A Queueing-Theoretical Delay Analysis Model for Intra-body Nervous Nanonetwork

Naveed A. Abbasi **Ozgur B. Akan**
Next-generation and Wireless Communications Laboratory,
Department of Electrical and Electronics Engineering,
Koc University, Istanbul, Turkey
Email: {nabbasi13, akan}@ku.edu.tr

***Abstract*—Nanonetworks is an emerging field of study where nanomachines communicate to work beyond their individual limited processing capabilities and perform complicated tasks. The human body is an example of a very large nanoscale communication network, where individual constituents communicate by means of molecular nanonetworks. Amongst the various intra-body networks, the nervous system forms the largest and the most complex network. In this paper, we introduce a queueing theory based delay analysis model for neuro-spike communication between two neurons. Using standard queueing model blocks such as servers, queues and fork-join networks, impulse reception and processing through the nervous system is modeled as arrival and service processes in queues. Simulations show that the response time characteristics of the model are comparable to those of the biological neurons.**

***Keywords*—Nanoscale communication, neuro-spike communication, intra-body nervous nanonetworks, queueing theory**

I. Introduction

Applications of nanotechnology are being realized from nano-switches and actuators [1], to intelligent drug delivery [2], nanoscale sensing [3] and bio-hybrid systems [4]. Although the promise of nanotechnology is huge, the associated challenges are not small by any means either. Nanomachines face very small dimensions, scarce processing, limited memory resources and simple networking capabilities.

The human body is a huge nanoscale communication network, where individual entities such as organs or cells communicate by means of nanomachines to make an intelligent system on a macro scale [5]. Understanding the dynamics of molecular communication not only helps us advance our work in development of nanomachines, but also gives a new perspective to the science of disease and treatment. Many diseases of the human body can in fact be quantified as various forms of communication system failures [6].

The nervous system forms one of the most complex communications systems. It is also one of the most studied systems because of its elegance and importance in the human body. Nervous diseases such as Alzheimers disease, Schizophrenia and Parkinsons disease are key challenges for the world in terms of human disease in the current age.

Although the applications and information in a particular communication network may differ significantly, the methods for communication remain similar. Several studies on molecular communication [5], [6], [7], [8] and [9] target the nervous nanonetwork to develop communication-theoretical understanding of the nervous system. These studies result in the formulation of various models of synaptic channels under different scenarios. On the other hand, [10], [11] and [12] introduce models for biological systems based on layered queueing networks.

The advantage of using queueing analysis for biological networks lies in the fact that big networks such as central nervous system and the neural cortex can be viewed on the whole as a single network, thus making us able to study the collective behavior of these networks. To date, no work exists on the modeling of nervous nanonetwork from the queueing theoretical perspective. The motivation for such a model lies in a variety of applications from drug delivery to nervous disorder diagnosis. For example, in case of drug delivery, the nervous system can be viewed as the wired network of the body and target-specific drug delivery can be triggered by means of nerve impulses and neurotransmitters. Parameters such as the time required for a particular drug delivery and optimal rates of impulses to produce these scenarios can be viewed as queueing parameters. Additionally, the diagnostics of nervous disorders, where reflex latencies help in diagnosis, can benefit from such a model. Practitioners can move from using basic reflexes such as patellar reflex to more complex responses to improve disease diagnostics.

The objective of this paper is to derive a model of nanoscale neuro-spike communications between one input and one output neuron by using the fundamentals of queueing theory. We first develop an understanding of the neuro-spike communication and identify key blocks of the system. We then model these blocks using queueing theory implements such as queues and servers. Finally, we perform an analysis of response time characteristics of the model.

The remainder of this paper is organized as follows. We provide a brief overview of neuro-spike communication in Section II. Section III presents the impulse transmission through the axon while Section IV discusses the neurotransmitter propagation and reception in the synapse. Based on these analyses, we develop a queueing model of the neuron in Section V. Results are presented in Section VI and concluding remarks are provided in Section VII.

II. Overview of Neuro-Spike Communication

The fundamental task of a neuron is to receive, conduct and transmit spikes or impulses which are generated in response to external or internal stimuli. These impulses travel between various body parts and the central nervous system (CNS).

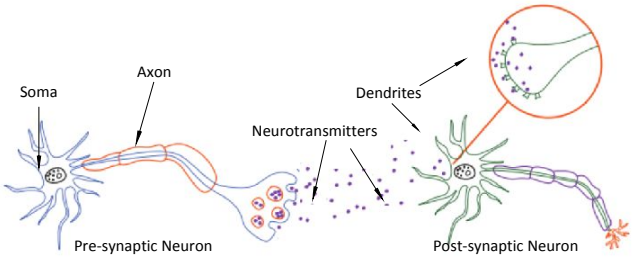

Fig. 1. Communication between a presynaptic and a postsynaptic neuron.

A single neuron can be divided into three main parts which are the dendrites, the soma or cell body and the axon. The communication between two neurons starts when an impulse traveling in a presynaptic neuron, also known as the action potential (AP), reaches the axon terminal. Fig. 1 shows two neurons in such a scenario. In order for the AP to traverse from one neuron to the next, it needs to travel across the cell gap known as the synaptic cleft. Synaptic communication can be done either electrically or chemically.

In electrical synaptic communication, APs are transferred directly between two neurons by means of direct physical connections between them. On the other hand, in chemical synaptic communication certain chemicals known as neurotransmitters are released in the synaptic space to achieve communication. Since, chemical synapses occur more frequently, we focus our study on these.

For chemical synaptic communication, when the AP reaches the axonal terminals of the presynaptic neuron, it causes membranous sacks called vesicles which are filled with neurotransmitters to move towards the cell membrane. Both the cell membrane and the vesicle membrane fuse at this stage enabling the vesicles to release the neurotransmitters into the synaptic space. The neurotransmitters drift across the synaptic space and bind to receptors present on the dendrites and the cell body of the postsynaptic neuron. The receptors are ligand-gated channels which open upon binding with a neurotransmitter. This causes positive Calcium ions (Ca^{2+}), which are higher in concentration outside the cell membrane, to flow inside the cell body causing a local depolarization of the cell membrane. Each of these local polarization causing channels can be thought to create a potential change known as Excitatory Post Synaptic Potential (EPSP). The overall depolarization is a sum of all the EPSP and is proportional to the size of the stimulus or the respective amount of neurotransmitters released.

This local depolarization is not powerful enough to traverse through the entire cell and in fact would need to be amplified along its way. Once it reaches the Axon Hillock and the depolarization is above a certain threshold (membrane threshold of usually around $-50mV$ from a base value of $-70mV$)[13], voltage gated channels Sodium ion (Na^+) and Potassium ion (K^+) channels are opened. This forms a positive feedback opening neighboring voltage gated channels further along the axonal body to open and causes further depolarization. The Na^+ channels become inactive after a while and that is the point when the depolarization peaks at around $70mV$. The K^+ channels then normalize the cell polarization afterwards. The AP travels through the axonal body until it reaches the axonal terminal again causing the same operations as explained above.

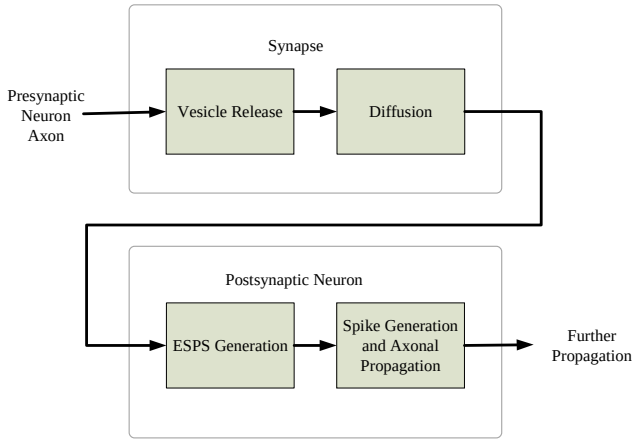

Fig. 2. Functional block diagram of neuro-spike communication.

A functional block diagram of neural communication is shown in Fig. 2. We see that the AP in a presynaptic neuron causes vesicle release in the synapse causing the diffusion of the neurotransmitter, EPSP generation at the postsynaptic neuron and spike propagation through the axon. All these processes can be loosely grouped under two types of transmissions namely, the axonal transmission and the synaptic transmission. Since our model is from a queueing perspective, we will be looking at these processes in terms of the service times required by these processes as well as the distributions of arrivals in these transmission networks.

III. AXONAL TRANSMISSION

In this section, we develop an understanding of axonal communication and identify the arrival and processing of APs through the axon. This analysis helps identify the queueing model for an axon. To simplify additional analysis, we will further identify dendritic transmission and somatic summation of the EPSP signals in this section as well.

Axons act as the transmission lines of the nervous system with diameters on the order of a few micrometers and are either unmyelinated or myelinated. Myelinated axons are covered by a sheath of a fatty dielectric substance called myelin. Since the axons carry electrical signals, insulation due to myelination creates a positive effect on the conduction enabling a rapid and better electrical propagation. Once the impulses reach the end of the axon, they terminate in the axonal terminal causing neurotransmitter release from vesicles. Although, the process is very complex, a few studies have identified the real-time latency of this release operation [13].

A. Arrival of Impulses

Axonal transmission starts with the arrival of impulses in the axon. Although the impulse sources for CNS are very diverse, most of these can be modeled as Poisson processes. In fact, several impulse sources for human nervous system are already known to be Poisson processes. These include the arrival rate of photons in human eyes, as well as sensations through the olfaction or gustation [7] and [14]. Therefore, the arrival of impulses in an axon can be modeled as a Poisson process.

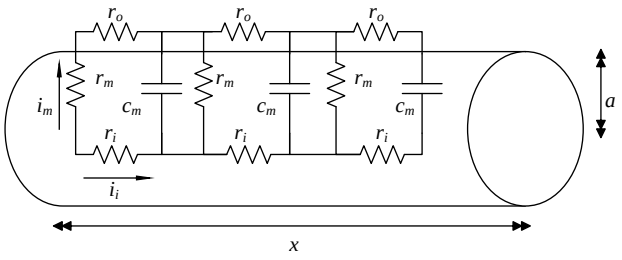

Fig. 3. Cable Theory Model of Axon.

B. Axon as a Cable

The best way to analyze the operation of an axon is to consider the operation as that of a transmission cable. Cable theory [15] provides us with such an axon model. Assuming that passive conduction occurs as the impulse is conducted through an axon, axons are modeled as cylinders composed of infinitesimal segments as shown in Fig. 3. Here c_m is the capacitance due to electrostatic forces across the axon membrane, r_m shows the membrane resistance per unit length, r_i is the axoplasmic resistance per unit length within the axon, r_o is the resistance per unit length outside the membrane and a is the radius of the axon.

The cable equation for the above model by [15] is

$$\lambda^2 \frac{\partial^2 V(x,t)}{\partial x^2} = V(x,t) + \tau \frac{\partial V(x,t)}{\partial t} \tag{1}$$

where $V(x,t)$ is the function for potential difference with respect to distance and time, λ is the length constant defined as $\lambda = \sqrt{r_m/r_i}$, and time constant τ is given as, $\tau = r_m c_m$.

The length constant indicates how far a charge can flow along the cable. Since λ is proportional to the square root of membrane resistance per unit length, r_m, the greater the values of r_m, the farther a charge can travel inside an axon. In myelinated axons membrane resistance is much higher so their length constants are higher as well. That is the reason why APs can travel more reliably for longer distance in much thinner myelinated neurons in comparison with unmyelinated neurons. [16] and [17] show that as the frequency of the input signal to a neuron increases, APs attenuate much faster because the length constant decrease that is occurring is proportional to increase in the internal resistance r_i. This shows that axonal transmission has a frequency dependence as well.

The time constant τ determines how fast the membrane potential responds to a current in the axoplasm. Thus, the larger the membrane capacitance, c_m, the longer it takes for a section of the membrane to get charged or discharged.

Suppose V_x is the voltage at a given distance x from one end of an axon, it is given as a solution for (1) by [15] as

$$V_x = V_0 e^{-x/\lambda} \tag{2}$$

where V_0 is the potential at the point of impulse initiation. Based on (2), we analyze and compute the time a signal takes to travel through an axon and the level of signal attenuation at a certain distance.

C. Conduction Speed of an Axon

One of the parameters most commonly found by experimental studies for various axons is the conduction speed through an axon. Lengths of various neurons are fairly easy to find experimentally. Thus, the time required for a transmission through that axon can be calculated based on the conduction speed through an axon.

Let us consider an AP traveling along an axon such that there it faces no attenuation. The conduction speed, s, is given as the derivative to distance traveled with respect to time.

$$s = \frac{dx}{dt} \tag{3}$$

If a voltage $V(x,t)$ of the AP is also moving in x direction, [18] shows that

$$\frac{d^2 V(x,t)}{dt^2} = \frac{s^2}{d} 4 \cdot \rho \cdot I_m \tag{4}$$

where d is the diameter of the elements from Fig. 3, $\rho = r_i \cdot pi \cdot a^2$ is the resistivity of the axoplasm and I_m is membrane current density given by $I_m = i_m/\pi \cdot d$. (4) can also be re-written as

$$s = \sqrt{\frac{d^2 V(x,t)}{dt^2} \frac{d}{4 \cdot \rho \cdot I_m}} \tag{5}$$

Thus, the conduction speed of an axon is directly proportional to the square-root of its diameter. This characteristic of the axonal communication is observed in most experiments based on axonal communication [13], [17].

In [19], the authors perform this analysis further and identify the conduction speed of unmyelinated axons as

$$s_{nm} = \sqrt{\frac{d}{8 \cdot \rho \cdot c^2 \cdot r^*}} \tag{6}$$

where c is the membrane capacitance per unit area and r^* is the resistance per unit area.

D. Na^+ Channel Inactivation

Before an AP is transferred, the axonal membrane is at rest and Na^+ channels are in a deactivated state. In response to an AP, these channels open, allowing the ions to flow into the axon and cause the action potential to grow. At the peak of an AP, when sufficient Na^+ ions enter the membrane, the Na^+ channels inactivate themselves. This stops further rise of an AP and slowly the membrane potential decreases back to its resting potential. When the membrane potential is low enough, the channels return to their deactivated state from the inactivated state.

Mentioning this phenomenon is important because for a given axon, this causes an additional delay before a new impulse can enter and propagate through the axon. In [17], the authors show that the during the inactivation after an AP, new impulses fail. They also show that the inactivation period increases at lower temperature, although, for the purpose of this work we are looking at neurons operating at a physiological temperature of 37^oC only. Hence, if the inter-arrival time

between two impulses should be larger than the inactivation period of the Na^+ ion channels.

IV. SYNAPTIC TRANSMISSION

In this section, we discuss synaptic transmission which starts with neurotransmitter release from the vesicles, diffusion through the synapse and the EPSP generation at the soma of the postsynaptic neuron. Neurotransmitter release can be identified by the arrival process in a queueing network whereas the diffusion and EPSP generation constitute the processing of the synaptic queueing network.

A. Neurotransmitter Arrival

Synaptic transmission starts by the arrival of neurotransmitters in the synapse. Neuro-spikes from the axon cause vesicles to move towards the cell membrane of the axon terminal and fuse with it. The neurotransmitters housed in the vesicles are then released in the synapse. Thus, synaptic neurotransmitter release can be thought of as the process by which the neurotransmitters arrive in the synaptic medium.

Neurotransmitter release is a random process in that even with the presence of neuro-spikes, vesicles might not be released or vice versa. The probability of such a scenario is quite low [20], therefore, if we ignore such cases, the release process depends directly on the arrival process of spikes.

B. Diffusion Through Synapse

The neurotransmitters released in the synapse reach the postsynaptic neuron by means of diffusion through the synaptic medium. The linear diffusion equation governs this process of diffusion [7] given as

$$\frac{\partial c(x,t)}{\partial t} = D\nabla^2 c(x,t) \tag{7}$$

where c denotes the concentration and D denotes the diffusion coefficient. Considering gaussian diffusion, the solution of the diffusion equation in an n-dimensional space from a position x is calculated by the theory of Greens function as

$$G(x,t) = (4\pi Dt)^{-n/2} exp\left[\frac{-(x-x')^2}{4Dt}\right] \tag{8}$$

where the Greens function tells us how a point of probability density initially at position x evolves over time and n-dimensional space. The above PDF exists under the condition of normalization stated as

$$\int G(x,t)dx = 1 \tag{9}$$

Assuming that diffusion for the current problem is a homogeneous process in one dimension, we modify (8) as

$$G(x,t) = \frac{1}{\sqrt{4\pi t}} exp\left[\frac{-(x-x')^2}{4t}\right] \tag{10}$$

(10) along with the condition provided in (9) describes the time distribution for diffusion of a neurotransmitter from one neuron to the next.

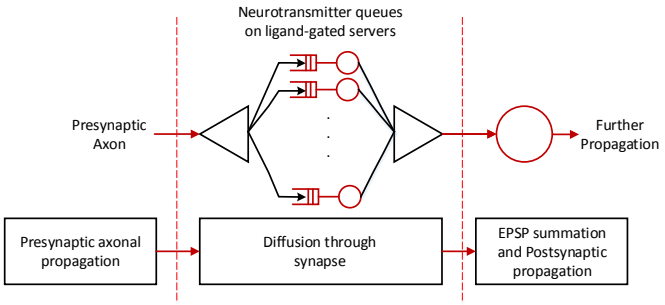

Fig. 4. Single-input single-output neuron queue model.

C. EPSP Generation

When neurotransmitters reach the ligand-gated channels present on the dendrites and soma of the postsynaptic neuron, they attach themselves to these channels and a flow of ions starts between the cell and its surroundings. As discussed earlier, the dendrites and axons have a similar structure and operation. The ion inactivation in dendrites is however, due to Ca^{2+} ions instead of the Na^+ ions. The flow of ions causes a change in membrane potential generating EPSP pulses. The analysis presented for axons in Section III is also valid for dendrites.

The EPSP generated at each of these ligand-gated channels is independent from EPSP generated from other channels. Subsequently, the cell body sums these independent EPSP to generate a potential difference. If this potential difference is above the threshold voltage, an AP is generated in the postsynaptic neuron, following a positive feedback amplification of the EPSP. Otherwise, the EPSP pulse dissipates without causing an AP in the postsynaptic neuron.

The number of neurotransmitters that successfully attach to the ligand-gated channels depends on the number of channels available at a time and the concentration of the neurotransmitters in the synapse. The number of neurotransmitters is usually very large as compared the channels so the excessive neurotransmitters either wait in the synapse till they are served, degraded or the presynaptic neuron re-uptakes them. This process is analogous to clients leaving a queue or packets being *dropped* from a queue after a certain time without service.

V. NEURON QUEUEING MODEL

Now that we have analyzed the key aspects of neural communication, we model it using a queueing perspective. To view the neuron from a queueing perspective, we need to break down the neuron model into queueing constituents namely, the servers and the buffers/queue.

A diagram for the system is provided in Fig. 4. We assume that there are two constituents of the system. The first constituent is composed of just a single server. This server models the propagation of neuro-spikes through the presynaptic neuron axon. The absence of a queue in the network relates to the fact that an axon can serve only one impulse at a time and does not contain any buffer where another incoming impulse can queue. Therefore, until an impulse passes through the entire system, we cannot have another impulse enter the axon. The spikes generated by the axon server then proceed to release the neurotransmitters in a synapse. The discipline of the axonal server is first come first serve (FCFS).

The second constituent of the system is a fork-join queueing network that deals with neurotransmitter release, synaptic diffusion, dendritic propagation and EPSP generation. The number of branches for this fork-join synaptic queue is be equal to the number of ligand-gated channels available for neurotransmitters attachment. To simplify the process, we assume that all the services of the synaptic communication are performed at the ligand-gated channels and each such receptor channel acts as a server. The synapse would thus act as the buffer of the system. The identification of the synapse with a queue is natural because when one neuron is trying to communicate with another, its respective neurotransmitters remain or buffer in the synapse while a previous impulse is being serviced by a neuron.

The summation of the EPSP in the soma may or may not result in further spike generation. Thus, depending on a threshold it might end up in a sink or might be transmitted through the axonal body. The probability of failure p_f to go beyond the threshold voltage is usually quiet low because neural communication is quite reliable as discussed in [13] and [17] but it is an essential parameter for any reliable model reliable model of communication [7]. In case there is a failure, further axonal propagation does not happen. Therefore, the second network is composed of single server queues with FCFS discipline queues in parallel.

A. Analysis of the Axonal Server

1) Arrival Process: Any sum of EPSP above the threshold level of the presynaptic neuron generates an AP in the axon. The arrival of this EPSP excitation to the presynaptic neuron is a stochastic process. Since there can be several independent sources that produce the excitation, the arrival process at the axonal server can be considered as a Poisson process.

2) Response Time: To determine the response time of an axon, we should first look at what constitutes the service of an axon. After an impulse enters an axon, as long as it is propagating within an axon, we consider it under service. Additionally, after an impulse leaves, the axon cannot have any further impulses pass through it until the Na^+ ion channels return from their inactivation state. Thus, the total response time of an axon is a summation of the Na^+ ion inactivation period and the response time for propagation through the axon body

$$R_A = R_{Na} + R_{AP} \tag{11}$$

where R_A is the average total response time of an axon, R_{Na} is the average time it takes for Na^+ ion channels to become reactivated and R_{AP} is the average response time for impulse propagation through the axon.

Considering that the axons of a particular neuron type have an average length of $E[L_A]$, and conduction speed s, the average response time for propagation through such axons, R_{AP}, is given as

$$R_{AP} = \frac{E[L_A]}{s} \tag{12}$$

Therefore, using (5), (6) and (12), we can provide expressions for response time of an axon for any general case and for an unmyelinated neuron case as

$$R_A = R_{Na} + \frac{E[L_A]}{\sqrt{\frac{d^2V(x,t)}{dt^2}\frac{d}{4\cdot\rho\cdot I_m}}} \tag{13}$$

$$R_{A-nm} = R_{Na} + \frac{E[L_A]}{\sqrt{\frac{d}{8\cdot\rho\cdot c^2\cdot r^*}}} \tag{14}$$

We see that the response time of the axonal server depends on the conduction speed, diameter, length, resistivity of the axoplasm, the Na^+ ion inactivation period and several other physiological factors for an axon.

B. Analysis of the Synaptic Transmission

1) Arrival Process: The arrival distribution of the synaptic distribution depends on the departure distribution of the axonal server. Several works on distributions of firing rate of neurons such as [21] and [22] have identified the distribution of firing neurons to be Poisson distribution. This is apparent from the analysis for axonal server given above as well. A Poisson input distribution to the axonal server, if deterministically served, generates an output which is also Poisson distributed.

2) Response Time: The service time of the fork-join synaptic queue network involves several processes such as time taken for diffusion through the synapse and the dendritic propagation. In such scenarios, we often assume network service time distributions to be some unknown general distribution. Therefore under the analysis presented, we can say that the fork-join synaptic queue is a network of parallel M/G/1 queues.

The total response time of synaptic transmission R_S is a summation of the average dendritic transmission time R_D and the average response time for diffusion through the synapse R_N.

$$R_S = R_D + R_N \tag{15}$$

Assuming the average dendritic length of $E[L_D]$ and a dendritic diameter of d_D, the average propagation time through a dendrite can be a modified form of (14) as the dendrites are also not coated by myelin [23].

$$R_D = R_{Ca} + \frac{E[L_D]}{\sqrt{\frac{d_D}{8\cdot\rho\cdot c^2\cdot r^*}}} \tag{16}$$

where R_{Ca} is the average Ca^{2+} ion channel inactivation time.

It is difficult to model fork-join queues analytically and in fact, to date, analytical results exist for only two server systems [24]. Usually for more than two queues, approximations for mean response time exist in case of homogeneous servers. This suits the current analysis because the servers we are considering here are homogeneous in nature.

We assume that the parallel queues of neurotransmitters in the synapse are independent and identically distributed (iid). In [25], the authors have presented an approximation of the mean response time of a set of iid fork-join queues with M/G/1 queues in parallel service as

$$R_N \approx R_1 + \sigma_1 F_N \alpha_N \tag{17}$$

where R_1 and σ_1 are the mean response time and standard deviation respectively for one M/G/1 queue with no fork-join properties. F_N is a constant which scales according to the service time distribution of the servers and α_N is a scaling factor that helps scale simulation results according to the results of a physical experiment.

If we consider a unit distance between the two neurons through the synapse, the PDF of neurotransmitter diffusion given in (10) can give us a service time distribution of neurotransmitter service for unit distance case. It should also be mentioned here that a normalization constant according to (9) also has to be multiplied so that the PDF does not exceed the unit area under the curve condition. Thus, the distribution of service time for a single M/G/1 neurotransmitter service is given as

$$G(t) = \frac{1}{\sqrt{4\pi t}} exp\left[\frac{-1}{4t}\right] exp\left[\frac{-t}{2.57 \cdot \pi}\right] \tag{18}$$

where the normalization constant $exp\left[\frac{-t}{2.57 \cdot \pi}\right]$ is approximated numerically to a precision of 10^{-6}.

Since the failure rate of neurotransmitters generating an AP from their individual EPSP is quiet low in case of a valid stimulation to the presynaptic neuron [17], there must be a high degree of synchronization in their diffusion though the synapse. In other words, this means that the mean response times of each of the fork-join queues is quite similar. Thereby, the standard deviation term from (17) can be neglected for successful transmission case resulting in

$$R_N \approx R_1 = \int tG(t)dt \tag{19}$$

C. Total Response Time

Both the axonal server and the fork-join synaptic queue are in series, therefore, the overall mean response time of a neuron to a single impulse is the summation of the mean response times of the axonal transmission and the synaptic transmission.

$$R = R_S + (1 - p_f)R_A \tag{20}$$

(20) can be written in its explicit form as

$$R = R_D + \int tG(t)dt + (1 - p_f)\left[R_{Na} + \frac{L}{\sqrt{\frac{d^2V(x,t)}{dt^2}\frac{d}{4 \cdot \rho \cdot I_m}}}\right] \tag{21}$$

This shows that the mean response time of a neuron depends on the diffusion characteristics, ion channel inactivation periods and the physical parameters of the neuronal structure. These characteristics match the results of previous neural communication models [6], [7], [8] and experimental results in [13], [14], [16], [17] and we see that the response time characteristics depend on propagation speed, axonal myelination and the diffusion distance through the synapse. Although our work is aimed at the Human nervous system, the results generated can be applied to any neuron type with provided characteristics.

An expression similar to (21) can be found specifically for unmyelinated neurons by using the value of R_{A-nm} instead of R_A in (20).

D. Impulses in the System

Apart from response time, other important measures of any queueing network are the number of customers in the system and its server utilization. These parameters become especially more useful when we are talking about large networks of queues. In our current scenario, customers in the system correspond to impulses present in a neuron.

We first consider the axonal server which is a single server system without any queue. Since there is no queue, at any given time, only one impulse can be served by the axon. Any other impulses that may come are rejected or dropped until the axon is ready to receive impulses again. The server utilization over a period of time, however, depends on the rate of impulse arrival at the axonal terminal. For any single server system, the arrival rate should be less than the service time for stable operation. If this rate is equal or beyond, the server utilization is hundred percent or in other words the neuron is maximally stimulated. The human body has several mechanisms in place to avoid such scenarios. However, tetanic contractions are one such example where a motor unit (muscle) is maximally stimulated by its associated neuron. This causes violent twitches in the muscles and can be lethal in certain cases. These contractions are usually the result of tetanus or the effect of toxic substances.

For the fork-join synaptic network, a queue exists in the synapse. To be stable, the arrival rate of the fork-join queue must be less than sum of the service rates of the servers. The number of neurotransmitters in the queue in such a case can be calculated by Littles law which states

$$N = \lambda_Q R_S \tag{22}$$

where N is the number of neurotransmitters in the queue, λ_Q is the arrival rate in the queue and R_S is the mean response time of the synaptic system. It must be noted here that the number of neurotransmitters in the synapse would be more than the number of neurotransmitters in the queue. This is because neurotransmitters stay in the queue until they are served, they degrade or the presynaptic neuron re-uptakes them.

VI. Simulation Results

A wide variety of neurons occurs naturally, with their own respective response types. Evolution over millions of years fine-tuned these various types of neurons to their specific tasks. For all this variability, the characteristics of most neurons are similar. Various characteristics of neurons such as propagation speed through axons, effects of myelination, failure rates of axonal communication, synaptic distances are known from several studies [13], [23], [26], [27], [28]. Some of the key results are compiled in I. Using these known parameters and assuming some of unknowns, we demonstrate that the characteristics of our current model are near a physical neuron by simulations conducted in MATLAB environment.

Authors of [35] identify the Na^+ ion absolute refractory period in human motor neurons to be nearly 2.65 ± 0.65 ms. It should, however, be noted that this refractory period might vary with stimulus strength [36], but since it is a small change, it can be neglected when the analysis of an entire system is

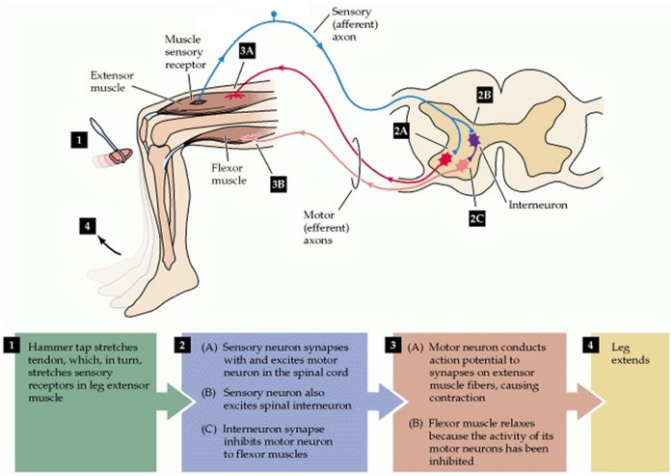

Fig. 5. The knee-jerk reflex arc [37].

being considered. Additionally, the time required for vesicle release is identified in real-time by the authors of [13] to be nearly 1.3 ms. We would be using the mean value of 2.65 ms for Na^+ ion absolute refractory period and 1.3 ms as the vesicle release latency in our analysis henceforth.

According to [24], the synaptic cleft between a presynaptic and a postsynaptic neuron is approximately 20 nm in humans. Furthermore, the probability of failure p_f of AP generation is taken at 3% according to the results of [17]. By (11), the time for diffusion of neurotransmitters with at least 97% confidence interval for a synaptic cleft width of 20 nm is derived to be nearly 0.7 ms.

Using these known parameters and assuming some of unknowns, we simulate a variety of scenarios in the next sections in MATLAB environment.

A. *Case study of a nervous circuit: the knee-jerk reflex arc*

1) Overview: One of the key advantages of using a queueing theory is the ease with which it can look at complex networks to find the average delays faced by packets. Under this motivation, we study a simple reflex action of the human body, known as the knee-jerk reflex, which is often used by physicians to test nervous diseases of the lower spinal cord. The knee-jerk reflex is an example of a monosynaptic nervous circuit in the body. This reflex is a reflex of proprioception which helps maintain posture and balance, allowing to keep ones balance with little effort or conscious thought. The process of eliciting a knee-jerk reflex is detailed in Fig. 5.

Striking the ligament of the knee with a reflex hammer just below the patella stretches the muscle spindle in the quadriceps muscle. Muscle spindles are sensory receptors present inside

TABLE I. AXON PARAMETERS [13][23][26][27] [28].

Type	Myelin	Diameter	Propagation Speed
Alpha Motor Neurons	Yes	$13-20\mu m$	$80-120m/s$
Gamma Motor Neurons	Yes	$5-8\mu m$	$4-24m/s$
Type I Sensory Neurons	Yes	$13-20\mu m$	$80-120m/s$
Type II Sensory Neurons	Yes	$6-12\mu m$	$33-75m/s$
Type III Sensory Neurons	Thin	$1-5\mu m$	$3-30m/s$
Type IV Sensory Neurons	No	$0.2-1.5\mu m$	$0.5-2.0m/s$
Preganglionic Neurons	Yes	$1-5\mu m$	$3-15m/s$
Postganglionic Neurons	No	$0.2-1.5\mu m$	$0.5-2.0m/s$

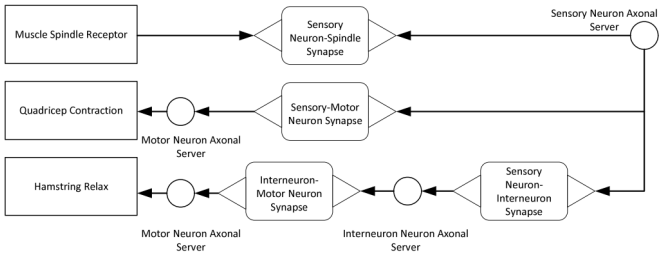

Fig. 6. Queueing model for knee-jerk reflex.

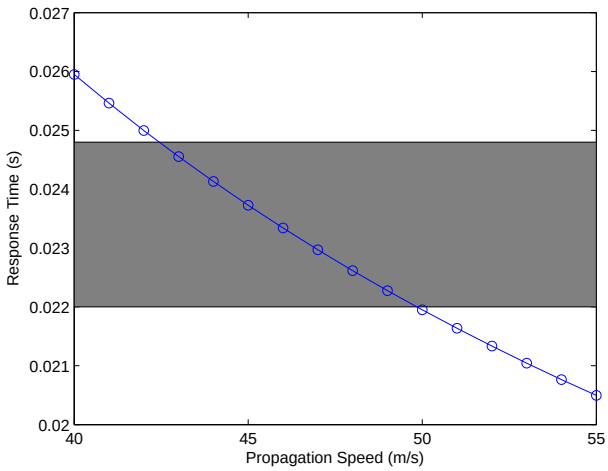

Fig. 7. Simulation of knee-jerk reflex over a range of propagation speeds.

a muscle which detect changes in the length of the muscle. These spindles then produce a signal which travels back to the spinal cord via a sensory neuron and synapses in the L4 spinal segment.

The sensory neuron synapses with two motor neurons. A femoral motor neuron then conducts the impulse back to the quadriceps femoris muscle, triggering its contraction. Through an interneuron, a tibial motor neuron carries the signal to the hamstring muscle to relax it. This contraction, coordinated with the relaxation of the hamstring muscle causes the leg to kick [37].

2) Experimental results: Since the knee-jerk response is a very simple and efficient way to detect nervous diseases, it has been the center of various scientific studies for more than a hundred years. Many studies are present where the latencies from the time of stimulation to the time of response have been conducted. In [38], the knee-jerk reflex latencies are found and further divided as reflex latency, which is the time taken by the nervous components of the reflex arc, and the reflex motor time which is the time required by the muscles to move after receiving signals from motor neuron.

3) Queueing model of knee-jerk circuit: Based on our discussion of the knee-jerk reflex, we present the queueing model for the circuit in Fig. 6. Application of an external impulse to the knee produces sensations in the muscle spindle. The muscle spindles synapse with a sensory neuron making the first fork-join synaptic queue. This queue feeds the resultant summed EPSP signals to the Sensory neuron server. The

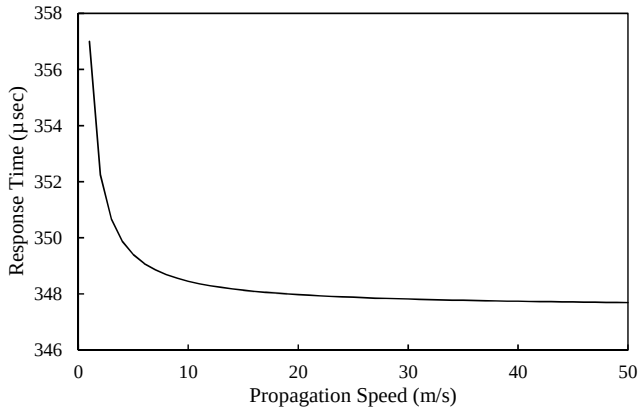

Fig. 8. Effect of propagation speed on response time.

sensory neuron is further connected with two fork-join queues, one for the motor neuron of qudricep and the other with an inhibitor interneuron. We can consider two independent fork-join queues for each of the nervous connections because the amount of neurotransmitter released is much higher than number of gates for either of the connections. The interneuron is further connected with the hamstring muscle though another fork-join queue. Each of these neurons has an axonal server as well.

4) Simulation of knee-jerk circuit: We now move on to simulate our knee-jerk reflex arc model. The average axonal lengths of the sensory and motor neurons for the knee-jerk reflex have not been estimated in any study, therefore, we have to approximate these. Since the axons of the sensory and the motor neurons run the length of the thigh to the base of the spinal cord near the hip, a good approximation would be to use the average distance between the hip and knee joint. According to US Center for Disease Control (CDC) statistics [39], the average upper leg length for males and females aged 2029 is approximately 40 cm. Using (21), we simulate for the entire reflex arc over the conductions speeds for the range of values provided in Table 1 and show the results in Fig. 7.

The shaded region of the graph represents the range of values for reflex latency as described by [38]. We see that a significant number of result data points fall within this range. This shows that application of the current model to the knee-jerk reflex nervous circuit gives similar results to those found by experimental studies. The results can be further improved if axonal lengths and the latencies are measured from the same population or studies such as [35] are conducted on more subjects.

B. Effect of Propagation Speed

Propagation speed is one of the key parameters found by most studies in neural operation. Usually, larger animals have higher propagation speed through the axon as the APs have to travel over a larger distance and give comparable response times as those of smaller animals [13]. Propagation speed also depends on the myelination of neurons, with myelinated neurons having higher speeds.

Assuming unmyelinated neurons from the human cortex and keeping a fixed response time across a unit length synapse of ($300\mu sec$) and the length of an axon as 10 microns, we generated Fig. 5. It shows that the as propagation speed increases, the response time decreases. In other words, we can say that as the signal travels longer distances over the same period of time because of a higher propagation speed. The result agrees with the vast majority of studies done on neurons [13][18][28] which show that higher propagation speeds result in quicker response from the neurons.

C. Effect of Axon Diameter

As discussed during the calculation of propagation speed, the diameter of an axon is directly proportional to its propagation speed. Taking the case of unmyelinated neuron, we varied the diameter of the axon to generate results shown Fig. 6. The values for capacitance per unit area, c, is taken as $1\mu F/cm^2$, the value resistance per unit area r^* is taken as $2000\Omega/cm^2$ and the resistivity is taken as $100\Omega m$ according to [23]. We see that as the diameter is increased, the propagation speed increases which in turn cause the response time to decrease. The response time characteristics are similar to those observed physical experiments in [13] and [23].

D. Voltage Decay along Myelinated and Unmyelinated Neurons

The Human nervous system contains both myelinated and unmyelinated neurons [14]. Myelinated neurons are typically found in sensory and motor neurons while non-myelinated neurons are found in the brain and spinal cord.

Voltage decay along the axonal length is plotted in Fig. 7 using (2). We assumed that the AP was at a maximum potential of $70mV$ at the start of two neurons, one myelinated and one unmyelinated. For a myelinated neuron, the value of membrane resistance r_m is about 10 to 2000 times that of an unmyelinated neuron [23] depending on the thickness of the sheath of myelin outside the axon. For our current result we used a value of 20 times. The resulting figure shows that over a similar length, the signal of the unmyelinated neuron decays more quickly. This was expected because as r_m increases, the length constant λ increases and the signal can travel longer distances. An interesting fact here is that as the membrane resistance increase for myelinated neurons, the membrane capacitance c_m decreases. This keeps the value of the time constant τ

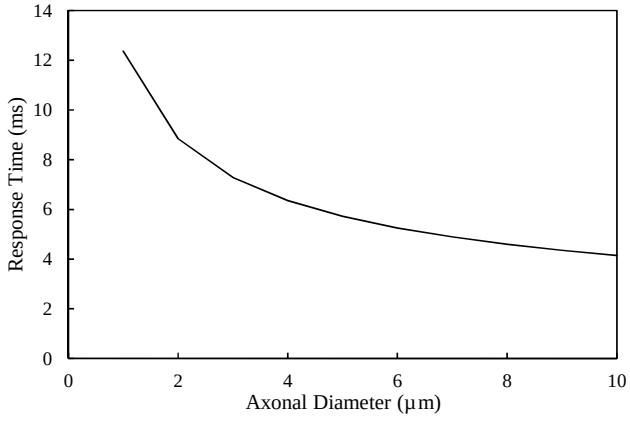

Fig. 9. Effect of axon diameter on response time.

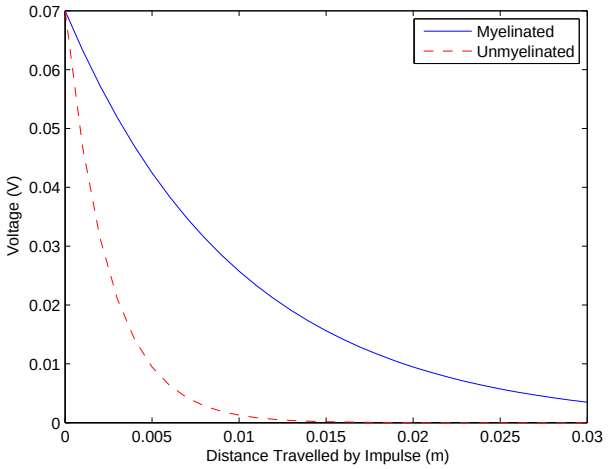

Fig. 10. Voltage decay along a neuron; myelinated vs unmyelinated neurons.

given by $\tau = c_m r_m$ nearly unchanged. The result agrees with experimental results that larger animals, needing slower AP decay, usually have myelinated neuron which operate at higher propagation speeds [13]. Similarly, sensory and motor neurons have myelination because they have to carry impulses over longer distances as compared to impulses carried over much smaller distances carried by the unmyelinated neurons in the human brain.

E. Effect of Frequency

The frequency of the input to a neuron is directly proportional to the internal resistance r_i and results in a decrease in the length constant. The results of Fig. 8 are generated using the PDF provided by (18) for a unit diffusion distance. Our results agree with those of [7] and [17] which state that as the frequency of the input signal to a neuron increases, the response time increases as well.

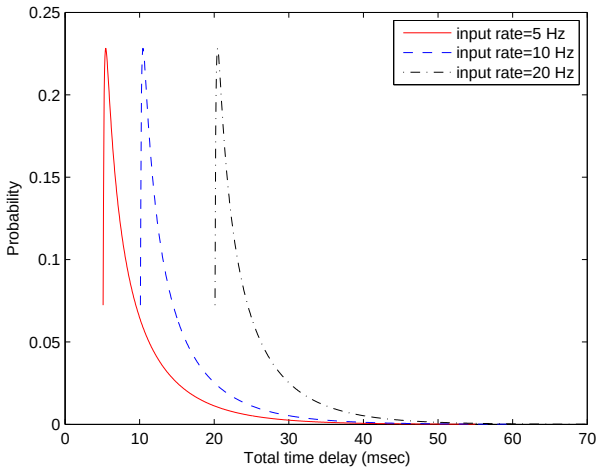

Fig. 11. Effect of frequency on response time distribution.

VII. Conclusion

In this paper, we characterized neuro-spike communication between a presynaptic and a postsynaptic neuron by providing a queueing theory based model of the neuron. The model was evaluated on the basis of its mean response time characteristics. We found that the response time of neurons depends on a host of features including their environment, type of neuron, their physical dimensions and the input signals they are being provided by their respective stimuli. We also motivated the use of our technique and applied it to knee-jerk reflex arc. Our results agree with experimental finding regarding the characteristics of neurons. The model is very flexible and can be applied to neurons from other animals as well.

The current model can be used to gauge the response time characteristics of bio-inspired networks for new nanomachines and it may also help build new benchmarks for the study of the nervous system. Our future works aim to develop queueing based network models for more complex neural circuits such as those involved in human memory to understand the network behavior of sensing.

VIII. Acknowledgment

This work was supported in part by the European Research Council (ERC) under grant ERC-2013-CoG #616922, by the Turkish Scientific and Technical Research Council under grant #109E257, by the Turkish National Academy of Sciences Distinguished Young Scientist Award Program (TUBA-GEBIP), and by IBM through IBM Faculty Award.

References

[1] M. Sauer, "Reversible molecular photoswitches: A key technology for nanoscience and fluorescence imaging," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9433–9434, 2005.

[2] S. Davis, "Biomedical applications of nanotechnologyimplications for drug targeting and gene therapy," *Trends in biotechnology*, vol. 15, no. 6, pp. 217–224, 1997.

[3] K. K. Jain, "Nanotechnology in clinical laboratory diagnostics," *Clinica Chimica Acta*, vol. 358, no. 1, pp. 37–54, 2005.

[4] I. F. Akyildiz, F. Brunetti, and C. Blzquez, "Nanonetworks: A new communication paradigm," *Computer Networks*, vol. 52, no. 12, pp. 2260–2279, 2008.

[5] D. Malak and O. B. Akan, "Molecular communication nanonetworks inside human body," *Nano Communication Networks*, vol. 3, no. 1, pp. 19–35, 2012.

[6] ——, "Communication theoretical understanding of intra-body nervous nanonetworks," *Communications Magazine, IEEE*, vol. 52, no. 4, pp. 129–135, 2014.

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

[8] D. Malak, M. Kocaoglu, and O. B. Akan, "Communication theoretic analysis of the synaptic channel for cortical neurons," *Nano Communication Networks*, vol. 4, no. 3, pp. 131–141, 2013.

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

[10] A. Sharp, A. Pannier, B. Wysocki, and T. Wysocki, "A novel telecommunications-based approach to hiv modeling and simulation," *Nano Communication Networks*, vol. 3, no. 2, pp. 129–137, 2012.

[11] B. J. Wysocki, T. M. Martin, T. A. Wysocki, and A. K. Pannier, "Modeling nonviral gene delivery as a macro-to-nano communication system," *Nano Communication Networks*, vol. 4, no. 1, pp. 14–22, 2013.

[12] T. M. Martin, B. J. Wysocki, J. P. Beyersdorf, T. A. Wysocki, and A. K. Pannier, "Integrating mitosis, toxicity, and transgene expression in a telecommunications packet-switched network model of lipoplex-mediated gene delivery," *Biotechnology and Bioengineering*, 2014.

[13] D. Debanne, E. Campanac, A. Bialowas, E. Carlier, and G. Alcaraz, "Axon physiology," *Physiological reviews*, vol. 91, no. 2, pp. 555–602, 2011.

[14] E. Goldstein, *Sensation and perception*. Cengage Learning, 2013.

[15] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, p. 500, 1952.

[16] R. Scott, A. Ruiz, C. Henneberger, D. M. Kullmann, and D. A. Rusakov, "Analog modulation of mossy fiber transmission is uncoupled from changes in presynaptic ca2+," *The Journal of Neuroscience*, vol. 28, no. 31, pp. 7765–7773, 2008.

[17] M. Raastad and G. M. Shepherd, "Single-axon action potentials in the rat hippocampal cortex," *The Journal of physiology*, vol. 548, no. 3, pp. 745–752, 2003.

[18] D. ATDLEY, *The physiology of excitable cells*. Soc Integ Comp Biol, 1971.

[19] I. Tasaki, "On the conduction velocity of nonmyelinated nerve fibers," *Journal of integrative neuroscience*, vol. 3, no. 02, pp. 115–124, 2004.

[20] A. Manwani and C. Koch, "Synaptic transmission: An information-theoretic perspective," *Advances in neural information processing systems*, pp. 201–207, 1998.

[21] W. R. Softky and C. Koch, "The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps," *The Journal of Neuroscience*, vol. 13, no. 1, pp. 334–350, 1993.

[22] A. Roxin, N. Brunel, D. Hansel, G. Mongillo, and C. van Vreeswijk, "On the distribution of firing rates in networks of cortical neurons," *The Journal of Neuroscience*, vol. 31, no. 45, pp. 16 217–16 226, 2011.

[23] J. G. Nicholls, A. R. Martin, B. G. Wallace, and P. A. Fuchs, *From neuron to brain*. Sinauer Associates Sunderland, MA, 2001, vol. 271.

[24] A. S. Lebrecht and W. J. Knottenbelt, "Response time approximations in fork-join queues," in *23rd UK Performance Engineering Workshop (UKPEW)*, Conference Proceedings.

[25] A. Thomasian and A. N. Tantawi, "Approximate solutions for m/g/1 fork/join synchronization," in *Proceedings of the 26th conference on Winter simulation*. Society for Computer Simulation International, Conference Proceedings, pp. 361–368.

[26] A. Siegel and H. N. Sapru, *Essential neuroscience*. Lippincott Williams & Wilkins, 2006.

[27] D. S. Stetson, J. W. Albers, B. A. Silverstein, and R. A. Wolfe, "Effects of age, sex, and anthropometric factors on nerve conduction measures," *Muscle & nerve*, vol. 15, no. 10, pp. 1095–1104, 1992.

[28] D. R. J. Laming, *The measurement of sensation*. Oxford University Press, 1997.