

# Neuroprostheses: method to evaluate the information content of stimulation strategies

Kevin Meng<sup>1</sup>, Hamish Meffin<sup>1,2</sup>, Michael R. Ibbotson<sup>1,2</sup>, Tatiana Kameneva<sup>3,4</sup>

**Abstract**—We propose a framework to evaluate the information content of different stimulation strategies used in neuroprosthetic implants. We analyze the responses of retinal ganglion cells to electrical stimulation using an information theory framework. This methodology allows us to calculate the information content by looking at the consistency of neural responses generated across multiple repetitions of the same stimulation protocol.

## I. INTRODUCTION

Retinal prosthetic systems work by translating an image of a visual scene into a pattern of electrical stimulation. One of the ways to encode visual information is to use electrode-coupled photodiodes mounted in the eye that convert light into current and inject it direct into the retina. An alternative method to stimulate the retina involves the acquisition of the images through an external camera. A specialized processing unit converts the recorded scenes into electrical instructions that are sent to the implanted stimulator chip. The camera-electrode system is versatile because visual scenes can be custom preprocessed prior to setting the electrical stimulation pattern, for example by enhancing edges or contrast.

When developing stimulation strategies for visual prosthetic systems, it is important to consider how neurons encode information. One of the ways to address this question is to use an information theory framework that has been applied in the context of visual neuroscience.

Information theory was developed in the 1940s to formalize signal-processing operations such as reliable communication, data compression and data storage [6]. The theory provides mathematical tools to determine the limits of communication systems. When applied in the context of neuroscience, the information theory quantifies how much information is carried by a neural response about the stimulus [1]. It is assumed that neural information is transmitted by the timing and frequency of action potentials [2]. Although action potential shapes can vary substantially, their widths and heights are assumed to have negligible variations for a given cell, such that the information is not contained in individual spikes, but rather in spike trains. Thus, the information theory framework considers only the interval between spikes, known as the inter spike interval. A neuron is assumed to be a transmission channel that encodes information. The neuron receives a stimulus and produces an observable response.

In mathematical terms, both the stimulus and response are random variables. The stimulus can be any parameter that modulates the neural response (i.e. intensity, direction or velocity of an image object). The response is any property of a neuron that is affected by the stimulus; it is often represented by the frequency of action potentials (i.e. firing rate).

Two important quantities constitute the basis of information theory. Entropy is a property of a random variable that represents the uncertainty in predicting its value. Mutual information is a property of the channel that represents transmission efficiency. Both quantities are expressed in units of bits. In this work, we calculate both quantities using data collected from rat retinal ganglion cells during electrical stimulation.

The aim of this paper is to provide a framework to compare the information content of different stimulation strategies using an information theory framework. The information content of the stimulation strategy indicates how consistent the neuron is at encoding the amplitude of the stimulus when a particular strategy is applied. We propose a methodology to compare the information content by looking at the consistency of neural responses generated across multiple repetitions of the same stimulation protocol. Our work brings together tools for retinal electrophysiology and information theory to evaluate stimulation strategies in retinal implants.

## II. METHODS

Experiments were conducted at the National Vision Research Institute. All protocols conformed to the policies of the National Health and Medical Research Council of Australia and were approved by the Animal Experimentation Ethics Committee of the University of Melbourne.

Data was collected from female Long-Evans rats whose age varied between 2 and 6 months. Each experimental session started with the extraction of retinal tissue. The retina was perfused at controlled room temperature with a constant flow of extracellular medium at 3-8 mL/min. The inner limiting membrane was scraped away to provide direct access to the underlying retinal ganglion cells (RGCs). Whole cell current-clamp recordings were obtained following standard procedures [3]. Initial pipette resistance ranged between 5 and 15 M $\Omega$ . The pipette was placed such that its tip slightly touched the membrane surface to form a gigaseal. The recorded voltage was amplified and digitally stored at an acquisition rate of 25 kHz.

<sup>1</sup> National Vision Research Institute, Australian College of Optometry, Australia. <sup>2</sup> Department of Optometry and Vision Sciences, University of Melbourne, Australia. <sup>3</sup> Faculty of Science, Engineering and Technology, Swinburne University, Australia. <sup>4</sup> Department of Biomedical Engineering, The University of Melbourne, Australia. Contact: Tatiana Kameneva, tkameneva@swin.edu.au

A disk-shaped stimulation electrode was placed in the extracellular medium prior to the patching process. The return electrode was placed at the other side of the chamber in the same bath. A custom-made MATLAB interface (MathWorks, R2016a) was used to command a multi-channel stimulator (Tucker Davis Technologies) and send different waveform signals for stimulation. The amplitudes of stimulation were bound by the limits of the stimulator at 300  $\mu\text{A}$ . Stimulation artifacts were blanked prior to spike detection.

The stimulation protocol involved the application of pulse trains. All pulses were biphasic, anodic-first, symmetric, and charge-balanced. No interphase gap was introduced. Pulse duration was set to 1 ms, pulse amplitude varied from 0 to 300  $\mu\text{A}$ , and pulse frequency was 10, 20 or 100 Hz.

To compare the amount of information content of stimulation strategies, we evaluated which strategy was the most consistent at encoding the pulse amplitude using the information theory methodology. The following two comparisons were made:

- Comparison 1: A strategy with regularly timed pulses versus a strategy with irregularly timed pulses. Both strategies had 100 pulses in 1 sec (100 Hz stimulation frequency);
- Comparison 2: A strategy with regularly timed pulses at 10 Hz versus a strategy with regularly timed pulses at 20 Hz. Both strategies had regularly timed pulses.

**Comparison 1.** The stimulation pulse train consisted of multiple blocks of pulses, each block being associated with a specific amplitude. The trains had a total duration of 60 seconds and consisted of 300 blocks. Each block lasted for 200 ms and contained 20 pulses of the same duration and amplitude. Pulse amplitude was kept constant within each block, but varied from one block to another in a randomized fashion. The range of pulse amplitudes was chosen as the range of the sigmoidal activation curve for each cell. Within this range, ten amplitudes were selected with linear step size and each amplitude was assigned to 30/300 blocks from the entire stimulation train.

Each session consisted of 5 pulse trains with regularly timed pulses and 5 others with irregularly timed pulses sent in a randomized order. A resting period of 10 seconds was applied between the trains. The regularly-timed strategy placed all pulses of the train at regular timed intervals of 100 ms (to achieve 100 Hz stimulation frequency). The irregularly-timed strategy placed 100 pulses in 1 sec at irregularly timed intervals derived from a Poisson distribution with a mean of 100 pulses (the mean stimulation frequency was 100 Hz). For a given strategy, each selected amplitude was assigned to a total of 150 blocks. Thus, according to the principles of information theory, 150 measures of the average firing rate were computed and formed the probability distribution for the corresponding amplitude.

**Comparison 2.** As above, the stimulation train consisted of multiple blocks of pulses, each block being associated with a specific amplitude. The trains had a total duration of 60 seconds and consisted of 60 blocks. Each block lasted for 1 second and contained either 10 pulses (10 Hz strategy)

or 20 pulses (20 Hz strategy). Pulse duration and pulse amplitude varied in the same way as in Comparison 1. The only difference was that only six amplitudes were selected from the range of activation sigmoidal curves. Therefore, each stimulation amplitude was assigned to 10 out of 60 blocks in the entire stimulation train.

Each session consisted of 10 pulse trains at 10 Hz and 10 pulse trains at 20 Hz placed in a randomized order. A resting period of 10 seconds was applied between the trains. For a given strategy, each selected amplitude was assigned to a total of 100 blocks and, therefore, 100 values of average firing rate formed the probability distribution for the corresponding amplitude.

In our work, the stimulus  $S$  is pulse amplitude and the response  $R$  is the average firing rate. The random variables  $S$  and  $R$  are appropriately discretized. Mutual information between stimulus and response provides a relative comparison different two stimulation strategies. For a given strategy, if the variability in the neuron's responses  $R$  is low, the mutual information  $I(S, R)$  is high, meaning that the strategy consistently encodes the amplitude of the stimulus,  $S$ . The mutual information is calculated according to the formula:

$$I(S, R) = \sum_{i,j} p(s_i, r_j) \log_2 \frac{p(s_i, r_j)}{p(s_i)p(r_j)}, \quad (1)$$

where  $p()$  is the probability distribution. Derivations and interpretations are given in the appendix.

### III. RESULTS

The voltage trace in Figure 1 shows typical behavior from a patched RGC following the presentation of biphasic pulses. In this particular example, five biphasic anodic-first pulses were injected. Artifacts occurred before action potentials. Spike detection was performed after blanking the artifacts and using a crossing amplitude threshold, usually at -20 mV. Most studies aim to control and optimize average quantities, such as the mean response of RGCs to a certain pattern of electrical stimulation. Very few studies look at the response variability as a useful measure to analyse stimulation strategies. In this section, we show how to apply information-theoretical principles to compare the amount of information content of different stimulation strategies.

**Comparison 1.** Ten amplitudes within the range of each neuron's activation were linearly selected. For both strategies (regular and irregular timed pulses), the associated distribution of average firing rates was estimated using 150 repetitions. Figure 2 illustrates the distributions of firing rates for 3 out of 10 selected amplitudes. The top row corresponds to the regularly-timed strategy; the bottom row corresponds to the irregularly-timed strategy.

Using the distributions of firing rates for all values of stimulation amplitudes, the following results for mutual information were computed using Equation (1): 0.71 bits for the regularly-timed strategy and 0.42 bits for the irregularly-timed strategy (see Table 2). This result should be interpreted as follows: under Strategy 1, the neuron is more consistent at encoding the stimulation amplitude; therefore, Strategy 1

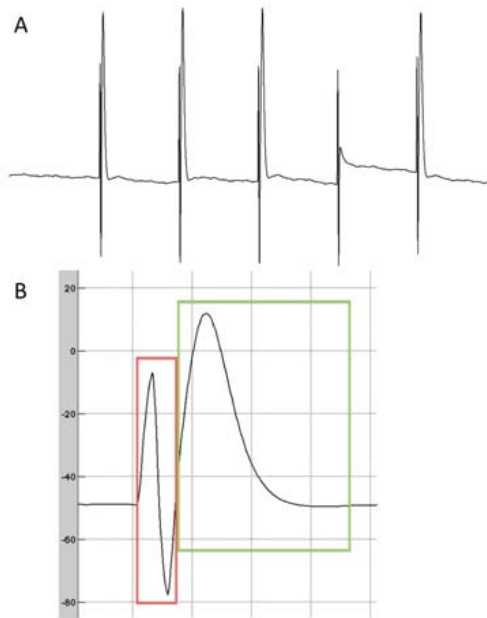


Fig. 1. Artifact blanking and spike detection. A. A typical voltage trace under continuous electrical stimulation. This example includes the injection of five biphasic anodic-first pulses. Artifacts are observed for each pulse. Four out of the five pulses elicited an action potential. The remaining pulse elicited a small depolarization, which did not reach spiking threshold. B. A detailed view of the voltage trace. The red rectangle highlights the artefact due to the anodic-first pulse, while the green rectangle frames the action potential. Artifacts never overlapped with action potentials and were blanked prior to spike detection.

(regular-timed pulses) should be preferred over Strategy 2 (irregular-timed pulses) in terms of response variability.

**Comparison 2.** Six amplitudes within the range of a neuron's activation were linearly selected. For both strategies (10 Hz and 20 Hz stimulation frequency), the associated distribution of average firing rates was estimated using 100 repetitions. Using the distributions of firing rates for all values of stimulation amplitudes, we found that the results were variable and dependent on the chosen discrimination

Table 2. Mutual information values [bits].

step.	Regular Pulses	Irregular Pulses	The calculated values of mutual information for 10 and 20 Hz stimulation were dependent on the discrimination step of the firing rates. One choice showed the lower frequency to be more informative (more consistent for encoding amplitudes within the dynamic range) than the higher one, whereas another choice led to the contrary. This illustrated that the choice of parameters has to be carefully considered. The probability distributions of firing rates for 10 and 20 Hz stimulation are illustrated in Figure 3.
$I(S,R)$	0.71	0.42	

#### IV. CONCLUSION AND DISCUSSION

We presented a new methodology to evaluate the information content of different stimulation strategies. The results provide a proof of principal example, which offer a framework to evaluate the amount of information content. More experiments with larger cell numbers should be conducted

to categorically conclude which strategy is more consistent at encoding the amplitude of the stimulus.

One of the practical issues associated with the application of tools to apply information theory to experimental neuroscience is extensive data collection. Often, cells cannot tolerate such prolonged stimulation protocols. In our experiments, we lost many cells before the protocol was completed. These cells were excluded from the analysis. Too little data may lead to a sampling bias problem. Entropy is defined in terms of probability distributions; therefore, there is a need for extensive data collection to approach the true distributions and provide accurate estimations of the entropy and mutual information. As a general rule, sampling makes neuronal responses appear less variable than they actually are, which means that estimated entropies are lower than their true values. More precisely, conditional entropies are biased towards lower values than the total entropy. Therefore, it appears that the mutual information is biased upward, which means that the neuron is assumed to be more informative than is actually the case [7]. Given the undesired implications of finite data collection, different computational studies have proposed methods to correct the sampling bias [4]. In our work, it is assumed that the number of repetitions is satisfactory and the probability distributions are accurate, such that no correction method to account for sampling bias is required.

Another important point to consider when applying information theory tools to neuroscience is discretization levels for both the stimulus and response variables. It has been shown that the number of repetitions per stimulus level should exceed the number of response variables by 2-4 times to reasonably estimate the true probability distributions [5]. The choice of quantization levels for stimuli and responses is not a trivial process and can lead to erroneous conclusions if it is badly designed. In our work, the following tradeoff has been considered: discretizing the range of stimulus using a small step to cover most of the possible amplitudes and collecting the largest amount of data per amplitude.

In our work, it is assumed that the total maximum information is the same for the stimulation strategies being compared. This is not proven at this time. Two theories have been proposed to explain how neurons encode information: rate code and temporal code. Both theories have been shown to agree with experimental results. Our work assumes that the visual encoding and decoding is based on the rate code (without assuming that the temporal code is wrong).

It is assumed that the evaluation of the information content of the stimulation strategies will be conducted *in vitro*. The best strategy will be applied in a visual prosthetic implant in the clinic. In order to apply the framework proposed here in the clinic directly, it is necessary for the implant to have the neural recording capacity, which is not a standard feature for the implants currently used developed.

#### REFERENCES

- [1] Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Nature neuroscience*, 2(11), 947-957.

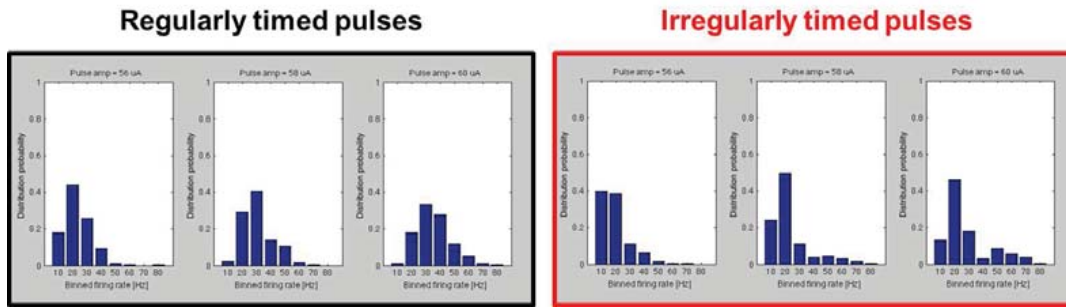


Fig. 2. Distribution of rates for 3/10 selected amplitudes. The top row corresponds to the regularly-timed strategy, while the bottom row corresponds to the irregularly-timed strategy. Despite the well-localized peaks of these distributions, which only constitute a sample of all distributions, the irregularly-timed strategy has an overall lower value of mutual information.

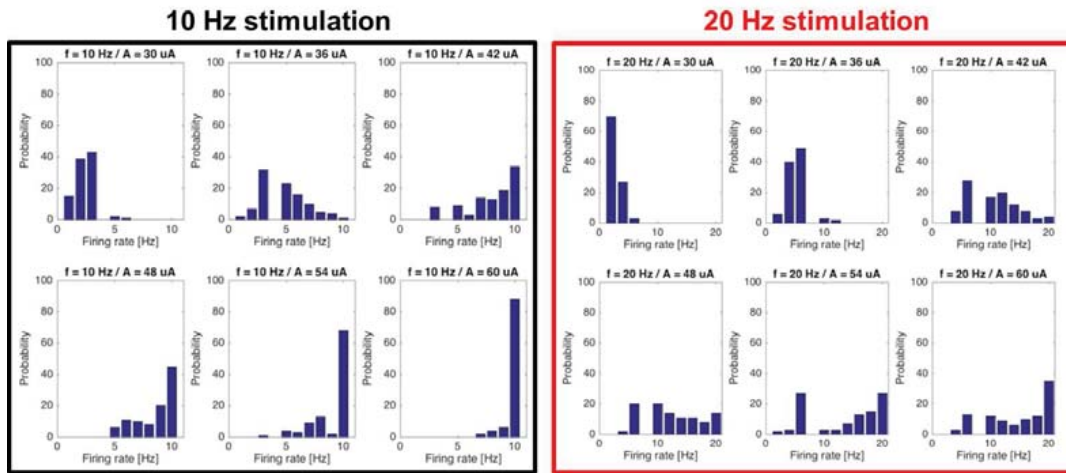


Fig. 3. Probability distributions of firing rates for 10 and 20 Hz stimulation for six different amplitudes of stimulation.

- [2] Dayan, P., & Abbott, L. F. (2001). Theoretical neuroscience (Vol. 10, pp. S0306-4522). Cambridge, MA: MIT Press.
- [3] Hamill, O. P., Marty, A., Neher, E., Sakmann, B., & Sigworth, F. J. (1981). Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Archiv*, 391(2), 85-100.
- [4] Panzeri, S., Senatore, R., Montemurro, M. A., & Petersen, R. S. (2007). Correcting for the sampling bias problem in spike train information measures. *Journal of neurophysiology*, 98(3), 1064-1072.
- [5] Pola, G., Schultz, S. R., Petersen, R. S., & Panzeri, S. (2003). A practical guide to information analysis of spike trains. In *Neuroscience Databases* (pp. 139-154). Springer US.
- [6] Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1), 3-55.
- [7] Treves, A., & Panzeri, S. (1995). The upward bias in measures of information derived from limited data samples. *Neural Computation*, 7(2), 399-407.

#### APPENDIX

The set of stimuli and the set of responses represented by discrete or discretised random variables  $S$  and  $R$ ,  $S = s_1, s_2, \dots, s_i, \dots, s_N$ ,  $R = r_1, r_2, \dots, r_j, \dots, r_M$ , with probability distribution  $p(S)$  and  $p(R)$ . The amount of information  $I(s_i)$  contained in the event  $s_i$  is related to the inverse of its probability of occurrence. The intuition behind this statement is that, the less probable the event, the more information it provides. This quantity is given by the logarithm to base 2 of the inverse of its probability and takes units of bits.  $I(s_i) =$

$\log_2\left(\frac{1}{p(s_i)}\right) = -\log_2(p(s_i))$ . Because the probability that the event  $s_i$  occurs lies in the interval  $[0,1]$ , the amount of information provided is always positive or zero.

The entropy  $H(S)$  can be seen as the uncertainty in predicting the stimulus  $S$ . It is defined as the expectation of the amount of information contained in each event  $s_i$ , over the probability distribution  $p(S)$ :  $H(S) = -\mathbb{E}_{p(S)}[\log_2 p(s_i)] = -\sum_i p(s_i) \log_2 p(s_i)$ . When response  $R$  is known, average conditional entropy  $H(S|R)$  is defined as the remaining uncertainty in predicting stimulus  $S$ . Therefore, it is expressed as a double expectation:  $H(S|R) = -\mathbb{E}_{p(R)}[\mathbb{E}_{p(S|r_i)}[\log_2 p(s_i|r_i)]] = -\sum_j p(r_j) \sum_i p(s_i|r_j) \log_2 p(s_i|r_j)$ . Mutual information  $I(S, R) = H(S) - H(S|R) = \sum_j p(r_j) \sum_i p(s_i|r_j) \log_2 \frac{p(s_i|r_j)}{p(s_i)}$ .

From Bayes' theorem, the following equality can be extracted for the joint probability distribution  $p(S, R)$ :  $p(S, R) = p(S|R) \cdot p(R) = p(R|S) \cdot p(S)$ . Therefore, mutual information can be rewritten as an expectation over the joint probability distribution:  $I(S, R) = \sum_{i,j} p(s_i, r_j) \log_2 \frac{p(s_i, r_j)}{p(s_i)p(r_j)} = \mathbb{E}_{p(S, R)}\left[\log_2 \frac{p(s_i, r_j)}{p(s_i)p(r_j)}\right]$ . Simple derivations starting from Bayes' theorem also show that mutual information is symmetric with respect to  $S$  and  $R$ :  $I(S, R) = H(S) - H(S|R) = H(R) - H(R|S)$ .