoded [28] in such a way that 1 is generated with probability p, and 0 is generated with probability q, where q is equal to $1 - p$. Following the definition, the Information Transmission Rate (3) of the Bernoulli process is

$$
H^B(p, q) = -p \log_2 p - q \log q = H_2(p).
\tag{11}
$$

2.2.3. Generalized entropy variants

The form of entropy H was derived under assumptions of monotonicity, joint entropy, continuity properties, and Grouping Axiom. In the classical case of the entropy rate $H^{\mathbf{M}}$ for Markov process, in formula (11) the terms $H(p_{1|0})$ and $H(p_{0|1})$ are clearly understood in the Shannon sense (2). To get a better insight into the asymptotic behavior of the relations studied in this paper, we additionally consider formula (11) with H replaced by its Taylor approximation (10 terms). We also studied the interesting case when instead of H we used famous unimodal map $U(p) = 4p(1 - p)$ [29] which is, in fact, close (Figure 1) to H in the supremum norm [30]. This idea is along the research direction related to generalized concepts of entropy developed, starting from Renyai [31], by many authors [32–36]. Figure 1 shows the approximation of entropy (11) by polynomials: unimodal map (black dash line) and 10 first terms in the

Taylor series of H (gray dash-dot line). We also included the square root of the unimodal map (black point line) in this Figure.
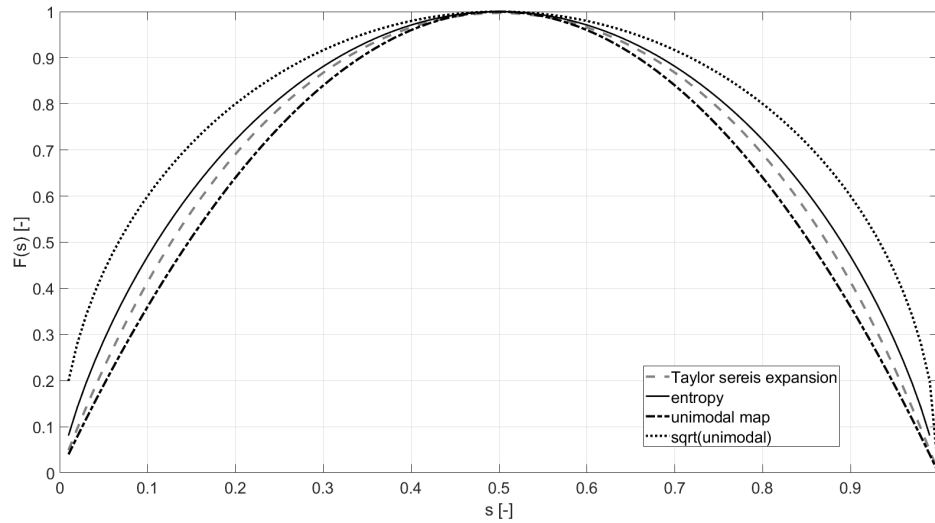

Figure 1. Approximation of the Shannon entropy (black solid lines) using the Taylor series expression (gray dash-dot line, 10 first terms), unimodal function (black dash line), and unimodal map root (black point line).

2.3. Fluctuations measure

It is commonly accepted that for a given random variable $\mathbf{X}$ the fluctuations of values of this random variable around its average can be characterized by the Standard Deviation σ [37]

$$\sigma := \left(\mathbf{E}(\mathbf{X} - \mathbf{EX})^2\right)^{\frac{1}{2}} \tag{12}$$

where symbol $\mathbf{E}$ means the average taken over the probability distribution associated with the values reached by $\mathbf{X}$.

Considering a stochastic process $\mathbf{Y} = (X_k), k = 1, 2, 3, \ldots$, where X_k are random variables each with the same probability distribution as $\mathbf{X}$, the fluctuation of trajectories of this process can be estimated by the Root-Mean-Square (RMS). For a given trajectory $(x_k)_{i=1}^{n}, k = 1, \ldots, n$ RMS is defined as the root from the arithmetic mean value of the squares, i.e.

$$RMS(\mathbf{Y}) := \left(\frac{1}{n}\sum_{k=1}^{n}(x_k - x_{n_{avr}})^2\right)^{\frac{1}{2}} \tag{13}$$

where $x_{n_{avr}}$ is the average value, i.e. $x_{n_{avr}} = \frac{1}{n}\sum_{i=1}^{n} x_k$. Note, that from this formula the form of σ for Markov processes can be derived when using stationary distribution (8) in formula (12).

The Standard Deviation σ for any random variable depends, in fact, not only on its probability distribution but also on the values taken by this random variable. Here we are interested in bits oscillation, i.e. if the spike train occurs or not. Thus, we have limited our considerations to the values 0 and 1.

To get a better insight into the relation between ITR and $signal/bits$ fluctuations we also included an analysis of the quotient $\frac{ITR}{V}$. This is interesting due to the specific form of Variation for the Bernoulli process what leads to interesting observations when consider, for example, the unimodal map to approximate

entropy (5). Moreover, when studying $\frac{ITR}{V}$ we, in fact, refer the quotient $\frac{ITR}{\sigma}$ to σ since we have simply $\frac{\left(\frac{ITR}{\sigma}\right)}{\sigma} = \frac{ITR}{V}$.

3. Results

In this Section, we study the quotients $\frac{ITR}{\sigma}$ and $\frac{ITR}{V}$ as a function of the transition probability $p_{1|0}$ from the state no-spike "0" to the spike state "1" for a fixed parameter s (10). Note, that the probability $0 < p_{1|0} < 1$ and parameter $0 < s < 2$ uniquely determined the transition probability matrix $\mathbf{P}$ (6) and consequently, they completely define the Markov process $\mathbf{M}$, provided that initial probabilities $P_0(0), P_0(1)$ are chosen. Here, as initial probabilities, to get a stationary process, we must assume the probabilities of the form (8).

3.1. Information against fluctuations for two-states Markov processes – general case

We start our considerations from the most general form of the two-states Markov process. To analyze the quotients $\frac{ITR}{\sigma}$ and $\frac{ITR}{V}$ we first express Standard Deviation of Markov process $\mathbf{M}$ in terms of conditional probability $p_{1|0}$ and parameter s.

3.1.1. Standard Deviation in the Markov process case

For a given Markov process $\mathbf{M}$ to evaluate its fluctuation, specifically to address its long time behavior, one considers its corresponding stationary probabilities as defined by (8). Thus, in the limiting case, the Standard Deviation σ for the Markov process can be assumed as

$$\sigma^{\mathbf{M}} = \sqrt{P_{eq}(0) \cdot P_{eq}(1)}. \tag{14}$$

Fixing parameter s and expressing $\sigma^{\mathbf{M}}$ as a function of the conditional probability $p_{1|0}$ we came to the following formula:

$$\sigma_s^{\mathbf{M}}(p_{1|0}) = \sqrt{\frac{p_{0|1}}{s} \cdot \frac{p_{1|0}}{s}} = \frac{\sqrt{(s - p_{1|0})p_{1|0}}}{s}. \tag{15}$$

Note that in the case of Variance $[V_s^{\mathbf{M}}(p_{1|0})] = [\sigma_s^{\mathbf{M}}(p_{1|0})]^2$ we have a polynomial dependence on $p_{1|0}$ (keeping in mind that s is fixed).

3.1.2. Relation between Information Transmission Rate ITR of Markov process and its Standard Deviation

Let's start by establishing the relation between Standard Deviation and ITR for the Bernoulli process. This means that in our notation s is equal to 1. Making use of the classical inequality $x - 1 \geq \ln x$ (for all $x > 0$) and doing a few simple operations one can come to the inequality $2 \cdot \log_2 e \leq \frac{ITR_2(p_{1|0})}{\sigma^2}$. To find the relations between entropy $H^{\mathbf{M}}$ and $\sigma^{\mathbf{M}}$ in more general cases, one can consider the quotient

$$Q_\sigma^{\mathbf{M,s}}(p_{1|0}) := \frac{H^{\mathbf{M,s}}(p_{1|0})}{\sigma_s^{\mathbf{M}}(p_{1|0})}. \tag{16}$$

Note that $Q_\sigma^{\mathbf{M,s}}(p_{1|0})$ is a symmetric function with respect to to the axe $p_{1|0} = \frac{s}{2}$ i.e.

$$Q_\sigma^{\mathbf{M,s}}(p_{1|0}) = Q_\sigma^{\mathbf{M,s}}(s - p_{1|0}). \tag{17}$$

For $0 \le s \le 2$, we consider the quotient $Q_\sigma^{\mathbf{M},\mathbf{s}}(p_{1|0})$ in two cases taking into account the range of $p_{1|0}$

$$\text{A)} \quad 0 \le s \le 1 \quad \text{and this implies} \quad 0 \le p_{1|0} \le s \tag{18}$$

$$\text{B)} \quad 1 < s < 2 \quad \text{and this implies} \quad s - 1 \le p_{1|0} \le 1. \tag{19}$$

Substituting (8), (10) and (14) into (16) we obtain

$$Q_\sigma^{\mathbf{M},\mathbf{s}}(p_{1|0}) := \frac{\frac{p_{0|1}}{s} H(p_{1|0}) + \frac{p_{1|0}}{s} H(p_{0|1})}{\sqrt{\frac{(s-p_{1|0})p_{1|0}}{s}}} \tag{20}$$

and after simple calculations we have

$$\begin{aligned} Q_\sigma^{\mathbf{M},\mathbf{s}}(p_{1|0}) &= \frac{\frac{s-p_{1|0}}{s} H(p_{1|0}) + \frac{p_{1|0}}{s} H(s-p_{1|0})}{\sqrt{\frac{s-p_{1|0}}{s} \cdot \frac{p_{1|0}}{s}}} \\ &= \sqrt{\frac{s-p_{1|0}}{p_{1|0}}} H(p_{1|0}) + \sqrt{\frac{p_{1|0}}{s-p_{1|0}}} H(s-p_{1|0}) \end{aligned} \tag{21}$$

One can check that for smaller $s \in (0,1)$, i.e in case (18), for a given fixed s when $p_{1|0}$ tends to interval bounds 0 or to s, the quotient $Q_\sigma^{H,s}(p_{1|0})$ tends to 0, i.e.:

$$\lim_{p_{1|0}\to 0^+} Q_\sigma^{H,s}(p_{1|0}) = \lim_{p_{1|0}\to s^-} Q_\sigma^{H,s}(p_{1|0}) = 0. \tag{22}$$

By the form of (20) and symmetry property (17) it is clear that the quotient $Q_\sigma^{H,s}(p_{1|0})$ reaches the maximum in the symmetry point $p_{1|0} = \frac{s}{2}$ and it is equal to

$$Q_\sigma^{H,s}\left(\frac{s}{2}\right) = 2H\left(\frac{s}{2}\right). \tag{23}$$

One can check that in the case B) i.e. for $s \in (1,2)$ for a given fixed s when $p_{1|0}$ tends to $s-1$ or to 1 the quotient $Q_\sigma^{H,s}(p_{1|0})$ tends to $\frac{H(s-1)}{\sqrt{(s-1)}}$, i.e.:

$$\lim_{p_{1|0}\to (s-1)^+} Q_\sigma^{H,s}(p_{1|0}) = \lim_{p_{1|0}\to 1^-} Q_\sigma^{H,s}(p_{1|0}) = \frac{H(s-1)}{\sqrt{s-1}}. \tag{24}$$

Thus, we have for $s \in (1,2)$

$$\frac{H(s-1)}{\sqrt{(s-1)}} \le Q_\sigma^{H,s}(p_{1|0}) \le 2H\left(\frac{s}{2}\right). \tag{25}$$
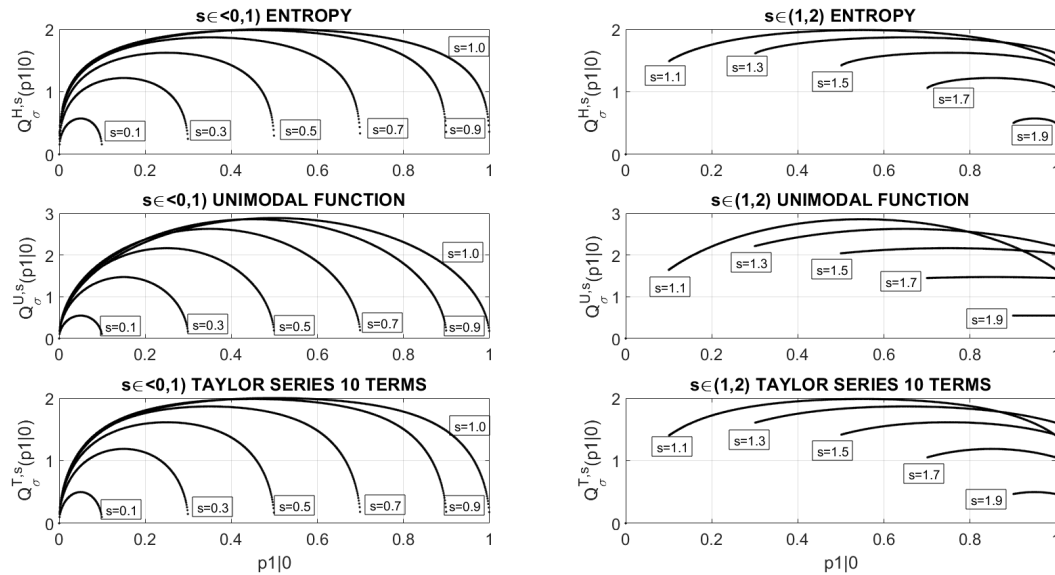
Figure 2. The quotient $\frac{ITR}{\sigma}$ as a function of the transition probability $p_{1|0}$ for chosen values of the jumping parameter s: A) For parameters $0 \le s \le 1$ due to (16) the range of $p_{1|0}$ is $[0, s]$ and B) for $1 < s < 2$ according to (17) the range of $p_{1|0}$ is $1 - s \le p_{1|0} \le 1$. The courses of the quotients $Q_\sigma^{H,s}(p_{1|0})$, $Q_\sigma^{U,s}(p_{1|0})$, $Q_\sigma^{T,s}(p_{1|0})$ for Shannon form, unimodal map, Taylor series being applied as H in formula (9) are presented.

The typical runnings of $Q_\sigma^{H,s}(p_{1|0})$ for some values of the parameter, s are shown in Figure 2. Column A is devoted to lower values of the jumping parameter $0 \le s \le 1$, while column B presents the $Q_\sigma^{H,s}$ courses for higher values of the jumping parameter $1 < s < 2$. Observe, that for $1 < s < 2$ the curves intersect contrary to the case $0 \le s \le 1$. This is mostly since the limiting value (24) is not a monotonic function of s while the maximal value (23) is already monotonic.

Note, that for the approximation of entropy H by polynomials, specifically by unimodal map U and by Taylor series T, the corresponding quotients $Q_\sigma^{U,s}$ B, $Q_\sigma^{T,s}$ behave similarly as for the Shannon form of H (see Figure 2).

3.1.3. Relation between Information Transmission Rate ITR of Markov process and its Variation

To find how the Variation of trajectories of Markov Information Source affects Information Transmission Rate one should consider now a modified quotient

$$Q_\sigma^{\mathbf{M},\mathbf{s}}(p_{1|0}) = \frac{H^{\mathbf{M}}(p_{1|0})}{V(p_{1|0})} = \frac{H^{\mathbf{M}}(p_{1|0})}{P_{eq}(0) \cdot P_{eq}(1)}. \tag{26}$$

Substituting (8) and (10) to (26) we obtain

$$Q_V^{\mathbf{M},\mathbf{s}}(p_{1|0}) = \frac{\frac{p_{0|1}}{s} H(p_{1|0}) + \frac{p_{1|0}}{s} H(p_{0|1})}{\frac{p_{0|1}}{s} \cdot \frac{p_{1|0}}{s}} = s\Big[\frac{H(p_{1|0})}{p_{1|0}} + \frac{H(s - p_{1|0})}{s - p_{1|0}}\Big]. \tag{27}$$

First, observe that clearly as in the standard deviation case we have symmetry property around the value $\frac{s}{2}$, i.e.

$$Q_V^{H,s}(p_{1|0}) = Q_V^{H,s}(s - p_{1|0}). \tag{28}$$

By this symmetry it is clear that $Q_V^{H,s}(p_{1|0})$ reaches extremum at the point $p_{1|0} = \frac{s}{2}$ and it is equal to $4H(\frac{s}{2})$.

Observe, that in the case A), i.e. for a given fixed $s \in (0,1)$, for $p_{1|0}$ tending interval bound i.e. to 0 or $s-$ the quotient $Q_V^{H,s}(p_{1|0})$, in opposite to $Q_\sigma^{T,s}(p_{1|0})$, tends to infinity, i.e.:

$$\lim_{p_{1|0} \to 0+} Q_V^{H,s}(p_{1|0}) = \lim_{p_{1|0} \to s-} Q_V^{H,s}(p_{1|0}) = +\infty. \tag{29}$$

Thus, it is clear that $Q_V^{H,s}(p_{1|0})$ reaches a minimum at the point $p_{1|0} = \frac{s}{2}$.

In the case of B), it turned out that the quotient $Q_V^{H,s}(p_{1|0})$ for any fixed $s \in (1,2)$ is bounded both from below and from above. We have:

$$\lim_{p_{1|0} \to (s-1)+} Q_V^{H,s}(p_{1|0}) = \lim_{p_{1|0} \to 1-} Q_V^{H,s}(p_{1|0}) = s\frac{H(s-1)}{s-1}. \tag{30}$$

Numerical calculations showed that for the parameters $s > s_0$ the point $p_{1|0} = \frac{s}{2}$ is a minimum while for $s < s_0$ at this point, there is a maximum, where the critical parameter $s_0 \approx 1.33$ can be calculated from the equality:

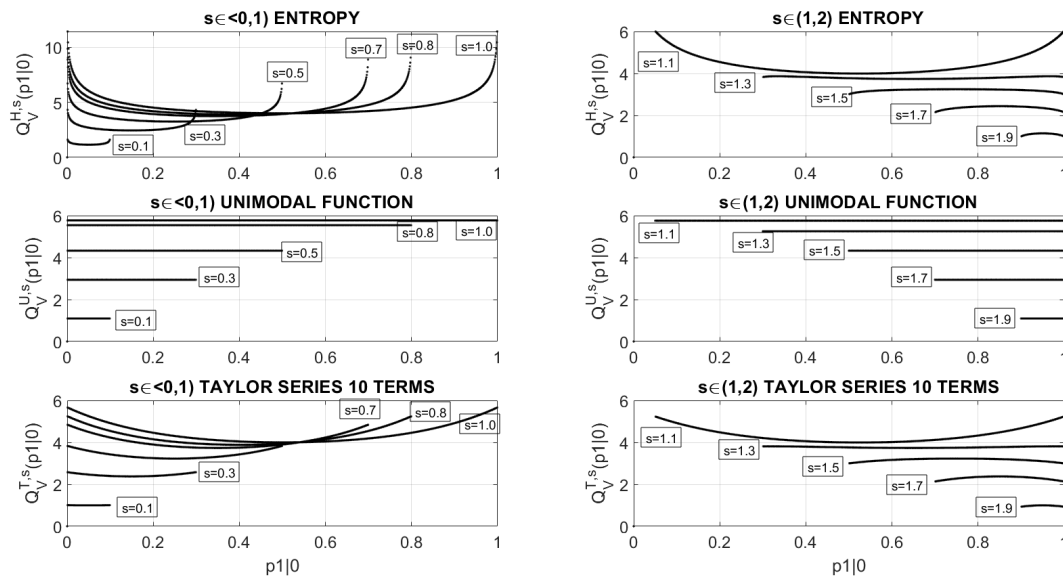$$s_0 \frac{H(s_0 - 1)}{s_0 - 1} = 4H(\frac{s_0}{2}). \tag{31}$$

Figure 3. The quotient $\frac{ITR}{V}$ as a function of the initial probability $p_{1|0}$ for the chosen values of the jumping parameter s: A) For parameters $0 \le s \le 1$ due to (16) the range is $0 < p_{1|0} < s$ and B) For parameters $1 < s < 2$ due to (16) the range is $s - 1 \le p_{1|0} \le 1$.

The typical running of the $Q_V^{H,s}(p_{1|0})$ for some values of the parameter, s is shown in Figure 3. Panel A (left column) is devoted to lower values of the jumping parameter $0 \le s \le 1$, while panel B presents graphs of $Q_V^{H,s}(p_{1|0})$ for higher values of the jumping parameter $1 < s < 2$.

It turned out that the approximation of entropy H by polynomials namely by the unimodal map and by Taylor series leads to the completely different behavior of $Q_V^{H,s}(p_{1|0})$. Note, that for the approximation of H in (2) with the unimodal map the quotient $Q_V^{H,s}(p_{1|0})$, for each s, is a constant and equal to $4s(2 - s)$,

while for the approximation by the Taylor series (10 terms) the quotient $Q_V^{T,s}(p_{1|0})$ preserves a similar courses as for H of Shannon form.

4. Discussion and Conclusions

In this paper, we study relation between the Information Transmission Rate carried out by sequences of bits and these bits fluctuations. These sequences are coming from Information Sources which are modeled by Markov processes. Our results show that the qualitative and quantitative character of the relation between the Information Transmission Rate and signal bits fluctuations strongly depends on the jumping parameter s, which we introduced in our previous papers [24,25]. This parameter characterizes the tendency of the process to transition from state to state. In some sense, it describes the variability of the signals.

It turned out that similarly as in our previous papers when we have studied relation between Information Transmission Rates, spikes correlations, and frequencies of these spikes appearance, the critical value of s is equal to 1 what corresponds to Bernoulli process. For all small s ($s < 1$) the quotient $\frac{ITR}{\sigma}$ can reach 0, while for larger s ($s > 1$) this quotient is always separated from 0. Specifically, for $1 < s < 1.7$ the ITR will be always, independently on transition probabilities which forming this s, above the level of fluctuations (i.e. $\sigma < ITR$). Thus, this shows an interesting fact that for s large enough the information is never completely lost independently on the level of fluctuations.

On the other hand, for each $0 < s < 2$ the quotient $\frac{ITR}{\sigma}$ is limited from above by 2 and it is reached for each s, for $p_{1|0} = \frac{s}{2}$, i.e. it is reached when $p_{1|0} = p_{0|1}$. This means that, when compare ITR to σ, the most effective transmission is for symmetric communication channels. Note, that the capacity $C(s)$ of such channels is already equal to

$$C(s) = 1 - H\left(\frac{s}{2}\right). \tag{32}$$

It turned out that $\frac{ITR}{\sigma}$ for the approximation of Shannon entropy H by polynomials, specifically by the unimodal map and its Taylor series behaves similarly.

For better insight, we also referred ITR to Variance. We observed that the behavior of the $\frac{ITR}{V}$ significantly differs from the behavior of $\frac{ITR}{\sigma}$. For each $s < 1$ the quotient $\frac{ITR}{V}$ can tend to infinity and it is separated from 0. For $1 < s < 2$ it is limited from above and it never reaches 0 for any s. However, it behaves in a more complex way than $\frac{ITR}{\sigma}$ by having even 3 local extreme points, eg. it is visible for $s = 1.3$ and $s = 1.5$. On the other hand approximations of Shannon entropy H by polynomials like the unimodal map or by its Taylor series, contrary to the case of $\frac{ITR}{\sigma}$, lead to a significant qualitative difference between the behavior of $\frac{ITR}{V}$.

To summarize, the results obtained show that for Markov information sources, regardless of the level of fluctuation, the level of Information Transmission Rate does not reduce to zero, provided that the transition parameter s is sufficiently large. This means that to get more reliable communication the spike trains should have a higher tendency of transition from the state no spike to spike state and *vice versa*.

The results are presented in the context of signal processing in the brain, due to the fact that information transmission in the brain is in this case a natural and fundamental phenomena. However, our results have, in fact, a general character and can be applied to any communication systems modeled by two states Markov processes.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.   Weber, A. I.; Shea-Brown, E.; Rieke, F. Identification of multiple noise sources improves estimation of neural responses across stimulus conditions, **2020**, DOI: 10.1101/2020.02.17.951830v1.

2. Huk, A. C.; Hart, E. Parsing signal and noise in the brain. *Science* **2019**, *364(6437)*, 236–237.

3. van Hemmen, J. L.; Sejnowski, T. 23 Problems in Systems Neurosciences (Oxford University Press, Oxford, **2006**.

4. Deco, G; Jirsa, V; McIntosh, A. R.; Sporns, O.; Kötter, R. Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106(25)* 10302–10307.

5. Fraiman, D.; Chialvo, D. R. What kind of noise is brain noise: anomalous scaling behavior of the resting brain activity fluctuations. *Frontiers in Physiology* **2012**, *3(307)*, 1–11.

6. Adrian, E. D.; Zotterman, Y. The impulses produced by sensory nerve endings. *Journal of Physiology* **1926**, *61*, 49–72.

7. MacKay, D.; McCulloch, W. S. The limiting information capacity of a neuronal link. *Bulletin of Mathematical Biology* **1952**, *14*, 127–135.

8. Cover, T. M.; Thomas, J. A. Elements of information theory, Wiley, New York, **1991**.

9. Rieke, F.; Warland, D. D.; de Ruyter van Steveninck, R. R.; Bialek, W. Spikes: Exploring the Neural Code, MIT Press, **1997**.

10. Shannon, C. E. A mathematical theory of communication. *Bell Labs Technical Journal* **1948**, *27(3)*,379–423.

11. Ash, R. B. Information Theory, John Wiley and Sons, New York, **1965**.

12. Teich, M. C.; Khanna, S. M. Pulse-number distribution for the neural spike train in the cat's auditory nerve. *The Journal of the Acoustical Society of America* **1985**, *77(3)*, 1110–1128.

13. Daley, D. H.; Vere-Jones, D. An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods, Springer, Berlin, **2003**.

14. Ross, S. M. Stochastic Processes, Wiley-Interscience, New York, **1996**.

15. Papoulis, A., Pillai, S. U. Probability, Random Variables, and Stochastic Processes, Tata McGraw-Hill Education, **2002**.

16. Kass, R. E.; Ventura, V. A spike-train probability model. *Neural Computation* **2001**, *13*, 1713–1720.

17. Bouchaud, J. P. Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quantitative Finance* **2004**, *4(2)*, 176-–190.

18. Knoblauch, A.; Palm, G. What is signal and what is noise in the brain? *Biosystems* **2005**, *79(1–3)*, 83–90.

19. Witte, R. S.; Witte, J. S. Statistics, John Wiley and Sons Ltd, **2009**.

20. Frisch, U. Turbulence, Cambridge University Press: Cambridge, The United Kingdom, **1995**.

21. Salinas S. R. A. Introduction to Statistical Physics, Springer, **2000**.

22. Kittel C., Elementary Statistical Physics, Dovel Publications, INC., Mineola, New York, USA, **2004**.

23. Pregowska, A.; Szczepanski, J.; Wajnryb, E. Mutual information against correlations in binary communication channels. *BMC Neuroscience* **2015**, *16* 32.

24. Pregowska, A.; Szczepanski, J.; Wajnryb, E. Temporal code versus rate code for binary Information Sources. *Neurocomputing* **2016**, *216*, 756–762.

25. Pregowska, A.; Kaplan, E.; Szczepanski, J. How Far can Neural Correlations Reduce Uncertainty? Comparison of Information Transmission Rates for Markov and Bernoulli Processes. *International Journal of Neural Systems* **2019**, *29(8)*, 1950003–1–13.

26. Amigo, J. M.; Szczepański, J.; Wajnryb, E.; Sanchez-Vives, M. V. *Neural Computation* **2004**, *16(4)*, 717–736.

27. van Kampen, N. G. Stochastic Processes in Physics and Chemistry, Elsevier, **2007**.

28. Bialek, W.; Rieke, F.; de Ruyter van Steveninck, R. R.; Warland, D. D. Reading a neural code. *Science* **1991**, *252(5014)*, 1854–1857.

29. Collet, P.; Eckmann, J. P. Iterated maps on the interval as dynamical systems. Reports on Progress in Physics, Birkhauser, **1980**.

30. Rudin, W. Principles of Mathematical Analysis. New York: McGraw-Hill, **1964**.

31. Renyi A. On measures of information and entropy. Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability **1960**, 547.

32. Amigo, J. M. Permutation complexity in dynamical systems: ordinal patterns, permutation entropy and all that, Springer Science and Business Media, Springer-Verlag Berlin Heidelberg, **2010**.

33.   Crumiller, M.; Knight, B.; Kaplan, E. The measurement of information transmitted by a neural population: promises and challenges. *Entropy* **2013**, *15(9)*, 3507–3527.

34.   Bossomaier, T.; Barnett, L.; Harré, M.; Lizier, J.T. An Introduction to Transfer Entropy, Information Flow in Complex Systems. Springer International Publishing Switzerland, **2016**.

35.   Amigo, J. M.; Balogh, S.G.; Hernandez, S. A Brief Review of Generalized Entropies. *Entropy* **2018**, *20(11)*, 813.

36.   Jetka, T.; Nienałtowski, K.; Filippi, S.; Stumpf, M.P.H.; Komorowski M. An information-theoretic framework for deciphering pleiotropic and noisy biochemical signaling. *Nature Communications* **2018**, *9*, 4591–1–9.

37.   Feller, W. An Introduction to Probability Theory and its Applications, Wiley Series Probability and Statistics John Wiley and Sons, New York, **1958**.