

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/322877551>

# Axonal Channel Capacity in Neuro-Spike Communication

Article in IEEE transactions on nanobioscience · February 2018

DOI: 10.1109/TNB.2018.2800899

CITATIONS

25

READS

347

3 authors:



**Keyvan Aghababaiyan**

University of Tehran

18 PUBLICATIONS 225 CITATIONS

SEE PROFILE



**Vahid Shah-Mansouri**

University of Tehran

96 PUBLICATIONS 1,250 CITATIONS

SEE PROFILE



**Behrouz Maham**

Nazarbayev University

156 PUBLICATIONS 2,321 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



A Stochastic Geometric Approach for the Analysis and Design of Cognitive Device-to-Device Networks [View project](#)



Distributed Optimization for Big Data [View project](#)

# Axonal Channel Capacity in Neuro-Spike Communication

Keyvan Aghababaiyan, *Student Member, IEEE*, Vahid Shah-Mansouri<sup>ID</sup>, *Senior Member, IEEE*,  
and Behrouz Maham, *Senior Member, IEEE*

**Abstract**—Novel nano-scale communication techniques are inspired by biological systems. Neuro-spike communication is an example of this communication paradigm which transfers vital information about external and internal conditions of the body through the nervous system. The analysis of this communication paradigm is beneficial to exploit in the artificial neural systems where nano-machines are linked to neurons to treat the neurodegenerative diseases. In these networks, nano-machines are used to replace the damaged segments of the nervous system and they exactly behave like biological entities. In neuro-spike communication, neurons / nano-machines exploit the electro-chemical spikes and molecular communication to transfer information. This communication paradigm can be divided into three main parts, namely the axonal pathway, the synaptic transmission, and the spike generation. In this paper, we focus on the axonal transmission part as a separate channel since the capacity of the axonal pathway has a significant effect on the capacity of neuro-spike communication channel. In thinner axons, the capacity of this part is the bottleneck of the neuro-spike communication channel capacity. Hence, we investigate the restricting factors of the axonal transmission which limit its capacity. We derive the capacity of single-input single-output and multiple-input single-output (MISO) axonal channels. In the MISO case, we investigate the effect of the correlation among inputs on the channel capacity. Moreover, we derive a closed form description for the optimum value of the input spike rate to maximize the capacity of the axonal channel when the information is encoded by firing rate of neurons / nano-machines.

**Index Terms**—Nano-scale communications, neuro-spike communications, axonal transmission, channel capacity.

## I. INTRODUCTION

ELECTROMAGNETIC-BASED communication approaches are unfortunately inapplicable or inappropriate in very small dimensions or in specific environments, such as nano-tunnels and nano-networks [1], [2]. Hence, recent breakthroughs in nano technology have motivated diffusion-based molecular communication. Molecular communication

Manuscript received August 17, 2017; revised October 30, 2017 and December 14, 2017; accepted January 19, 2018. Date of publication February 1, 2018; date of current version March 20, 2018. This paper was presented at the 2017 IEEE International Conference on Communications (ICC 2017) under the title of Axonal Transmission Analysis in Neuro-Spike Communication. (*Corresponding author: Vahid Shah-Mansouri.*)

K. Aghababaiyan and V. Shah-Mansouri are with the School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran 14395-515, Iran (e-mail: aghababaiyan@ut.ac.ir; vmansouri@ut.ac.ir).

B. Maham is with the Department of Electrical and Electronic Engineering, School of Engineering, Nazarbayev University, 010000 Astana, Kazakhstan (e-mail: behrouz.maham@nu.edu.kz).

Digital Object Identifier 10.1109/TNB.2018.2800899

is a new promising field that can be used in nano-networks and provides communication among nano-machines by using molecules as information carriers [3], [4]. The calcium signaling [5] and transmitting information by pheromones and neuro-transmitters [3] are some examples of molecular communication mechanisms. Nervous system is the most complex and advanced intra-body nano-network. Neurons are notable among the cells of the body in their capability to propagate signals rapidly over large distances. They perform this by producing electrical pulses called action potentials or, more simply, spikes that can propagate along nerve fibers. The communication which is used in the neuronal nano-network is called neuro-spike communication. It is a hybrid model that involves both molecular communication in the synaptic transmission part, which occurs between two adjacent neurons, and electrical transmission of action potentials in the axonal pathway. An action potential is a roughly 100 mV oscillation in the electrical potential across the cell velum that lasts for about 1 ms. Action potentials can be propagated over large distances since they are the only form of velum potential fluctuation.

Several work has been performed in neuro-spike communication field. It has been investigated in [6] that the efficiency of neuro-spike communication is acceptable in terms of robustness, speed and reliability. In [7], a mathematical model of how a neuron stochastically processes data and communicates information is introduced and analyzed. A physical channel model for neuro-spike communication has been proposed in [8] to characterize its fundamental properties. This model describes the neuro-spike communication channels in the Cornu Ammonis (CA) region, a specific zone in the hippo-campus location of the brain. In [9], an alternative representation of the neuron-to-neuron communication process has been proposed. A synaptic model has been suggested in [10]. This model shows that redundancy of synapses provides an improvement in the information transmission efficiency. In [8], Balevi and Akan proposed a theoretical lower bound on the capacity of a simple cortical synapse model. In [11], Veletić *et al.* derived theoretical upper bounds on the information capacity of both bipartite and tripartite synapses. In [12], we have derived the capacity bounds of the neuro-spike communication systems by exploiting temporal modulations. Johnson [13] showed that connections among neurons can enhance the capacity. In [14], Malak and Akan studied the multiple-access communications among neurons which are occurred through exchanging of molecules via chemical synapses. The axonal propagation in the CA region and its reliability have been investigated in [15] and [16].

In [17], we have investigated the effect of the axonal noise on the axonal pathway capacity. In [4], the impact of axonal variability on the synaptic response has been investigated. Moreover, an optimum detection policy has been proposed in [18] to detect spikes under different stochastic impairments such as axonal and synaptic noises and random vesicle release. Most of the existing works, such as [8], [18], and [19] have considered the axonal noise as a Bernoulli random variable regardless of the noise spike rate. Moreover, their model is inefficient to describe the modality of the axonal noise.

The axonal noise depends on action potential physical characteristics. The action potential is produced by voltage-gated ion channels, which control the flow of ionic currents through the velum. Therefore, oscillations in voltage-gated ion channels result in probabilistic gating, generating random electrical currents called axonal channel noise [20]. In thin axons, the behavior of individual ion channels can have a substantial impact on the velum potential dynamics owing to the higher input resistance of these axons [21], [22], [23]. Since in these axons fewer channels sustain conduction of action potential, oscillations in a single ion channel have more significant effect on the velum potential. In axons thinner than  $0.3 \mu\text{m}$  diameter, the input resistance is large enough such that spontaneous opening of single  $\text{Na}^+$  channel at the resting potential can generate  $\text{Na}^+$  sparks which can trigger action potentials in the absence of any other inputs. These spontaneous action potentials become exponentially more frequent when the axon diameter reduces. Faisal *et al.* [21] have shown that channel noise sets a lower bound to reliable axonal communication at  $0.08 - 0.1 \mu\text{m}$  diameter. Above this bound, in axons of  $0.1 - 0.5 \mu\text{m}$  diameter, channel noise causes fluctuation in the rising phase of the spike and the resting input resistance of axons. Thus, spikes are added along the axon [22]. Hence, axonal noise has a significant effect on the capacity of neuro-spike communication channel so that it can be the bottleneck of the capacity of neuro-spike communication.

In this paper, we model the input spike train of this channel by a doubly Poisson process which is a Poisson process with a time-varying rate. Moreover, we consider the axonal noise as a Poisson process. We investigate how noise of axons affects the capacity of axonal transmission. We derive the capacity for a simple Single-Input Single-Output (SISO) axonal transmission channel. We obtain the distribution function of input rate which maximizes the mutual information between input and output of the axonal channel. Then, we extend our analytic works by considering a Multiple-Input Single-Output (MISO) axonal transmission channel and investigate the effect of correlation among inputs on the capacity of this channel. We demonstrate when the correlation among inputs increases, the channel capacity decreases. Moreover, in contrast to previous works, we investigate the effect of optimization of the input spike rate to design the efficient input encoding rate with the goal of reduction in the adverse effect of axonal noise. To derive this optimum distribution, we exploit the convexity of the mutual information and Jensens inequality. The results of our work could be beneficial to exploit in the artificial neural systems where nano-machines are linked to neurons to treat the neurodegenerative diseases, e.g., Alzheimers, where some

cells lost their ability to communicate and the nervous system needs to be re-connected.

The rest of this paper is organized as follows. In Section II a simple mathematical model for axonal channel, its input and noise characteristics are presented. In Section III, we derive the capacity of SISO and MISO axonal channel, analytically. In this section, we investigate the effect of correlation of inputs on the capacity of MISO axonal channels. In Section IV, we derive an optimum input spike rate which maximizes the channel capacity. Finally, Section V concludes this paper.

## II. AXONAL TRANSMISSION MODEL

The communication among neurons is called neuro-spike communication, since spikes are used as information carriers from one neuron to another one. Neurons exploit firing rate and temporal coding to transmit information by action potentials. Neural coding refers to the mapping from the stimulus to the response of neuron. The neuro-spike communication includes three main parts. In the first part, i.e., the axonal pathway, spikes are diffused along the axon. The pre-synaptic terminals which are located at the end of the axon release the neuro-transmitter packets to the gaps among neurons, i.e., synapses. The synaptic transmission, i.e., the second part of the neuro-spike communication, begins by releasing the neuro-transmitter packets. Every packet encompasses many neuro-transmitter molecules. By the release of neuro-transmitters, each neuro-transmitter propagates towards the output neuron. There are many post-synaptic terminals at the dendrites, in which the receptors are located [14], and they receive the propagated molecules. When neuro-transmitters are absorbed, the final part of neuro-spike communication, called spike generation, is commenced by the movement of ions. Moving ions excite the velum potential of the output neuron, and lead to the action potential generation. There is a threshold value for each neuron to be depolarized. Thus, action potentials are generated at the neuron axon by any excitation above this threshold level. The arrival of this excitation to the input neuron is a stochastic process. This stochastic model can be assumed as a Poisson process [8].

We use Fano factor to verify that the Poisson process is a realistic model for the input of axonal transmission channel. The Fano factor describes the relevance between the mean spike count over a specified interval and the spike-count variance. Mean spike count  $\bar{n}$  and variance  $\sigma_n^2$  from a wide variety of neuronal recording are fitted to the equation  $\sigma_n^2 = A\bar{n}^B$ , and the multiplier  $A$  and exponent  $B$  are determined. The values of both  $A$  and  $B$  typically lie between 1 and 1.5 [24]. Since the Poisson model predicts  $A = B = 1$ , this indicates that the data shows a higher degree of variability than the Poisson model would predict. The results for spike-count means and variances extracted from recording of MT (Medial temporal) neuron in alert macaque monkeys using a number of various stimuli depict  $A$  and  $B$  values are close to 1 [25]. The MT area is a visual region of the primate cortex where many neurons are sensitive to image motion. Hence, we can conclude that although many neural responses cannot be described by Poisson statistics, it is reassuring to see a case where the Poisson model seems a reasonable approximation.

The refractory effect is often the main cause that spike trains are not described exactly by a Poisson model. For a few millisecond just after an action potential has been fired, it may be virtually impossible to generate another spike. This is called the absolute refractory period. For a longer interval known as the relative refractory period, lasting up to tens of milliseconds after a spike, it is more hard to initiate an action potential. Thus, the Poisson model with refractories provides a reasonably good description for input spike train. This description can be modeled by a Poisson process with a time-variant rate. Moreover, neurons fire action potentials in cluster or bursts of spikes that cannot be described by a Poisson process with a fixed rate. Bursting can be included in a Poisson model by allowing the firing rate to fluctuate in order to describe the high rate of firing during a burst period. Hence, the distribution of bursts themselves can be described by a doubly Poisson process which is called a Cox process [26]. Therefore, we model the input spike train of axonal transmission part by a doubly Poisson process in this paper.

We consider a simple model for axonal transmission channel as a separate part of neuro-spike communication system. The input spike sequence of axonal transmission can be represented as a sum of Dirac delta functions. As mentioned before, the channel input is taken as a random process  $x(t)$  where  $\{x(t), t \in (0, T)\}$  is a doubly Poisson stochastic process with a non-negative random intensity of  $\lambda(t)$  as follows:

$$x(t) = \sum_k \delta(t - t_k), \quad (1)$$

where  $t_k$  is the time that a spike is generated. We define the signal in the pre-synaptic terminal of the input neuron,  $q(t)$ , as follows:

$$q(t) = x(t) + n(t), \quad (2)$$

where  $n(t)$  is the axonal noise. As mentioned before, fluctuations in voltage-gated ion channels result to random electrical currents called axonal channel noise [20]. Thus, some action potentials are added along the axon [22]. We model the axonal noise by arbitrary spikes as:

$$n(t) = \sum_m \delta(t - t_m), \quad (3)$$

where  $t_m$  is the time that an undesirable spike is generated. Since the axonal noise mainly is a result of the stochastic opening of the ion channels, and since there may be numerous ion channels, the axonal noise can be considered as a Poisson process by a constant rate. This rate is a function of the axon thickness and it is more for thinner axons. Fig. 1 shows the rate of spontaneous spikes for some types of axons with different diameters. It can be observed this rate is more significant for axons thinner than  $0.3 \mu\text{m}$ . Since the input resistance in thinner axons is large enough, spontaneous opening of single ion channel at the resting potential can trigger action potentials in the absence of any other inputs. Besides, since stochastic fluctuations of the ion channels are regardless of the input spikes train, the spikes due to the axonal noise are independent of the actual input [4]. Therefore, we can assume  $x(t)$  and  $n(t)$  are independent processes. The capacity of axonal transmission channel is the maximum of the mutual information between  $x(t)$  and  $q(t)$ . Therefore, we investigate

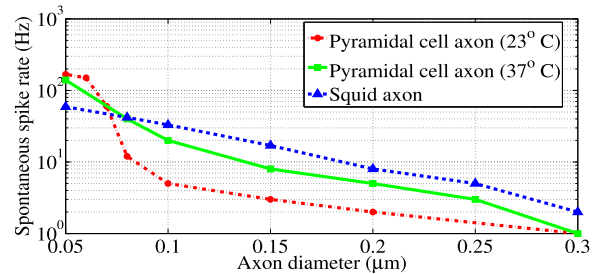


Fig. 1. The spontaneous spike rate for diverse types of axons with different diameters [21].

the capacity of axonal channel and attempt to maximize this capacity by allocating the optimum rate to the input spike train. We can exploit this result to design optimum artificial neural networks by using nano-machines instead of infirm neurons.

### III. CAPACITY OF AXONAL CHANNELS

In the next three subsections, we investigate the capacity of different axonal channels. First, we consider a SISO model for axonal channel to derive the capacity as a function of input and axonal noise rates. Then, we extend our model to a MISO channel to derive the channel capacity and study the effect of correlation between inputs in this case.

#### A. Capacity of SISO Axonal Channel

In this part, we consider a Single-Input and Single-Output model for the axonal channel. A SISO axonal channel model can be used when the dendrites of a neuron receive a spike train from a pre-synaptic terminal and the neuron transmits this spike train along its axon. As we assume, the input of the channel is  $x(t)$  where  $\{x(t), t \in (0, T)\}$  is a doubly Poisson stochastic process with a non-negative random intensity,  $\lambda(t)$ . In SISO model, the input of each neuron is the output spike train of another neuron. Since the firing rate of each neuron is dependent on its own past due to the refractory effect, the  $\lambda(t)$  has memory. In addition, the noise of this channel is the axonal noise which is a Poisson stochastic process with a constant non-negative intensity,  $n$ . As mentioned before, this rate is a function of the axon thickness. Therefore, the output of this channel, i.e.,  $q(t)$  is a doubly Poisson stochastic process with an intensity equal to the sum of  $\lambda(t)$  and  $n$ . If we ignore the brief duration of action potential, an action potential sequence can be characterized simply by a list of the times when spikes occurred. A stochastic process that describes a sequence of events, such as action potential, is called a point process. Let  $Q^T$  be the samples of the output point process  $q(t)$  in the time interval  $(0, T)$  and  $X^T$  be the samples of  $x(t)$  in this interval. We define  $P(Q^T)$  as the probability density function of the  $q(t)$  samples. Since  $\lambda(t)$  is dependent on its own past values, from [29, Th. 2] for doubly Poisson processes, we have

$$P(Q^T) = e^{-\int_0^T (\bar{\lambda}(t)+n)dt + \int_0^T \log(\bar{\lambda}(t)+n)dq(t)}, \quad (4)$$

where  $\bar{\lambda}(t)$  is the estimate of the  $\lambda(t)$  when we have the output intensity. The  $P(Q^T | X^T)$  is defined as the probability density function of the  $q(t)$  samples when the  $x(t)$  is given in the time interval  $(0, T)$ . Then, from [29, Th. 4], we have

$$P(Q^T | X^T) = e^{-\int_0^T (\lambda(t)+n)dt + \int_0^T \log(\lambda(t)+n)dq(t)}. \quad (5)$$

Thus, we can define the mutual information between input and output of our channel as:

$$I(X^T; Q^T) = \mathbb{E}_{\lambda(t)} \left[ \log \left( \frac{P(Q^T | X^T)}{P(Q^T)} \right) \right], \quad (6)$$

where  $\mathbb{E}_{\lambda(t)}[\cdot]$  is the expectation operation over  $\lambda(t)$ . By substituting the expressions in (4) and (5) into (6), we have

$$I(X^T; Q^T) = \mathbb{E}_{\lambda(t)} \left[ - \int_0^T (\lambda(t) - \bar{\lambda}(t)) dt + \int_0^T \log \left( \frac{\lambda(t) + n}{\bar{\lambda}(t) + n} \right) dq(t) \right]. \quad (7)$$

We have  $\mathbb{E}_{\lambda(t)}[\bar{\lambda}(t)] = \mathbb{E}_{\lambda(t)}[\hat{\mathbb{E}}[\lambda(t)|q(t)]] = \mathbb{E}_{\lambda(t)}[\lambda(t)]$ , and thus, we can simplify the expression in (7) as follows:

$$I(X^T; Q^T) = \mathbb{E}_{\lambda(t)} \left[ \int_0^T \log \left( \frac{\lambda(t) + n}{\bar{\lambda}(t) + n} \right) dq(t) \right]. \quad (8)$$

In addition,  $q(t) - \int_0^T \log(\lambda(t) + n) dt$  is a martingale [30, eq. (3.20)], then from a theorem of stochastic integrals [31, p. 437] the expression in (8) is converted to

$$I(X^T; Q^T) = \mathbb{E}_{\lambda(t)} \left[ \int_0^T (\lambda(t) + n) \log \left( \frac{\lambda(t) + n}{\bar{\lambda}(t) + n} \right) dt \right]. \quad (9)$$

Thus, we have

$$I(X^T; Q^T) = \int_0^T (\mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\lambda(t) + n)] - \mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\bar{\lambda}(t) + n)]) dt. \quad (10)$$

Then, by replacing the  $\lambda(t)$  with its estimation when  $q(t)$  is given in the second term, we have

$$\begin{aligned} I(X^T; Q^T) &= \int_0^T (\mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\lambda(t) + n)] \\ &\quad - \mathbb{E}_{\lambda(t)}[\hat{\mathbb{E}}[(\lambda(t) + n) \log(\bar{\lambda}(t) + n) | q(t)]] dt). \end{aligned} \quad (11)$$

Since  $\hat{\mathbb{E}}[(\lambda(t) + n) | q(t)] = \bar{\lambda}(t) + n$ , we can simplify the expression in (11) as follows:

$$I(X^T; Q^T) = \int_0^T (\mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\lambda(t) + n)] - \mathbb{E}_{\lambda(t)}[(\bar{\lambda}(t) + n) \log(\bar{\lambda}(t) + n)]) dt, \quad (12)$$

and the capacity of the SISO axonal channel is defined as the maximum of the expression in (12) over  $\lambda(t)$ .

$$C_{\text{SISO}} = \max_{\lambda(t)} I(Q^T; X^T). \quad (13)$$

Therefore, we solve the following optimization problem subject to average and peak rate constraints for obtaining the capacity of the SISO axonal channel.

$$\begin{aligned} \max_{\lambda(t)} & \left\{ \int_0^T (\mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\lambda(t) + n)] \right. \\ & \quad \left. - \mathbb{E}_{\lambda(t)}[(\bar{\lambda}(t) + n) \log(\bar{\lambda}(t) + n)]) dt \right\}, \\ \text{s.t.} & \frac{1}{T} \mathbb{E}_{\lambda(t)} \left[ \int_0^T \lambda(t) dt \right] \leq \sigma R, \quad 0 \leq \lambda(t) \leq R, \\ & 0 \leq \sigma \leq 1, \end{aligned} \quad (14)$$

where  $R$  is the maximum rate of the input spike train in response to stimuli and  $\sigma$  is the ratio of average to peak rate. We consider an upper bound expression for  $I(Q^T; X^T)$  denoted as  $U(Q^T; X^T)$ . Note that since the function  $f(x) = x \log x$  is convex, Jensen's inequality ( $\mathbb{E}[f(x)] \geq f(\mathbb{E}[x])$ ) gives

$$\begin{aligned} \mathbb{E}_{\lambda(t)} & [(\bar{\lambda}_s(t) + n) \log(\bar{\lambda}_s(t) + n)] \\ & \geq [\mathbb{E}_{\lambda(t)}(\bar{\lambda}_s(t) + n)] \log[\mathbb{E}_{\lambda(t)}(\bar{\lambda}_s(t) + n)], \end{aligned} \quad (15)$$

and since  $\mathbb{E}_{\lambda(t)}[\bar{\lambda}(t)] = \mathbb{E}_{\lambda(t)}[\hat{\mathbb{E}}[\lambda(t)|q(t)]] = \mathbb{E}_{\lambda(t)}[\lambda(t)]$ , we have

$$\begin{aligned} [\mathbb{E}_{\lambda(t)}(\bar{\lambda}_s(t) + n)] \log[\mathbb{E}_{\lambda(t)}(\bar{\lambda}_s(t) + n)] \\ = [\mathbb{E}_{\lambda(t)}(\lambda_s(t) + n)] \log[\mathbb{E}_{\lambda(t)}(\lambda_s(t) + n)]. \end{aligned} \quad (16)$$

Thus, we have

$$\begin{aligned} I(Q^T; X^T) & \leq U(Q^T; X^T) \\ & = \int_0^T (\mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\lambda(t) + n)] \\ & \quad - [\mathbb{E}_{\lambda(t)}(\lambda(t) + n)] \log[\mathbb{E}_{\lambda(t)}(\lambda(t) + n)]) dt. \end{aligned} \quad (17)$$

Hence,  $U(Q^T; X^T)$  is the upper bound on the capacity of the SISO axonal channel. According to (15) and (16), the axonal channel capacity converges to its upper bound when  $\bar{\lambda}(t) = \mathbb{E}_{\lambda(t)}[\lambda(t)]$ . Now, we try to obtain the optimum input intensity distribution which maximizes the capacity of SISO axonal channel. Therefore, to find the optimum input distribution we solve the following optimization problem instead of (14) when we assume  $\bar{\lambda}(t) = \mathbb{E}_{\lambda(t)}[\lambda(t)]$ .

$$\begin{aligned} \max_{\lambda(t)} & \left\{ \int_0^T (\mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\lambda(t) + n)] \right. \\ & \quad \left. - [\mathbb{E}_{\lambda(t)}(\lambda(t) + n)] \log[\mathbb{E}_{\lambda(t)}(\lambda(t) + n)] dt \right\}, \\ \text{s.t.} & \frac{1}{T} \mathbb{E}_{\lambda(t)} \left[ \int_0^T \lambda(t) dt \right] \leq \sigma R, \quad 0 \leq \lambda(t) \leq R, \quad 0 \leq \sigma \leq 1. \end{aligned} \quad (18)$$

We assume  $\mathbb{E}_{\lambda(t)}[\lambda(t)] = R_k$ , where  $0 \leq R_k \leq R$ . For achieving the optimum input distribution which maximizes the capacity of the SISO axonal channel, we solve the dual problem of the optimization problem in (18) as follow:

$$\max_{\lambda(t)} L(\lambda(t), \zeta), \quad (19)$$

where the  $L(\lambda(t), \zeta)$  is defined as:

$$\begin{aligned} L(\lambda(t), \zeta) & = \int_0^T (\mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\lambda(t) + n)] \\ & \quad - [\mathbb{E}_{\lambda(t)}(\lambda(t) + n)] \log[\mathbb{E}_{\lambda(t)}(\lambda(t) + n)] \\ & \quad + \frac{\zeta}{T} (\mathbb{E}_{\lambda(t)}[\lambda(t)] - \sigma R)) dt, \end{aligned} \quad (20)$$

where  $\zeta$  is the Lagrangian multiplier. When  $\mathbb{E}_{\lambda(t)}[\lambda(t)] = R_k$  and  $\mathbb{E}[n] = R_n$ , the expression in (20) is converted to

$$\begin{aligned} \int_0^T (\mathbb{E}_{\lambda(t)}[(\lambda(t) + n) \log(\lambda(t) + n)] \\ - (R_k + R_n) \log(R_k + R_n) + \frac{\zeta}{T} (R_k - \sigma R)) dt. \end{aligned} \quad (21)$$

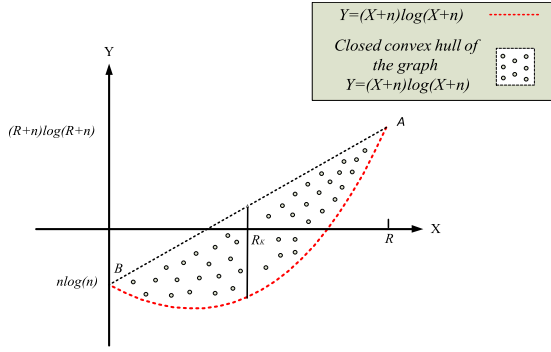


Fig. 2. The closed convex hull of the graph of  $Y = (X + n)\log(X + n)$  for  $0 \leq X \leq R$ .

Note from Fig. 2 that  $0 \leq \lambda(t) \leq R$ , then the possible values of  $\mathbb{E}_{\lambda(t)}[(\lambda(t) + n)\log(\lambda(t) + n)]$  must lie in the closed convex hull of the graph of  $Y = (X + n)\log(X + n)$  for  $0 \leq X \leq R$ . Hence the largest possible values lie on the cord AB. These values can be achieved by using the following distribution:

$$\text{Prob}(\lambda = R) = 1 - \text{Prob}(\lambda = 0) = \alpha, \quad 0 \leq \alpha \leq 1. \quad (22)$$

Since  $\mathbb{E}_{\lambda(t)}[\lambda(t)] = R_k$ , we have

$$\begin{aligned} \mathbb{E}_{\lambda(t)}[\lambda(t)] &= \sum \lambda(t) \text{Prob}(\lambda) \rightarrow R_k = R \text{Prob}(\lambda = R) \\ &= R\alpha \rightarrow \alpha = \frac{R_k}{R}. \end{aligned} \quad (23)$$

Hence

$$\begin{aligned} \max_{\lambda(t)} \mathbb{E}_{\lambda(t)}[(\lambda(t) + n)\log(\lambda(t) + n)] \\ = \left(\frac{R_k}{R}\right)(R + R_n)\log(R + R_n) + \left(1 - \frac{R_k}{R}\right)R_n \log(R_n), \end{aligned} \quad (24)$$

Thus, by inserting (24) in (17) the capacity of the SISO axonal channel is obtained as follows:

$$\begin{aligned} C_{\text{SISO}} = \frac{R_k}{R} (R + R_n) \log(R + R_n) + \left(1 - \frac{R_k}{R}\right) R_n \log(R_n) \\ - (R_k + R_n) \log(R_k + R_n), \end{aligned} \quad (25)$$

where  $R_n$  is the expectation value of the rate of axonal noise.

### B. Capacity of MISO Channel With Independent Inputs

In this subsection, we consider the MISO axonal channel with independent inputs model which can be used when the dendrites of a neuron receive several independent spike trains from different pre-synaptic terminals and the neuron transmits them along its axon. For simplicity, let us consider two independent inputs for this channel. As expressed in Section II, the inputs of the channel are assumed as two independent random processes  $x_1(t)$  and  $x_2(t)$  where  $\{x_i(t), t \in (0, T)\}$ ,  $i = 1, 2$  are two doubly Poisson stochastic processes with non-negative random intensities  $\lambda_1(t)$  and  $\lambda_2(t)$ . In MISO model, the input of each neuron contains the spike trains of multiple neurons. Since the firing rates of neurons are dependent on their own past due to the refractory effect, the  $\lambda_1(t)$  and  $\lambda_2(t)$  have memory. In addition, the noise of this channel is the axonal noise which is a Poisson stochastic process with a constant non-negative intensity  $n$ . Therefore, the output of this channel, i.e.,  $q(t)$  is a doubly Poisson stochastic process with an intensity equal to the sum rates of  $\lambda_1(t)$ ,  $\lambda_2(t)$  and  $n$ . Let  $Q^T$  be the samples of the output point process  $q(t)$  in the

time interval  $(0, T)$ , and also  $X_1$  and  $X_2$  be the samples of  $x_1(t)$  and  $x_2(t)$  in this interval, respectively. We define  $P(Q^T)$  as the probability density function of  $q(t)$  samples. Since  $\lambda_1(t)$  and  $\lambda_2(t)$  are dependent on their own past values, from [29, Th. 2] for doubly Poisson processes, we have

$$P(Q^T) = e^{-\int_0^T (\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n) dt + \int_0^T \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n) dq(t)}, \quad (26)$$

where  $\bar{\lambda}_1(t)$  and  $\bar{\lambda}_2(t)$  are the estimates of  $\lambda_1(t)$  and  $\lambda_2(t)$ , respectively; when we have the output intensity. The  $P(Q^T | X_1, X_2)$  is defined as the probability density function of  $q(t)$  samples when  $x_1(t)$  and  $x_2(t)$  are given in the time interval  $(0, T)$ . Then, from [29, Th. 4], we have

$$\begin{aligned} P(Q^T | X_1, X_2) \\ = e^{-\int_0^T (\lambda_1(t) + \lambda_2(t) + n) dt + \int_0^T \log(\lambda_1(t) + \lambda_2(t) + n) dq(t)}. \end{aligned} \quad (27)$$

Thus, we can define the mutual information between the inputs and the output of this channel as:

$$I(X^T; Q^T) = \mathbb{E}_{\Lambda} \left[ \log \left( \frac{P(Q^T | X_1, X_2)}{P(Q^T)} \right) \right], \quad (28)$$

where  $\mathbb{E}_{\Lambda}[\cdot]$  is the expectation operation over the vector  $\Lambda = [\lambda_1(t), \lambda_2(t)]$ . By substituting the expressions in (26) and (27) into (28), we have

$$\begin{aligned} I(X^T; Q^T) \\ = \mathbb{E}_{\Lambda} \left[ -\int_0^T (\lambda_1(t) - \bar{\lambda}_1(t)) dt - \int_0^T (\lambda_2(t) - \bar{\lambda}_2(t)) dt \right] \\ + \mathbb{E}_{\Lambda} \left[ \int_0^T \log \left( \frac{\lambda_1(t) + \lambda_2(t) + n}{\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n} \right) dq(t) \right]. \end{aligned} \quad (29)$$

Since  $\mathbb{E}_{\Lambda}[\bar{\lambda}_1(t) + \bar{\lambda}_2(t)] = \mathbb{E}_{\Lambda}[\hat{\mathbb{E}}[\lambda_1(t) + \lambda_2(t) | q(t)]] = \mathbb{E}_{\Lambda}[\lambda_1(t) + \lambda_2(t)]$ , and thus, we can simplify the expression in (29) as follows:

$$I(X^T; Q^T) = \mathbb{E}_{\Lambda} \left[ \int_0^T \log \left( \frac{\lambda_1(t) + \lambda_2(t) + n}{\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n} \right) dq(t) \right]. \quad (30)$$

In addition,  $q(t) - \int_0^T \log(\lambda_1(t) + \lambda_2(t) + n) dt$  is a martingale [30, eq. (3.20)], then from a theorem of stochastic integrals [31, p. 437] the expression in (30) is converted to

$$\begin{aligned} I(X^T; Q^T) = \mathbb{E}_{\Lambda} \left[ \int_0^T (\lambda_1(t) + \lambda_2(t) + n) \right. \\ \left. \times \log \left( \frac{\lambda_1(t) + \lambda_2(t) + n}{\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n} \right) dt \right]. \end{aligned} \quad (31)$$

Then, we have

$$\begin{aligned} I(X^T; Q^T) \\ = \int_0^T (\mathbb{E}_{\Lambda}[\lambda_1(t) + \lambda_2(t) + n] \log(\lambda_1(t) + \lambda_2(t) + n) \\ - \mathbb{E}_{\Lambda}[\lambda_1(t) + \lambda_2(t) + n] \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n)) dt. \end{aligned} \quad (32)$$

Then, by replacing the  $\lambda_i(t)$ ,  $i = 1, 2$  with their estimations when  $q(t)$  is given in the second term, we have

$$\begin{aligned} I(X^T; Q^T) = \int_0^T (\mathbb{E}_{\Lambda}[\lambda_1(t) + \lambda_2(t) + n] \log(\lambda_1(t) + \lambda_2(t) + n) \\ - \mathbb{E}_{\Lambda}[\hat{\mathbb{E}}[\lambda_1(t) + \lambda_2(t) + n] \\ \times \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n) | q(t)]) dt. \end{aligned} \quad (33)$$

Since  $\mathbb{E}_{\lambda_i(t)} [(\lambda_1(t) + \lambda_2(t) + n) | q(t)] = \bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n$ , we can simplify the expression in (33) as follows:

$$\begin{aligned} I(X^T; Q^T) &= \int_0^T (\mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)] \\ &\quad - \mathbb{E}_\Lambda [(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n) \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n)]) dt, \end{aligned} \quad (34)$$

and the capacity of the MISO axonal channel is defined as the maximum of the expression presented in (34) over  $\Lambda = [\lambda_1(t), \lambda_2(t)]$ .

$$C_{\text{MISO}} = \max_{\Lambda} I(Q^T; X^T). \quad (35)$$

Therefore, we solve the following optimization problem subject to average and peak rate constraints, for obtaining the capacity of the MISO axonal channel with independent inputs.

$$\begin{aligned} \max_{\Lambda} \{ &\int_0^T (\mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)] \\ &- \mathbb{E}_\Lambda [(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n) \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n)]) dt, \\ \text{s.t. } &\frac{1}{T} \mathbb{E}_\Lambda \left[ \int_0^T \lambda_1(t) + \lambda_2(t) dt \right] \leq \sigma R, \quad 0 \leq \lambda_1(t) \leq R, \\ &0 \leq \lambda_2(t) \leq R, \quad 0 \leq \sigma \leq 1, \end{aligned} \quad (36)$$

where  $R$  is the maximum rate of the input spike trains in response to stimuli and  $\sigma$  is the ratio of average to peak rate. We consider an upper bound for  $I(Q^T; X^T)$  as  $U(Q^T; X^T)$ . Note that since the function  $f(x) = x \log x$  is convex, Jensen's inequality gives

$$\begin{aligned} \mathbb{E}_\Lambda [(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n) \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n)] \\ \geq [\mathbb{E}_\Lambda (\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n)] \log [\mathbb{E}_\Lambda (\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n)], \end{aligned} \quad (37)$$

and since  $\mathbb{E}_\Lambda [\bar{\lambda}_i(t)] = \mathbb{E}_\Lambda [\hat{\mathbb{E}}[\lambda_i(t) | q(t)]] = \mathbb{E}_\Lambda [\lambda_i(t)]$ , we have

$$\begin{aligned} [\mathbb{E}_\Lambda (\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n)] \log [\mathbb{E}_\Lambda (\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + n)] \\ = [\mathbb{E}_\Lambda (\lambda_1(t) + \lambda_2(t) + n)] \log [\mathbb{E}_\Lambda (\lambda_1(t) + \lambda_2(t) + n)], \end{aligned} \quad (38)$$

and thus, we have

$$\begin{aligned} I(Q^T; X^T) &\leq U(Q^T; X^T) \\ &= \int_0^T (\mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)] \\ &\quad - [\mathbb{E}_\Lambda (\lambda_1(t) + \lambda_2(t) + n)] \log [\mathbb{E}_\Lambda (\lambda_1(t) + \lambda_2(t) + n)]) dt. \end{aligned} \quad (39)$$

Thus,  $U(Q^T; X^T)$  is the upper bound on the capacity of the MISO axonal channel. According to (37) and (38), the axonal channel capacity converges to its upper bound when  $\bar{\lambda}_i(t) = \mathbb{E}_\Lambda [\lambda_i(t)]$ . Now, we try to obtain the optimum input intensity distribution which maximizes the capacity of MISO axonal channel. Therefore, to find the optimum input distribution we solve the following optimization problem instead of (36) when

we assume  $\bar{\lambda}_i(t) = \mathbb{E}_\Lambda [\lambda_i(t)]$ .

$$\begin{aligned} \max_{\Lambda} \{ &\int_0^T (\mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)] \\ &- [\mathbb{E}_\Lambda (\lambda_1(t) + \lambda_2(t) + n)] \\ &\quad \times \log [\mathbb{E}_\Lambda (\lambda_1(t) + \lambda_2(t) + n)]) dt, \\ \text{s.t. } &\frac{1}{T} \mathbb{E}_\Lambda \left[ \int_0^T \lambda_1(t) + \lambda_2(t) dt \right] \leq \sigma R, \quad 0 \leq \lambda_1(t) \leq R, \\ &0 \leq \lambda_2(t) \leq R, \quad 0 \leq \sigma \leq 1. \end{aligned} \quad (40)$$

We assume  $\mathbb{E}_\Lambda [\lambda_i(t)] = R_{k_i}$ , where  $0 \leq R_{k_i} \leq R$ . For achieving the optimum input distribution which maximizes the capacity of the MISO axonal channel, we solve the dual problem of the optimization problem in (40) as follow:

$$\max_{\Lambda} L(\Lambda, \zeta), \quad (41)$$

where the  $L(\Lambda, \zeta)$  is defined as:

$$\begin{aligned} L(\Lambda, \zeta) &= \int_0^T (\mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)] \\ &\quad - [\mathbb{E}_\Lambda (\lambda_1(t) + \lambda_2(t) + n)] \log [\mathbb{E}_\Lambda (\lambda_1(t) + \lambda_2(t) + n)] \\ &\quad + \frac{\zeta}{T} (\mathbb{E}_\Lambda [\lambda_1(t) + \lambda_2(t)] - \sigma R)) dt, \end{aligned} \quad (42)$$

where  $\zeta$  is the Lagrangian multiplier. When  $\mathbb{E}_\Lambda [\lambda_i(t)] = R_{k_i}$  and  $\mathbb{E}[n] = R_n$ , the expression in (42) is converted to

$$\begin{aligned} \int_0^T (\mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)] \\ - (R_{k_1} + R_{k_2} + R_n) \log(R_{k_1} + R_{k_2} + R_n) \\ + \frac{\zeta}{T} (R_{k_1} + R_{k_2} - \sigma R)) dt. \end{aligned} \quad (43)$$

For  $0 \leq R_{k_i} \leq R$  the possible values of  $\mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)]$  must lie in the closed convex hull of the graph of  $Z = (X + Y + n) \log(X + Y + n)$  for  $0 \leq X \leq R$  and  $0 \leq Y \leq R$ . Based on the argumentation was presented in Subsection III-A, the largest possible values for  $\mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)]$  can be achieved by using the following distribution:

$$\begin{aligned} \text{Prob}(\lambda_i = R) = 1 - \text{Prob}(\lambda_i = 0) = \alpha_i, \\ 0 \leq \alpha_i \leq 1, \quad \alpha_1 + \alpha_2 \leq 1. \end{aligned} \quad (44)$$

Since  $\mathbb{E}_\Lambda [\lambda_i(t)] = R_{k_i}$ , we have

$$\begin{aligned} \mathbb{E}_\Lambda [\lambda_i(t)] = \sum \lambda_i(t) \text{Prob}(\lambda_i) \rightarrow R_{k_i} = R \text{Prob}(\lambda_i = R) \\ = R \alpha_i \rightarrow \alpha_i = \frac{R_{k_i}}{R}. \end{aligned} \quad (45)$$

Hence

$$\begin{aligned} \max_{\Lambda} \mathbb{E}_\Lambda [(\lambda_1(t) + \lambda_2(t) + n) \log(\lambda_1(t) + \lambda_2(t) + n)] \\ = \left( \frac{R_{k_1} + R_{k_2}}{R} \right) (R + R_n) \log(R + R_n) \\ + \left( 1 - \frac{R_{k_1} + R_{k_2}}{R} \right) R_n \log(R_n). \end{aligned} \quad (46)$$

Thus, by inserting (46) in (39) the capacity of the MISO axonal channel is obtained as follows:

$$\begin{aligned} C_{\text{MISO (Independent)}} &= \left( \frac{R_{k1}}{R} + \frac{R_{k2}}{R} \right) (R + R_n) \log(R + R_n) \\ &+ \left( 1 - \frac{R_{k1}}{R} - \frac{R_{k2}}{R} \right) R_n \log(R_n) \\ &- (R_{k1} + R_{k2} + R_n) \log(R_{k1} + R_{k2} + R_n), \end{aligned} \quad (47)$$

where  $R_n$  is the expectation value of the rate of axonal noise.

### C. Capacity of MISO With Correlated Inputs

We consider the MISO axonal channel with correlated inputs model which can be used when the dendrites of a neuron receive several correlated spike trains from different pre-synaptic terminals. Then, the neuron transmits them along its axon. The correlated spike trains contain redundant information from a stimulus. Since the axonal pathway is unreliable, the redundancy is used commonly to enhance the axonal transmission reliability. For simplicity, let us consider two inputs for this channel. The inputs of the channel are assumed as two correlated random processes  $x_1(t)$  and  $x_2(t)$  where  $\{x_i(t), t \in (0, T)\}$ ,  $i = 1, 2$  are two point processes. Besides, it is assumed that we can divide each of the inputs in to two independent terms  $s_i(t)$  and  $s_0(t)$ . The  $s_i(t)$  describes independent parts of  $x_1(t)$  and  $x_2(t)$ . On the other hand, the  $s_0(t)$  describes common parts of  $x_1(t)$  and  $x_2(t)$  as follows:

$$x_i(t) = s_i(t) + s_0(t), \quad i = 1, 2. \quad (48)$$

As expressed in Section II, it is supposed that  $s_0(t)$ ,  $s_1(t)$  and  $s_2(t)$  are doubly Poisson stochastic processes with non-negative random intensities  $\lambda_0(t)$ ,  $\lambda_1(t)$  and  $\lambda_2(t)$ , respectively. In MISO model, the input of each neuron contains the spike trains of multiple neurons. In this scenario the spike trains are correlated. Since the firing rates of neurons are dependent on their own past due to the refractory effect, the  $\lambda_0(t)$ ,  $\lambda_1(t)$  and  $\lambda_2(t)$  have memory. We can obtain the correlation coefficient between  $x_1(t)$  and  $x_2(t)$  as follows:

$$\rho = \frac{\lambda_0(t)}{\sqrt{(\lambda_1(t) + \lambda_0(t))(\lambda_2(t) + \lambda_0(t))}}. \quad (49)$$

Moreover, it is assumed that, the noise of this channel is the axonal noise which is a Poisson stochastic process with a constant non-negative intensity  $n$  similar to Subsection III-A. Therefore, the output of this channel, i.e.,  $q(t)$  will be a doubly Poisson stochastic process with an intensity equal to the sum of  $\lambda_0(t)$ ,  $\lambda_1(t)$ ,  $\lambda_2(t)$  and  $n$ . Let  $Q^T$  be the samples of the output point process  $q(t)$  in the time interval  $(0, T)$ . Moreover,  $X_1$  and  $X_2$  be the samples of  $x_1(t)$  and  $x_2(t)$  in this interval, respectively. We define  $P(Q^T)$  as the probability density function of  $q(t)$  samples. Since  $\lambda_0(t)$ ,  $\lambda_1(t)$  and  $\lambda_2(t)$  are dependent on their own past values, from [29, Th. 2] for doubly Poisson processes, we have

$$\begin{aligned} P(Q^T) &= e^{-\int_0^T (\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n) dt + \int_0^T \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n) dq(t)}, \end{aligned} \quad (50)$$

where  $\bar{\lambda}_0(t)$ ,  $\bar{\lambda}_1(t)$  and  $\bar{\lambda}_2(t)$  are the estimates of the  $\lambda_0(t)$ ,  $\lambda_1(t)$  and  $\lambda_2(t)$ , respectively. The  $P(Q^T | X_1, X_2)$  is defined as the probability density function of  $q(t)$  samples when the  $x_1(t)$  and  $x_2(t)$  are given in the time interval  $(0, T)$ . Then, from [29, Th. 4], we have

$$\begin{aligned} P(Q^T | X_1, X_2) &= e^{-\int_0^T (\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) dt + \int_0^T \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) dq(t)}. \end{aligned} \quad (51)$$

Thus, we can define the mutual information between the inputs and the output of this channel as (28). By substituting the expressions which are described in (50) and (51) into (28), we have

$$\begin{aligned} I(X^T; Q^T) &= \mathbb{E}_\Lambda \left[ -\int_0^T (\lambda_0(t) - \bar{\lambda}_0(t)) dt - \int_0^T (\lambda_1(t) - \bar{\lambda}_1(t)) dt \right. \\ &\quad \left. - \int_0^T (\lambda_2(t) - \bar{\lambda}_2(t)) dt \right] \\ &\quad + \mathbb{E}_\Lambda \left[ \int_0^T \log \left( \frac{\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n}{\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n} \right) dq(t) \right], \end{aligned} \quad (52)$$

where  $\mathbb{E}_\Lambda[\cdot]$  is the expectation operation over the vector  $\Lambda = [\lambda_1(t), \lambda_2(t), \lambda_0(t)]$ . Since  $\mathbb{E}_\Lambda[\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t)] = \mathbb{E}_\Lambda[\hat{\mathbb{E}}[\lambda_1(t) + \lambda_2(t) + \lambda_0(t) | q(t)]] = \mathbb{E}_\Lambda[\lambda_1(t) + \lambda_2(t) + \lambda_0(t)]$ , and thus, we can simplify the expression in (52) as follows:

$$\begin{aligned} I(X^T; Q^T) &= \mathbb{E}_\Lambda \left[ \int_0^T \log \left( \frac{\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n}{\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n} \right) dq(t) \right]. \end{aligned} \quad (53)$$

In addition,  $q(t) - \int_0^T \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) dt$  is a martingale [30, eq. (3.20)], then from a theorem of stochastic integrals [31, p. 437] the expression in (53) is converted to

$$\begin{aligned} I(X^T; Q^T) &= \mathbb{E}_\Lambda \left[ \int_0^T (\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \right. \\ &\quad \left. \times \log \left( \frac{\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n}{\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n} \right) dt \right]. \end{aligned} \quad (54)$$

Then, we have

$$\begin{aligned} I(X^T; Q^T) &= \int_0^T (\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \\ &\quad \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] \\ &\quad - \mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \\ &\quad \times \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n)]) dt. \end{aligned} \quad (55)$$

Next, by replacing the  $\lambda_i(t)$ ,  $i = 0, 1, 2$  with their estimations when  $q(t)$  is given in the second term, we have

$$\begin{aligned} I(X^T; Q^T) &= \int_0^T (\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \\ &\quad \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] \\ &\quad - \mathbb{E}_\Lambda[\hat{\mathbb{E}}[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \\ &\quad \times \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n) | q(t)]] dt). \end{aligned} \quad (56)$$



Since  $\hat{\mathbb{E}}[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)|q(t)] = \bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n$ , we can simplify the expression in (56) as follows:

$$I(Q^T; X^T) = \int_0^T (\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] - \mathbb{E}_\Lambda[(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n) \times \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n)]) dt, \quad (57)$$

and the capacity of the MISO axonal channel with correlated inputs is defined as the maximum of the expression presented in (57) over  $\Lambda = [\lambda_1(t), \lambda_2(t), \lambda_0(t)]$ .

$$C_{MISO} = \max_{\Lambda} I(Q^T; X^T). \quad (58)$$

Therefore, we solve the following optimization problem subject to average and peak rate constraints for obtaining the capacity of the MISO axonal channel with correlated inputs.

$$\begin{aligned} \max_{\Lambda} \{ & \int_0^T (\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] \\ & - \mathbb{E}_\Lambda[(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n) \times \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n)]) dt, \\ \text{s.t. } & \frac{1}{T} \mathbb{E}_\Lambda \left[ \int_0^T \lambda_1(t) + \lambda_2(t) + \lambda_0(t) dt \right] \leq \sigma R, \\ & 0 \leq \lambda_1(t) + \lambda_0(t) \leq R, \\ & 0 \leq \lambda_2(t) + \lambda_0(t) \leq R, \quad 0 \leq \sigma \leq 1, \end{aligned} \quad (59)$$

where  $R$  is the maximum rate of the input spike trains in response to stimuli and  $\sigma$  is the ratio of average to peak rate. We consider an upper bound for  $I(Q^T; X^T)$  as  $U(Q^T; X^T)$ . Note that since the function  $f(x) = x \log x$  is convex, Jensen's inequality gives

$$\begin{aligned} & \mathbb{E}_\Lambda[(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n) \times \log(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n)] \\ & \geq [\mathbb{E}_\Lambda(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n) \times \log[\mathbb{E}_\Lambda(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n)]], \end{aligned} \quad (60)$$

and since  $\mathbb{E}_\Lambda[\bar{\lambda}_i(t)] = \mathbb{E}_\Lambda[\hat{\mathbb{E}}[\lambda_i(t)|q(t)]] = \mathbb{E}_\Lambda[\lambda_i(t)]$ , we have

$$\begin{aligned} & [\mathbb{E}_\Lambda(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n) \times \log[\mathbb{E}_\Lambda(\bar{\lambda}_1(t) + \bar{\lambda}_2(t) + \bar{\lambda}_0(t) + n)]] \\ & = [\mathbb{E}_\Lambda(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log[\mathbb{E}_\Lambda(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)]], \end{aligned} \quad (61)$$

and thus, we have

$$\begin{aligned} I(Q^T; X^T) & \leq U(Q^T; X^T) \\ & = \int_0^T (\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] \\ & \quad - [\mathbb{E}_\Lambda(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log[\mathbb{E}_\Lambda(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)]]) dt. \end{aligned} \quad (62)$$

Thus,  $U(Q^T; X^T)$  is the upper bound on the capacity of the MISO axonal channel with correlated inputs. According to (60) and (61), the axonal channel capacity converges to its upper bound when  $\bar{\lambda}_i(t) = \mathbb{E}_\Lambda[\lambda_i(t)]$ . Now, we try to obtain the optimum input intensity distribution which maximizes the capacity of MISO axonal channel with correlated inputs. Therefore, to find the optimum input rate distribution we solve the following optimization problem instead of (59) when we assume  $\bar{\lambda}_i(t) = \mathbb{E}_\Lambda[\lambda_i(t)]$ .

$$\begin{aligned} \max_{\Lambda} \{ & \int_0^T (\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] \\ & - [\mathbb{E}_\Lambda(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log[\mathbb{E}_\Lambda(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)]]) dt, \\ \text{s.t. } & \frac{1}{T} \mathbb{E}_\Lambda \left[ \int_0^T \lambda_1(t) + \lambda_2(t) + \lambda_0(t) dt \right] \leq \sigma R, \\ & 0 \leq \lambda_1(t) + \lambda_0(t) \leq R, \\ & 0 \leq \lambda_2(t) + \lambda_0(t) \leq R, \quad 0 \leq \sigma \leq 1. \end{aligned} \quad (63)$$

We assume  $\mathbb{E}_\Lambda[\lambda_i(t)] = R_{ki}$ , where  $0 \leq R_{ki} \leq R$ . For achieving the optimum input distribution which maximizes the capacity of the MISO axonal channel, we solve the dual problem of the optimization problem in (63) as follow:

$$\max_{\Lambda} L(\Lambda, \zeta), \quad (64)$$

where the  $L(\Lambda, \zeta)$  is defined as:

$$\begin{aligned} L(\Lambda, \zeta) & = \int_0^T (\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] \\ & \quad - [\mathbb{E}_\Lambda(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log[\mathbb{E}_\Lambda(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)]]) \\ & \quad + \frac{\zeta}{T} (\mathbb{E}_\Lambda[\lambda_1(t) + \lambda_2(t) + \lambda_0(t)] - \sigma R) dt, \end{aligned} \quad (65)$$

where  $\zeta$  is the Lagrangian multiplier. When  $\mathbb{E}_\Lambda[\lambda_i(t)] = R_{ki}$  and  $\mathbb{E}[n] = R_n$ , the expression in (65) is converted to

$$\begin{aligned} & \int_0^T (\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] \\ & \quad - (R_{k1} + R_{k2} + R_{k0} + R_n) \times \log(R_{k1} + R_{k2} + R_{k0} + R_n) \\ & \quad + \frac{\zeta}{T} (R_{k1} + R_{k2} + R_{k0} - \sigma R)) dt. \end{aligned} \quad (66)$$

If we assume  $\lambda_1(t) = \lambda_2(t)$ , the  $\lambda_0(t)$  can be wrote as  $\lambda_0(t) = A \times \lambda_i(t)$  for  $i = 1, 2$ , based on (49), where  $A = \frac{\rho}{1-\rho}$ . For  $0 \leq R_{ki} + R_{k0} \leq R$  the possible values of  $\mathbb{E}_\Lambda[(\lambda_1(t) +$

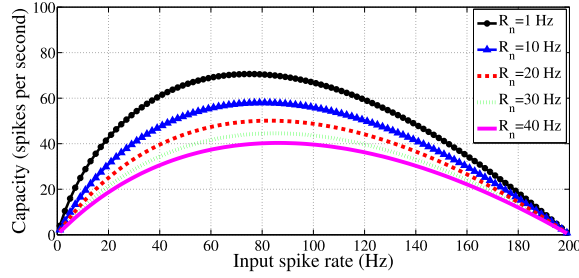


Fig. 3. The capacity of SISO axonal channel versus different input rates for the maximum input rate  $R = 200$  Hz and various rates of axonal noise.

$\lambda_2(t) + \lambda_0(t) + n \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)$  must lie in the closed convex hull of the graph of  $Z = (X+Y+n) \log(X+Y+n)$  for  $0 \leq X \leq (1 - \frac{\rho}{2})R$  and  $0 \leq Y \leq (1 - \frac{\rho}{2})R$ . Based on the argumentation was presented in Subsection III-A, the largest possible values for  $\mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)]$  can be achieved by using the following distribution:

$$\text{Prob}(\lambda_i + \lambda_0 = R) = 1 - \text{Prob}(\lambda_i + \lambda_0 = 0) = \alpha_i, \quad (67)$$

where  $0 \leq \alpha_i \leq 1$  and  $\alpha_1 + \alpha_2 \leq 1$ . Moreover, from (49) we have  $R_{k0} = \frac{\rho}{1-\rho} R_{ki}$ . Thus, we have:

$$\begin{aligned} R_{k1} + R_{k0} &= R \text{Prob}(\lambda_1 + \lambda_0 = R) = R\alpha_1 \rightarrow \alpha_1 \\ &= \frac{1}{1-\rho} \frac{R_{k1}}{R}, \\ R_{k2} + R_{k0} &= R \text{Prob}(\lambda_2 + \lambda_0 = R) = R\alpha_2 \rightarrow \alpha_2 \\ &= \frac{1}{1-\rho} \frac{R_{k2}}{R}. \end{aligned} \quad (68)$$

Hence

$$\begin{aligned} \max_{\Lambda} \mathbb{E}_\Lambda[(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n) \\ \times \log(\lambda_1(t) + \lambda_2(t) + \lambda_0(t) + n)] \\ = (1 - \alpha_1 - \alpha_2) R_n \log R_n \\ + (\alpha_1 + \alpha_2) (R + R_n) \log(R + R_n). \end{aligned} \quad (69)$$

Thus, by inserting (69) in (63) the capacity of the MISO axonal channel with correlated inputs is obtained as follows:

$$\begin{aligned} C_{\text{MISO (Correlated)}} \\ = (1 - \alpha_1 - \alpha_2) R_n \log R_n + (\alpha_1 + \alpha_2) (R + R_n) \\ \times \log(R + R_n) - (R_{k1} + R_{k2} + R_n) \log(R_{k1} + R_{k2} + R_n), \end{aligned} \quad (70)$$

where  $R_n$  is the expectation value of the rate of axonal noise.

#### IV. OPTIMUM INPUTS RATE

In this section, we propose an optimum rate allocation scheme for the input spike train to encode information in a way that this optimum rate enhances the capacity of axonal transmission. Encoding of spikes based on the received stimuli is carried out by input neurons/nano-machines. When we consider the neuro-spike communication channel between two nano-machines, they can design their coding. For example, we assume a neuron transmits a spike train which contains  $R_1$  spikes to inform the occurrence of event  $A$  and transmits no spike to inform occurrence of event  $A'$ . When the nano-machines are replaced instead of neurons, they can alter the rate  $R_1$  to  $R_2$  to inform the occurrence of event  $A$ , so that the

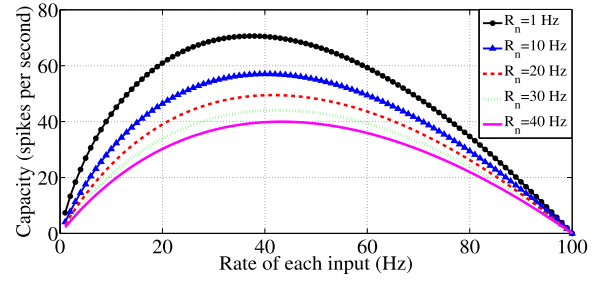


Fig. 4. The capacity of MISO axonal channel versus different inputs rates for the maximum inputs rate  $R = 200$  Hz and various rates of axonal noise.

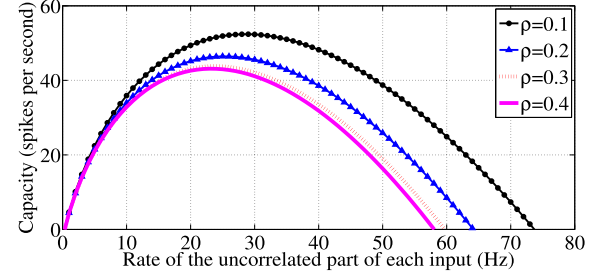


Fig. 5. The capacity of MISO axonal channel with correlated inputs versus different input rates for the maximum input rate  $R = 200$  Hz and the rate of axonal noise  $R_n = 10$  Hz for various inputs correlation coefficients ( $\rho$ ).

average input rate tends to the optimum input rate. Fig. 3 shows the capacity of the SISO axonal channel based on different input rates. It can be observed that there is an optimum input rate which maximizes the channel capacity. For deriving this optimum rate for the SISO channel, we solve the following optimization problem:

$$\max_{R_k} I(Q^T; X^T). \quad (71)$$

Note that  $I(Q^T; X^T)$  is a concave function with respect to  $R_k$ , since its second derivative is negative as  $\frac{\partial^2 I(Q^T; X^T)}{\partial R_k^2} = \frac{-1}{R_k + R_n}$ . Hence, based on the concavity of  $I(Q^T; X^T)$ , we solve the problem  $\frac{\partial I(Q^T; X^T)}{\partial R_k} = 0$  to obtain the optimum input rate. This optimum rate is derived as:

$$R_k^* = (R + R_n) e^{-1 + \frac{R_n}{R} \log\left(1 + \frac{R}{R_n}\right)} - R_n. \quad (72)$$

Fig. 4 shows the capacity of the MISO channel with independent inputs versus different input rates when we set  $R_k = R_{k1} = R_{k2}$  and  $R = R_1 = R_2$  for various rates of the axonal noise. For deriving the optimum input rates for the MISO channel with independent inputs, we solve the following optimization problem:

$$\max_{R_{k1}, R_{k2}} I(Q^T; X^T). \quad (73)$$

By considering  $R_{k1} = L \times R_{k2} = R_k$ , the  $I(Q^T; X^T)$  is concave with respect to  $R_k$ , since its second derivative is negative as  $\frac{\partial^2 I(Q^T; X^T)}{\partial R_k^2} = \frac{-(1+L)^2}{(1+L)R_k + R_n}$ . Hence, we solve the problems  $\frac{\partial I(Q^T; X^T)}{\partial R_{k1}} = 0$  and  $\frac{\partial I(Q^T; X^T)}{\partial R_{k2}} = 0$  to obtain the optimum input rates. These optimum rates are derived as:

$$R_k^* = (R + R_n) e^{-1 + \frac{R_n}{R} \log\left(1 + \frac{R}{R_n}\right)} - R_n. \quad (74)$$

In addition, Fig. 5 shows the capacity of the MISO channel with correlated inputs versus different input rates for various

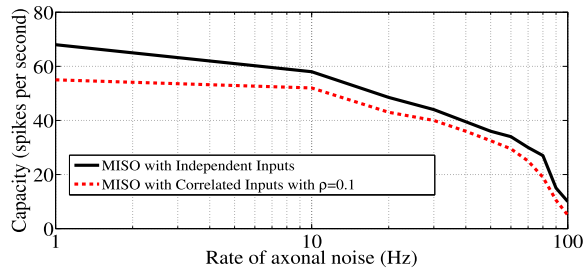


Fig. 6. The capacity of different axonal channels for different rates of axonal noise ( $R_n$ ) when the maximum inputs rates are assumed  $P = 200$  Hz.

correlation coefficients when we set  $R_k = R_{k1} = R_{k2}$ . For deriving the optimum input rate for the MISO channel with correlated inputs, we solve the optimization problem in (73) again. For simplicity, we assume that  $R_{k1} = R_{k2} = R_k$ . Moreover, the rate of common part of inputs is described by  $R_{k0}$ , that we have  $R_{k0} = A \times R_k$ , where  $A = \frac{\rho}{1-\rho}$ . Thus, the  $I(Q^T; X^T)$  is concave with respect to  $R_k$ , since its second derivative is negative as  $\frac{\partial^2 I(Q^T; X^T)}{\partial R_k^2} = \frac{-(2+A)^2}{(2+A)R_k + R_n}$ . Hence, we solve the problems  $\frac{\partial I(Q^T; X^T)}{\partial R_{k1}} = 0$  and  $\frac{\partial I(Q^T; X^T)}{\partial R_{k2}} = 0$  to obtain the optimum input rates. These optimum rates are derived as:

$$R_k^* = \frac{1}{2} (R + R_n) e^{-1 + \frac{1}{1-\rho} \frac{R_n}{R} \log\left(1 + \frac{R}{R_n}\right)} - \frac{1}{2} R_n. \quad (75)$$

Fig. 5 shows that the capacity of MISO axonal channels reduces with increasing the correlation coefficient of the inputs. Thus, redundancy decreases the capacity of axonal channels. Although redundancy reduces the capacity of axonal channel, it is necessary to overcome the unreliability of axonal transmission, especially for thin axons. Fig. 6 compares the capacity of MISO axonal channel with independent inputs and the capacity of MISO axonal channel with correlated inputs ( $\rho = 0.1$ ) for different rate of axonal noise. We can see that the difference between the capacity of the both types of MISO axonal channels is clearer when the rate of noise is low. On the other hand, when the rate of noise is high, the both types of MISO axonal channels behave similarly.

## V. CONCLUSION

In this paper, we derived the capacity of the axonal channel in different scenarios. We showed the capacity of axonal channels is a function of the average and maximum rate of the inputs. We provided a closed form expression for the capacity of axonal channel with different inputs rates for various scenarios such as SISO, MISO with independent inputs and MISO with correlated inputs. As a result, we noticed that the capacity of MISO axonal channels reduces with increasing the correlation coefficient of the inputs. Thus, redundancy decreases the capacity of axonal channels. Moreover, we expressed axonal channels can reach to their maximum capacity by assigning optimum rates to their inputs. In this regard, we presented a closed form formulation for the optimum rate of inputs in different scenarios.

## REFERENCES

[1] T. Palacios, A. Hsu, and H. Wang, "Applications of graphene devices in RF communications," *IEEE Commun. Mag.*, vol. 48, no. 6, pp. 122–128, Jun. 2010.

[2] J. M. Jornet and I. F. Akyildiz, "Graphene-based nano-antennas for electromagnetic nanocommunications in the terahertz band," in *Proc. 4th Eur. Conf. Antennas Propag.*, Barcelona, Spain, Apr. 2010, pp. 1–5.

[3] I. F. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: A new communication paradigm," *Comput. Netw.*, vol. 52, no. 12, pp. 2260–2279, Jun. 2008.

[4] M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 602–611, May 2010.

[5] J. M. Bekkers, G. B. Richerson, and C. F. Stevens, "Origin of variability in quantal size in cultured hippocampal neurons and hippocampal slices," *Proc. Nat. Acad. Sci. USA*, vol. 87, no. 14, pp. 5359–5362, 1990.

[6] P. Dayan and L. Abbott, "Theoretical neuroscience: Computational and mathematical modeling of neural systems," *J. Cognit. Neurosci.*, vol. 15, no. 1, pp. 154–155, 2003.

[7] T. Berger and W. B. Levy, "A mathematical theory of energy efficient neural computation and communication," *IEEE Trans. Inf. Theory*, vol. 56, no. 2, pp. 852–874, Feb. 2010.

[8] E. Balevi and O. B. Akan, "A physical channel model for nanoscale neuro-spike communications," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1178–1187, Mar. 2013.

[9] M. Veletić, P. A. Floor, Z. Babić, and I. Balasingham, "Peer-to-peer communication in neuronal nano-network," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1153–1166, Mar. 2016.

[10] A. Manwani, "Information-theoretic analysis of neuronal communication," Ph.D. dissertation, Dept. Elect. Eng., California Inst. Technol., Pasadena, CA, USA, 2000.

[11] M. Veletić, P. A. Floor, Y. Chahibi, and I. Balasingham, "On the upper bound of the information capacity in neuronal synapses," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5025–5036, Dec. 2016.

[12] K. Aghababaiyan, V. Shah-Mansouri, and B. Maham, "Capacity bounds of neuro-spike communication by exploiting temporal modulations," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.

[13] D. H. Johnson, "Information theory and neural information processing," *IEEE Trans. Inf. Theory*, vol. 56, no. 2, pp. 653–666, Feb. 2010.

[14] D. Malak and O. B. Akan, "A communication theoretical analysis of synaptic multiple-access channel in hippocampal-cortical neurons," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2457–2467, Jun. 2013.

[15] J. P. Meeks and S. Mennerick, "Action potential initiation and propagation in CA3 pyramidal axons," *J. Neurophysiol.*, vol. 97, no. 5, pp. 3460–3472, May 2007.

[16] M. Raastad and G. M. G. Shepherd, "Single-axon action potentials in the rat hippocampal cortex," *J. Physiol.*, vol. 548, no. 3, pp. 745–752, May 2003.

[17] K. Aghababaiyan and B. Maham, "Axonal transmission analysis in neuro-spike communication," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[18] B. Maham, "A communication theoretic analysis of synaptic channels under axonal noise," *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 1901–1904, Nov. 2015.

[19] K. Aghababaiyan and B. Maham, "Error probability analysis of neuro-spike communication channel," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jul. 2017, pp. 932–937.

[20] J. A. White, J. T. Rubinstein, and A. R. Kay, "Channel noise in neurons," *Trends Neurosci.*, vol. 23, no. 3, pp. 131–137, Mar. 2000.

[21] A. A. Faisal, J. A. White, and S. B. Laughlin, "Ion-channel noise places limits on the miniaturization of the brain's wiring," *Current Biol.*, vol. 15, no. 12, pp. 1143–1149, Jun. 2005.

[22] A. A. Faisal and S. B. Laughlin, "Stochastic simulations on the reliability of action potential propagation in thin axons," *PLoS Comput. Biol.*, vol. 3, no. 5, pp. e79–e89, May 2007.

[23] A. A. Faisal, "Stochastic simulation of neurons, axons and action potentials," in *Stochastic Methods in Neuroscience*. New York, NY, USA: Oxford Univ. Press, 2010, pp. 297–343.

[24] J. Tworzzydło, B. Trauzettel, M. Titov, A. Rycerz, and C. W. Beenakker, "Sub-Poissonian shot noise in graphene," *Phys. Rev. Lett.*, vol. 96, no. 24, p. 246802, Jun. 2006.

[25] L. P. O'Keefe and J. A. Movshon, "Processing of first- and second-order motion signals by neurons in area MT of the macaque monkey," *Vis. Neurosci.*, vol. 15, no. 2, pp. 305–317, Feb. 1998.

[26] J. F. C. Kingman, *Poisson Processes*. Hoboken, NJ, USA: Wiley, 1993.

[27] I. Rubin, "Regular point processes and their detection," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 5, pp. 547–557, Sep. 1972.

[28] A. Segall and T. Kailath, "The modeling of randomly modulated jump processes," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 135–143, Mar. 1975.

[29] J. L. Doob, *Stochastic Processes*. Hoboken, NJ, USA: Wiley, 1990.