Article

Signal Fluctuations and the Information Transmission Rates in Binary Communication Channels

Agnieszka Pregowska [ID]

Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawinskiego 5B, 02-106 Warsaw, Poland; aprego@ippt.pan.pl; Tel.: +48-22-826-12-81 (ext. 417)

Abstract: In the nervous system, information is conveyed by sequence of action potentials, called spikes-trains. As MacKay and McCulloch suggested, spike-trains can be represented as bits sequences coming from Information Sources (*IS*). Previously, we studied relations between spikes' Information Transmission Rates (*ITR*) and their correlations, and frequencies. Now, I concentrate on the problem of how spikes fluctuations affect *ITR*. The *IS* are typically modeled as stationary stochastic processes, which I consider here as two-state Markov processes. As a spike-trains' fluctuation measure, I assume the standard deviation σ, which measures the average fluctuation of spikes around the average spike frequency. I found that the character of *ITR* and signal fluctuations relation strongly depends on the parameter s being a sum of transitions probabilities from a no spike state to spike state. The estimate of the Information Transmission Rate was found by expressions depending on the values of signal fluctuations and parameter s. It turned out that for smaller $s < 1$, the quotient $\frac{ITR}{\sigma}$ has a maximum and can tend to zero depending on transition probabilities, while for $1 < s$, the $\frac{ITR}{\sigma}$ is separated from 0. Additionally, it was also shown that *ITR* quotient by variance behaves in a completely different way. Similar behavior was observed when classical Shannon entropy terms in the Markov entropy formula are replaced by their approximation with polynomials. My results suggest that in a noisier environment $(1 < s)$, to get appropriate reliability and efficiency of transmission, *IS* with higher tendency of transition from the no spike to spike state should be applied. Such selection of appropriate parameters plays an important role in designing learning mechanisms to obtain networks with higher performance.

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

1. Introduction

Information transmission processes in natural environments are typically affected by signal fluctuations due to the presence of noise-generating factors [1]. It is particularly visible in biological systems, in particular during signal processing in the brain [2–6]. The physical information carriers in the brain are small electrical currents [7]. 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 [8]. 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 [9,10].

In 1948, C. Shannon developed his famous Communication Theory where he introduced the concept of information and its quantitative measure [11]. 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 [9,12]. 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 [13,14]. Poisson point processes provide a good approximation of the experimental data, especially

when the refractory time scale or, more generally, any memory time scale in the spike generation mechanism is short compared to the time scales such as mean interspike interval. The use of Poisson processes to model spike trains has been proposed from the earliest descriptions [15,16] due to the proportional relationship between the mean and variance of multiple neuronal responses. The Poisson property has been observed for many experimental data [17]. On the other hand, it is known that such processes exhibit Markov properties [18,19]. This is because in these processes when describing spikes arrival times, current time, and the time from the last spike are primarily taken into account [20]. There is a number of papers devoted to the modeling of spike-trains by different types of Markov processes (Markov Interval Models, Hidden Markov processes, Poisson Point processes) successfully applied to a variety of experimental data [20–22].

Description of complex systems dynamics, from financial markets [23] to the neural networks of living beings [24,25], requires appropriate mathematical tools. Among them, there are stochastic processes, Information Theory and statistical methods and recently, fuzzy numbers [26,27]. In recent years, to limit or even to exploit the effect of noise and fluctuations on information transmission efficiency extensive effort has been conducted, specifically to design spiking neuronal networks ($SNNs$) with appropriate learning mechanisms [28–32]. Moreover, different models of a neuron have been used to address noise resistance [33]. Traditionally, the complex nature of systems is characterized, mostly due to the presence of noise, by using fluctuations, variations, or other statistical tools [18,19,34–36]. 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 quantifyied using the Standard Deviation [37–39].

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

In the analysis of neuronal coding, specifically when studying neuronal signals, finding relationships between these signals' main characteristics is of great importance [10]. Addressing this issue in our previous papers, I successively analyze the relations:

1. between signal Information Transmission Rates (also Mutual Information) and signal correlations [40]. I show that neural binary coding cannot be captured by straightforward correlations among input and output signals.
2. between signals information transmission rates and signal firing rates (spikes' frequencies) [41]. By examining this dependence, I have found the conditions in which temporal coding rather than rate coding is used. It turned out that this possibility depends on the parameter characterizing the transition from state to state.
3. between information transmission rates of signals (which are (auto)correlated) coming from Markov information sources and information transmission rates of signals coming from corresponding (to this Markov processes) Bernoulli processes. Here, "corresponding" means limiting the Bernoulli process with stationary distributions of these Markov processes [42]. I have shown in the case of correlated signals that the loss of information is relatively small, and thus temporal codes, which are more energetically efficient, can replace rate codes effectively. These results were confirmed by experiments.

In this paper, I consider the next important issue, namely I study the relation between Information Transmission Rates of signals and fluctuations of these signals [40]. I 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. A similar result appears when the Shannon entropy formula is replaced by appropriate polynomials.

On the other hand, I found that when I refer the quotient $\frac{ITR}{\sigma}$ to σ, i.e., when I consider, in fact, the quotient $\frac{ITR}{V}$, this quotient behaves in a completely different way. This behavior is not regular. Specifically, I 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, I briefly recall Shannon Information Theory concepts (entropy, information, binary Information Sources, Information Transmission Rate), and fluctuation measures (Standard Deviation and Root Mean Square). In Section 3, I 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, I briefly recall Shannon Information Theory's basic concepts [9,11,12] 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 **Z** is the information transmitted on average by a single symbol, i.e., *ITR* [9,12]

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

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

This limit exists if and only if the stochastic process **Z** is stationary [9].

In the special case of a two-letter alphabet $Z = \{0, 1\}$ and the length of words $L = 1$, I 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 [12]. Index 2 in (5) indicates that I consider logarithm with base 2, meaning that I consider the information expressed in bits.

2.2. Information Sources

In general, Information Sources are modeled as stationary stochastic processes [9,12]. The information is represented by trajectories of such processes. Here, to study the relation between Information Transmission Rate (ITR) and trajectories fluctuations, I consider Information Sources which are modeled as two-states Markov processes. The trajectories of these processes can be treated as encoded spike-trains [3,10,43]. The commonly accepted natural encoding procedure leads to binary sequences [10,43]. Spike-trains are, in fact,

the main objects that carry information [3,7]. I additionally consider among the Markov processes, as a special case, the Bernoulli processes.

2.2.1. Information Sources—Markov Processes

I consider a time-discrete, two-states Markov process **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 **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 **P** has to sum to 1 (i.e., it is a stochastic matrix [9]).

The time evolution of the states probabilities is governed by the Master Equation [34]

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

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 [9,12] that for Markov process **M**, the Information Transmission Rate as defined by (4) is of the following form

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

In previous papers [40–42], when I studied the relation between *ITRs* and firing rates, and when I compared *ITR* for Markov processes and for corresponding Bernoulli processes, I 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 in this paper also. Note that $s = 2 - tr\mathbf{P}$ and $0 \leq s \leq 2$. 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, I 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). I assume that for a given size of time-bin applied (this depends, in turn, on the time resolution assumed), spike trains are encoded [44] 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

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

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 the 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 ontain a better insight into the asymptotic behavior of the relations studied in this paper, I additionally consider formula (11) with H replaced by its Taylor approximation (10 terms). I also studied the interesting case when instead of H, I used a well known unimodal map $U(p) = 4p(1-p)$ [45] which is, in fact, close (Figure 1) to H in the supremum norm [46]. This idea is along the research direction related to generalized concepts of entropy developed, starting from Renyai [47], by many authors [48–52]. 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). I also included the square root of the unimodal map (black point line) in this Figure.
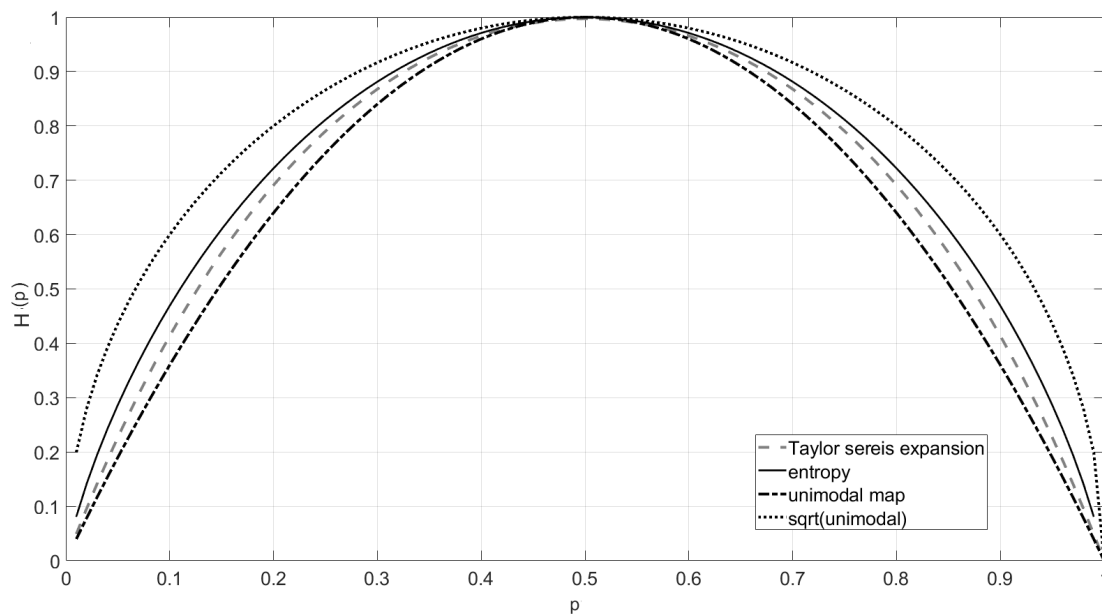
Figure 1. Approximation of the Shannon entropy (black solid lines) using the Taylor series expression (gray dash-dot line, 10 first terms), unimodal function (black dash line), and unimodal map root (black point line).

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 [35]

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

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 $(x_k)_{k=1}^n, k = 1, \ldots, n$ 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., $x_{n_{avr}} = \frac{1}{n}\sum_{k=1}^{n} x_k$. 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, I am interested in bits oscillation, i.e., if the spike train occurs or not. Thus, I have limited our considerations to the values 0 and 1.

To gain a better insight into the relation between *ITR* and *signal/bits* fluctuations, I also included an analysis of the quotient $\frac{ITR}{V}$. This is interesting due to the specific form of Variation for the Bernoulli process, which leads to interesting observations when one considers, for example, the unimodal map to approximate entropy (5). Moreover, when studying $\frac{ITR}{V}$ I, in fact, refer the quotient $\frac{ITR}{\sigma}$ to σ since I have simply $\frac{\left(\frac{ITR}{\sigma}\right)}{\sigma} = \frac{ITR}{V}$.

3. Results

In this section, I 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 **P** (6) and consequently, they completely define the Markov process **M**, provided that initial probabilities $P_0(0), P_0(1)$ are chosen. Here, as initial probabilities, to get a stationary process, I must assume the probabilities of the form (8). To study the relation of *ITR* and σ, I decided to consider the quotients of these two quantities, which seems to be natural and already easy to interpret. This idea was successfully applied in [40–42], which compared ITR and signal correlations, as well as the frequency of jumps. I found that the key role is played by the parameter s, the value of which determines the qualitative and quantitative form of these relations. In fact, *ITR* and σ depend on s, and this is the reason why I analyze the quotient for fixed s.

3.1. Information against Fluctuations for Two-States Markov Processes—General Case

I start my considerations from the most general form of the two-states Markov process. To analyze the quotients $\frac{ITR}{\sigma}$ and $\frac{ITR}{V}$, I first express Standard Deviation of Markov process **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 **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 $p_{1|0}$, I 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 $[V_s^{\mathbf{M}}(p_{1|0})] = [\sigma_s^{\mathbf{M}}(p_{1|0})]^2$, I 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 us 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 $ITR^{\mathbf{M}}$ and $\sigma^{\mathbf{M}}$ in more general cases, one can consider the quotient

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

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

(A)　$0 \leq s \leq 1$　and this implies　$0 \leq p_{1|0} \leq s$ $\tag{18}$

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

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

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

and after simple calculations, I have

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

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 $Q_\sigma^{M,s}(p_{1|0})$ tends to 0, i.e.,

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

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

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

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 $Q_\sigma^{M,s}(p_{1|0})$ tends to $\frac{H(s-1)}{\sqrt{(s-1)}}$, i.e.,

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

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

$$\frac{H(s-1)}{\sqrt{(s-1)}} \leq Q_\sigma^{M,s}(p_{1|0}) \leq 2H\left(\frac{s}{2}\right). \tag{25}$$

Finally, I obtained an interesting estimation of Information Transmission Rate *ITR* by the level of fluctuation σ:

$$\frac{H(s-1)}{\sqrt{s-1}}\sigma_s^M(p_{1|0}) \le ITR^{M,s}(p_{1|0}) \le 2H(\frac{s}{2})\sigma_s^M(p_{1|0}). \tag{26}$$

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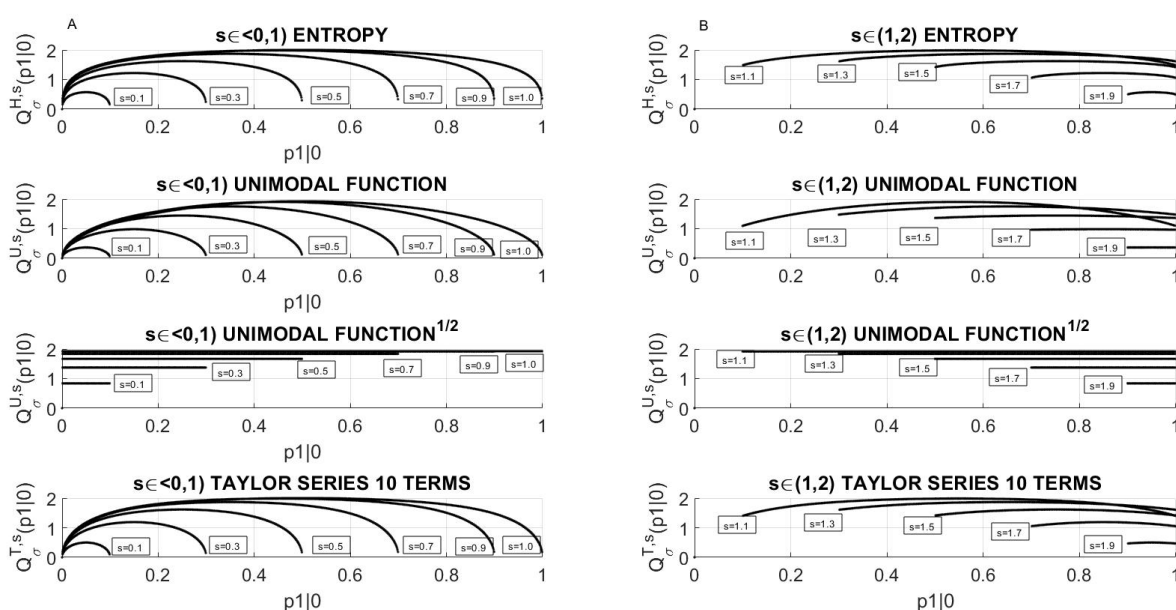
Figure 2. The quotient $\frac{ITR}{\sigma}$ as a function of the transition probability $p_{1|0}$ for chosen values of the jumping parameter *s*: (**A**) For parameters $0 \le s \le 1$ due to (16) the range of $p_{1|0}$ is $[0, s]$ and (**B**) for $1 < s < 2$ according to (17) the range of $p_{1|0}$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

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 the Information Transmission Rate, one should consider a modified quotient

$$Q_\sigma^{\mathbf{M,s}}(p_{1|0}) = \frac{ITR^{\mathbf{M}}(p_{1|0})}{V(p_{1|0})} = \frac{ITR^{\mathbf{M}}(p_{1|0})}{P_{eq}(0) \cdot P_{eq}(1)}. \tag{27}$$

Substituting (8) and (10) to (27) I obtain

$$Q_V^{\mathbf{M,s}}(p_{1|0}) = \frac{\frac{p_{0|1}}{s}H(p_{1|0}) + \frac{p_{1|0}}{s}H(p_{0|1})}{\frac{p_{0|1}}{s} \cdot \frac{p_{1|0}}{s}} = s[\frac{H(p_{1|0})}{p_{1|0}} + \frac{H(s - p_{1|0})}{s - p_{1|0}}]. \tag{28}$$

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

$$Q_V^{M,s}(p_{1|0}) = Q_V^{M,s}(s - p_{1|0}). \tag{29}$$

By this symmetry, it is clear that $Q_V^{M,s}(p_{1|0})$ 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 $Q_V^{M,s}(p_{1|0})$, in opposite to $Q_\sigma^{T,s}(p_{1|0})$, tends to infinity, i.e.,

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

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

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

$$\lim_{p_{1|0} \to (s-1)^+} Q_V^{M,s}(p_{1|0}) = \lim_{p_{1|0} \to 1^-} Q_V^{M,s}(p_{1|0}) = s\frac{H(s-1)}{s-1}. \tag{31}$$

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:

$$s_0 \frac{H(s_0 - 1)}{s_0 - 1} = 4H\left(\frac{s_0}{2}\right). \tag{32}$$

The typical running of the $Q_V^{M,s}(p_{1|0})$ 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 $0 \leq s \leq 1$, while panel B presents graphs of $Q_V^{M,s}(p_{1|0})$ for higher values of the jumping parameter $1 < s < 2$.
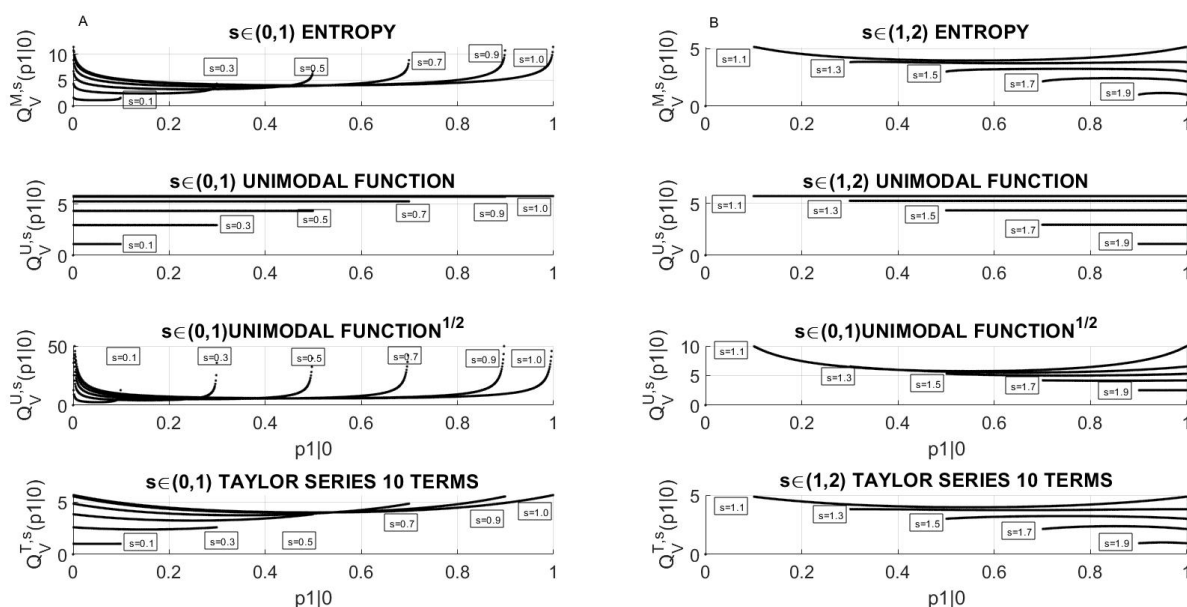
Figure 3. The quotient $\frac{ITR}{V}$ as a function of the initial probability $p_{1|0}$ for the chosen values of the jumping parameter s: (**A**) For parameters $0 \leq s \leq 1$ due to (16) the range is $0 < p_{1|0} < s$ and (**B**) For parameters $1 < s < 2$ due to (16) the range is $s - 1 \leq p_{1|0} \leq 1$. Observe that $\frac{ITR}{V}$ has a completely different course to that of the quotient $\frac{ITR}{\sigma}$ presented in Figure 2.

It turned out that the approximation of entropy H by polynomials, namely by the unimodal map and by Taylor series, leads to completely different behavior of $Q_V^{M,s}(p_{1|0})$. Note, that for the approximation of H in (2) with the unimodal map the quotient $Q_V^{U,s}(p_{1|0})$, 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 course as thay of H of the Shannon form.

4. Discussion and Conclusions

In this paper, I studied the relation between the Information Transmission Rate carried out by sequences of bits and the fluctuations of these bits. These sequences originate from Information Sources which are modeled by Markov processes. During the last 30 years, authors have modeled neuron activity by different variants of Markov processes, e.g., inhomogeneous Markov Interval Models and Hidden Markov Processes [20–22]. The Poisson Point processes commonly used to model experimental data of neuronal activity also exhibit the Markov property [18,19]. Our results show that the qualitative and quantitative character of the relation between the Information Transmission Rate and fluctuations of signal bits strongly depends on the jumping parameter s, which we introduced in our previous papers [41,42]. 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 have studied the relation between Information Transmission Rates, spikes correlations, and frequencies of these spikes appearance, the critical value of s was equal to 1, which corresponds to the Bernoulli process. For all small s ($s < 1$), the quotient $\frac{ITR}{\sigma}$ could reach 0, while for larger s ($s > 1$), this quotient was always separated from 0. Specifically, for $1 < s < 1.7$, the ITR will always be, independent of transition probabilities which form this s, above the level of fluctuations (i.e., $\sigma < ITR$). Thus, this highlights an interesting fact that for large enough s, the information is never completely lost, independent of 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 $p_{1|0} = \frac{s}{2}$ Thus, I have that the maximum is reached when $p_{1|0} = p_{0|1}$. This means that, when one compares ITR to σ, the most effective transmission was for symmetric communication channels. Note, that the capacity $C(s)$ of such channels is already equal to

$$C(s) = 1 - H(\frac{s}{2}). \tag{33}$$

Additionally, 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. Observe, that for all s these quotients, independent of the approximation applied, reach the maximum for $p_{1|0}$ equal to $\frac{s}{2}$ and monotonically increase for $p_{1|0}$ less than $\frac{s}{2}$, while monotonically decrease for $p_{1|0}$ below $\frac{s}{2}$.

For a better insight into the relation between ITR and signal variability, I also referred ITR to Variance. I 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 three local extreme points, e.g., it is visible for $s = 1.3$ and $s = 1.5$. On the other hand, approximations of Shannon entropy H by polynomials such as 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 obtain 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 inequality (26)

allows for estimatio of the amount of information being transmitted by the level of signal fluctuations. Signal fluctuations characterize, in fact, the level of noise.

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 system modeled by two-states Markov processes.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Huk, A.C.; Hart, E. Parsing signal and noise in the brain. *Science* **2019**, *364*, 236–237.
2. Mainen, Z.F.; Sejnowski, T.J. Reliability of spike timing in neocortical neurons. *Science* **1995**, *268*, 1503–1506.
3. van Hemmen, J.L.; Sejnowski, T. *23 Problems in Systems Neurosciences*; Oxford University Press: Oxford, UK, 2006.
4. Deco, G.; Jirsa, V.; McIntosh, A.R.; Sporns, O.; Kötter, R. Key role of coupling, delay, and noise in resting brain fluctuations. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10302–10307.
5. Fraiman, D.; Chialvo, D.R. What kind of noise is brain noise: anomalous scaling behavior of the resting brain activity fluctuations. *Front. Physiol.* **2012**, *3*, 1–11.
6. Gardella, C.; Marre, O.; Mora, T. Modeling the correlated activity of neural populations: A review. *Neural Comput.* **2019**, *31*, 233–269.
7. Adrian, E.D.; Zotterman, Y. The impulses produced by sensory nerve endings. *J. Physiol.* **1926**, *61*, 49–72.
8. MacKay, D.; McCulloch, W.S. The limiting information capacity of a neuronal link. *Bull. Math. Biol.* **1952**, *14*, 127–135.
9. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley: New York, NY, USA, 1991.
10. Rieke, F.; Warland, D.D.; de Ruyter van Steveninck, R.R.; Bialek, W. *Spikes: Exploring the Neural Code*; MIT Press: Cambridge, MA, USA, 1997 .
11. Shannon, C.E. A mathematical theory of communication. *Bell Labs Tech. J.* **1948**, *27*, 379–423.
12. Ash, R.B. *Information Theory*; John Wiley and Sons: New York, NY, USA, 1965.
13. Teich, M.C.; Khanna, S.M. Pulse-number distribution for the neural spike train in the cat's auditory nerve. *J. Acoust. Soc. Am.* **1985**, *77*, 1110–1128.
14. Daley, D.H.; Vere-Jones, D. *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*; Springer: Berlin, Germany, 2003.
15. Werner, G.; Mountcastle, V.B. Neural activity in mechanoreceptive cutaneous afferents: Stimulusresponse relations, weber functions, and information transmission. *J. Neurophysiol.* **1965**, *28*, 359–397.
16. Tolhurst, D.J.; Movshon, J.A.; Thompson, I.D. The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Exp. Brain Res.* **1981**, *41*, 414–419.
17. de Ruyter van Steveninck, R.R.; Lewen, G.D.; Strong, S.P.; Koberle, R.; Bialek, W. Reproducibility and variability in neural spike trains. *Science* **1997**, *275*, 1805–1808.
18. Ross, S.M. *Stochastic Processes*; Wiley-Interscience: New York, NY, USA, 1996.
19. Papoulis, A.; Pillai, S.U. *Probability, Random Variables, and Stochastic Processes*; Tata McGraw-Hill Education: New York, NY, USA, 2002 .
20. Kass, R.E.; Ventura, V. A spike-train probability model. *Neural Comput.* **2001**, *13*, 1713–1720.
21. Radons, G.; Becker, J.D.; Dülfer, B.; Krüger, J. Analysis, classification, and coding of multielectrode spike trains with hidden Markov models. *Biol. Cybern.* **1994**, *71*, 359–373.
22. Berry, M.J.; Meister, M. Refractoriness and neural precision. *J. Neurosci.* **1998**, *18*, 2200–2211.
23. Bouchaud, J.P. Fluctuations and response in financial markets: The subtle nature of 'random' price changes. *Quant. Financ.* **2004**, *4*, 176–190.
24. Knoblauch, A.; Palm, G. What is signal and what is noise in the brain? *Biosystems* **2005**, *79*, 83–90.
25. Mishkovski, I.; Biey, M.; Kocarev, L. Vulnerability of complex networks. *J. Commun. Nonlinear Sci. Numer. Simul.* **2011**, *16*, 341–349.
26. Zadeh, L. Fuzzy sets. *Inf. Control.* **1965**, *8*, 338–353.
27. Prokopowicz, P. The use of ordered fuzzy numbers for modeling changes in dynamic processe. *Inf. Sci.* **2019**, *470*, 1–14.
28. Zhang, M.L.; Qu, H.; Xie, X.R.; Kurths, J. Supervised learning in spiking, neural networks with noise-threshold. *Neurocomputing* **2017**, *219*, 333–349.
29. Lin, Z.; Ma, D.; Meng, J.; Chen, L. Relative ordering learning in spiking neural network for pattern recognition. *Neurocomputing* **2018**, *275*, 94–106.
30. Antonietti, A.; Monaco, J.; D'Angelo, E.; Pedrocchi, A.; Casellato, C. Dynamic redistribution of plasticity in a cerebellar spiking neural network reproducing an associative learning task perturbed by tms. *Int. J. Neural Syst.* **2018**, *28*, 1850020.
31. Kim, R.; Li, Y.; Sejnowski, T.J. Simple framework for constructing functional spiking recurrent neural networks. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 22811–22820.

32. Sobczak, F.; He, Y.; Sejnowski, T.J.; Yu, X. Predicting the fmri signal fluctuation with recurrent neural networks trained on vascular network dynamics. *Cereb. Cortex* **2020**, *31*, 826–844.

33. Qi, Y.; Wang, H.; Liu, R.; Wu, B.; Wang, Y. M.; Pan, G. Activity-dependent neuron model for noise resistance. *Neurocomputing* **2019**, *357*, 240–247.

34. van Kampen, N.G. *Stochastic Processes in Physics and Chemistry*; Elsevier: Amsterdam, The Netherlands, 2007 .

35. Feller, W. *An Introduction to Probability Theory and Its Applications*; Wiley Series Probability and Statistics; John Wiley and Sons: New York, NY, USA, 1958.

36. Salinas, E.; Sejnowski, T.J. Correlated neuronal activity and the flow of neural information. *Nat. Rev. Neurosci.* **2001**, *2*, 539–550.

37. Frisch, U. *Turbulence*; Cambridge University Press: Cambridge, UK, 1995.

38. Salinas, S.R.A. *Introduction to Statistical Physics*; Springer: Berlin, Germany, 2000 .

39. Kittel C. *Elementary Statistical Physics*; Dovel Publications, INC.: Mineola, NY, USA, 2004.

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

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

42. Pregowska, A.; Kaplan, E.; Szczepanski, J. How Far can Neural Correlations Reduce Uncertainty? Comparison of Information Transmission Rates for Markov and Bernoulli Processes. *Int. J. Neural Syst.* **2019**, *29*, doi:10.1142/S0129065719500035.

43. Amigo, J.M.; Szczepański, J.; Wajnryb, E.; Sanchez-Vives, M.V. *Neural Comput.* **2004**, *16*, 717–736.

44. Bialek, W.; Rieke, F.; Van Steveninck, R.D.R.; Warl, D. Reading a neural code. *Science* **1991**, *252*, 1854–1857.

45. Collet, P.; Eckmann, J.P. *Iterated Maps on the Interval as Dynamical Systems*; Reports on Progress in Physics; Birkhauser: Basel, Switzerland, 1980 .

46. Rudin, W. *Principles of Mathematical Analysis*; McGraw-Hill: New York, NY, USA, 1964.

47. Renyi, A. On measures of information and entropy. In Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960 ; pp. 547–561 .

48. Amigo, J.M. *Permutation Complexity in Dynamical Systems: Ordinal Patterns, Permutation Entropy and all that*; Springer Science and Business Media; Springer: Berlin/Heidelberg, Germany, 2010.

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

50. Bossomaier, T.; Barnett, L.; Harré, M.; Lizier, J.T. *An Introduction to Transfer Entropy, Information Flow in Complex Systems*; Springer: Berlin, Germany, 2016.

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

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