Vol. 49 (2018) *ACTA PHYSICA POLONICA B* No 12

THE KINEMATICS OF SPIKE TRAINS*

Daniel K. Wójcik

Nencki Institute of Experimental Biology of Polish Academy of Sciences
Pasteura 3, 02-093 Warszawa, Poland

(Received November 6, 2018)

Neural cells, the main agents in the brain responsible for processing of sensory information, animating our limbs, our thoughts, desires, and actions, communicate with each other by sending electric pulses called action potentials or spikes. This communication can be described with point processes which we introduce here simply.

DOI:10.5506/APhysPolB.49.2127

1. Introduction

When we look at a lecturer, we usually do not doubt his reality. We can see him, hear him, we are convinced that were we to come up and touch him he would be there. And yet, all this information perceived and expected is really a construct of our brain. Entering our brain, all sensory information is transformed into a distributed code of impulses called action potentials or spikes sent by the neural cells — neurons — to each other. This intricate network collectively processes incoming inputs constructing an image of the world, building memories, turning percepts and thoughts into actions.

Consider human eye. Each retina contains 125 million receptors which respond to incoming light by sending electric signals to other cells which process it through several layers of connected network leading to 1 million ganglion cells which send this processed information down to the brain. So the visual percept of the lecturer is built from images coming from two one-megapixel cameras updated a thousand times per second. Indeed, all the sensory stimuli are turned into sequences of identical impulses which are called spike trains. The interpretation of a specific sequence depends on the context, where is the signal coming from or where it goes.

The basics of neural coding were discovered at the beginning of 20th century by Adrian [1]. He recalled: *I had arranged electrodes on the optic nerve of a toad in connection with some experiments on the retina. The room*

---

[*] Presented at the LVIII Cracow School of Theoretical Physics “Neuroscience: Machine Learning Meets Fundamental Theory”, Zakopane, Poland, June 15–23, 2018.

was nearly dark and I was puzzled to hear repeated noises in the loudspeaker attached to the amplifier, noises indicating that a great deal of impulse activity was going on. It was not until I compared the noises with my own movements around the room that I realized I was in the field of vision of the toad's eye and that it was signaling what I was doing. Adrian observed three basic facts [1, 2]. He noticed that sensory neurons generate stereotypical impulses (action potentials) and the generation is all or nothing, pulse frequency encodes the stimulus amplitude, and prolonged stimuli lead to decrease of spiking frequency (adaptation).

Since then, numerous studies addressed the problem of coding in the neural system (see [3] for review). Recording responses of single cells to multiple repetitions of the same stimulus revealed their stochastic nature. It is most practical to think of the relation between possible stimuli $\{s(t)\}$ and responses $\{t_i\}$ in terms of joint probability distribution $P(\{s(t)\};\{t_i\})$. The problem of coding is the problem faced by a researcher: we present the same stimulus multiple times and we want to find out the distribution of admissible responses. The problem of decoding is the real problem the animal faces: given a specific spike train, identify the most probable stimulus which caused it. To discuss the codes used by the nervous system, we must have a language we can use to precisely describe this reality. This language is provided by the theory of point processes which we introduce here in a tutorial manner.

2. Spike trains as point processes

To set the stage consider the following experiment [4, 5]. An anesthetized cat is watching a screen where a light bar is moving with fixed velocity left and right with 1 s stops between changing the motion direction. We identify a cell in the visual system which responds to the stimulus by changing its activity in a way correlated with the stimulus manipulation. We want to characterize this response. How shall we do it?

A comprehensive representation of the results of such an experiment is shown in Fig. 1. How should we summarize these results? If we believe that every spike matters, then the whole raster plot is the only complete representation. At the other extreme, we might just provide a single number as a summary, such as the average spiking frequency. Indeed, this was the practice in the times of Lord Adrian. What if we are not happy with either extreme, how shall we proceed? After about 40 years since average frequency came into use, Gerstein proposed to summarize the results of evoked response experiments by averaging binned responses, which he called post-stimulus time histogram (PSTH) [6]. Forty years later, again, more complex models were proposed [7–9]. So should we expect progression of consecutively more

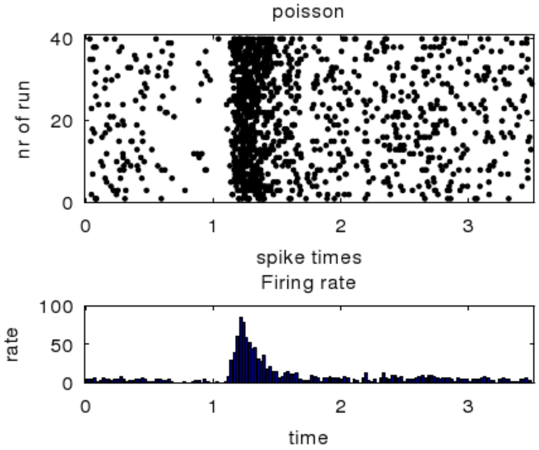

Fig. 1. Top: A raster plot representing results of a single experiment consisting of recordings of responses of a single cell to 40 repetitions of the same stimulus (simulated data). Bottom: Post-stimulus time histogram (PSTH), which is the average over repetitions of the response.

complex models? Are there ways to prefer one model over the other? These questions are addressed today within the theory of point processes, which we will now present.

Stochastic point process is a stochastic process whose realizations are sequences of point events. We assume these events are momentary and indistinguishable other than by the time of their appearance. For example, for spikes, we can take the time the membrane potential takes maximum value. Thus, a stochastic process is a rule which assigns a series of points to any random event from some probabilistic space.

We focus here on the case of evoked responses, when the same stimulus is presented repeatedly and the responses are collected. Then, during the time of experiment, we can register $0, 1, 2, \ldots, n$ spikes with probabilities $P_{(0,T]}[n]$ such that

$$\sum_{n=0}^{\infty} P_{(0,T]}[n] = 1\,.$$

The probability of each distribution is given by

$$Q_n(u_1, u_2, \ldots, u_n)\,, \qquad \text{where} \qquad u_1 < u_2 \ldots < u_n$$

and

$$\int_0^T \mathrm{d}\tau_1 \int_{\tau_1}^T \mathrm{d}\tau_2 \ldots \int_{\tau_{n-1}}^T \mathrm{d}\tau_n \, Q_n(\tau_1, \tau_2, \ldots, \tau_n) = P_{(0,T]}[n]\,,$$

where Q_n is probability density

$$Q_n(u_1, u_2, \ldots, u_n)\mathrm{d}u_1\mathrm{d}u_2\ldots\mathrm{d}u_n = \Pr(\text{one event in each interval } (u_1, u_1+\mathrm{d}u_1], (u_2, u_2+\mathrm{d}u_2], \ldots, (u_n, u_n+\mathrm{d}u_n])\,.$$

A stochastic process is defined uniquely by specifying all $Q_n(u_1, u_2, \ldots, u_n)$. If $Q_n(u_1, u_2, \ldots, u_n) = Q_n(u_1+t, u_2+t, \ldots, u_n+t)$, we say the process is stationary.

Note that we may also define $\tilde{Q}_n = \frac{1}{n!}$ for arbitrary $u_1, u_2, \ldots, u_n$

$$\int\limits_0^T \mathrm{d}\tau_1 \int\limits_0^T \mathrm{d}\tau_2 \ldots \int\limits_0^T \mathrm{d}\tau_n \tilde{Q}_n(\tau_1, \tau_2, \ldots, \tau_n) = P_{(0,T]}[n]\,.$$

3. Regular point processes

Among all possible point processes to describe spike trains, we consider a class of regular point processes. In such a process, the probability to observe an event during interval $(t, t+\mathrm{d}t]$ is

$$\begin{cases} \Pr[0 \text{ events in } (t, t+\mathrm{d}t]|N_{0:t}] = 1 - \lambda(t; H_t)\mathrm{d}t + o(t, \Delta t)\,, \\ \Pr[1 \text{ event in } (t, t+\mathrm{d}t]|N_{0:t}] = \lambda(t; H_t)\mathrm{d}t\,, \\ \Pr[\text{more than 1 event in } (t, t+\mathrm{d}t]|N_{0:t}] = o(t, \Delta t)\,, \end{cases} \qquad (1)$$

where

$$\begin{aligned} N_{0:T} &= \{0 < u_1 < u_2 < \ldots < u_j \le t \cap N(t) = j\}\,, \\ N(t) &\text{ — number of events during time}(0, T]\,, \\ H_t &= [u_1, \ldots, u_s] \text{ — vector of those } N(t) \text{ spike times}\,, \\ u_s &\equiv u_{N(t)}\,. \end{aligned}$$

From the definition above, we see that during short-time intervals at most one event may take place. $N_{[0,T]}$ is a single realization of the process during the time 0 to T. It is a right-continuous function which jumps by 1 during the spike times, u_k, and otherwise it is constant. Its properties are:

1. $N(t) \ge 0$;
2. $N(t)$ is integer;
3. if $s \le t$, then $N(s) \le N(t)$;
4. for $s < t$, $N(t) - N(s) =$ the number of spikes during $(s, t]$.

Quantity $\lambda(t|H_t)$ we introduced is called the *conditional intensity* for $t \in (0, T]$ (conditional intensity of the process, stochastic intensity, or hazard function)

$$\lambda(t|H_t) := \lim_{\Delta t \to 0} \frac{\Pr(N(t+\Delta t) - N(t) = 1|H_t)}{\Delta t}\,. \tag{2}$$

In general, λ may depend on the whole history of spiking.

In experiments, it is often convenient to use the *interspike interval (ISI) distribution* $p(t|H_t)$

$$p(t|H_t)\Delta t := \Pr(\text{one event in } (t, t+\Delta t) \cap \text{no events in } (u_{N(t)}, t]|u_1, \ldots, u_s)\,. \tag{3}$$

Another useful concept is *survival function*

$$\begin{aligned} S(t|u_1, \ldots, u_s) &:= \Pr(\text{no events in } (u_{N(t)}, t]|u_1, \ldots, u_s) \\ &= \int_t^\infty \mathrm{d}\tau\, p(\tau|u_1, u_2, \ldots, u_s) \\ &= 1 - \int_{u_s}^t \mathrm{d}\tau\, p(\tau|u_1, u_2, \ldots, u_s)\,. \end{aligned}$$

Hazard function, ISI distribution and the survival function are related as follows:

$$\begin{aligned} &p(t|u_1, \ldots, u_s)\Delta t = \\ &\Pr(\text{one event in } (t, t+\Delta t) \cap \text{no events in } (u_{N(t)}, t]|u_1, \ldots, u_s) = \\ &\Pr(\text{one event in } (t, t+\Delta t)|\text{no events in } (u_{N(t)}, t] \cap u_1, \ldots, u_s) \\ &\times\Pr(\text{no events in } (u_{N(t)}, t]|(u_1, \ldots, u_s)) = \\ &\lambda(t|u_1, \ldots, u_s)\Delta t\, S(t|u_1, \ldots, u_s)\,. \end{aligned}$$

Using this relation, it is possible to express each of them in terms of any other one. For example,

$$\lambda(t|u_1, \ldots, u_s) = \frac{p(t|u_1, \ldots, u_s)}{S(t|u_1, \ldots, u_s)} = \frac{p(t|u_1, \ldots, u_s)}{1 - \int_{u_s}^t \mathrm{d}\tau\, p(\tau|u_1, u_2, \ldots, u_s)}\,. \tag{4}$$

So we found a formula which allows us to compute the hazard function, λ, when we know the distribution of interspike intervals, p. Taking advantage of this formula, we can easily compute the remaining relations between λ, p and S

$$p(t|u_1, \ldots, u_s) = -\frac{\mathrm{d}}{\mathrm{d}t} S(t|u_1, \ldots, u_s)\,, \tag{5}$$

$$\begin{aligned}\lambda(t|u_1,\ldots,u_s) &= -\frac{\mathrm{d}S(t|u_1,\ldots,u_s)}{\mathrm{d}t}/S(t|u_1,\ldots,u_s)\\ &= -\frac{\mathrm{d}}{\mathrm{d}t}\ln S(t|u_1,\ldots,u_s)\,.\end{aligned}$$

Integrating this last formula from u_s to t, we obtain

$$\int\limits_{u_s}^{t}\mathrm{d}\tau\,\lambda(\tau|u_1,\ldots,u_s) = -\ln S(t|u_1,\ldots,u_s) + \ln S(u_s|u_1,\ldots,u_s)\,.$$

However, $S(u_s|u_1,\ldots,u_s)$ is the probability that during $(u_s,u_s]$, there was no spike, so it is 1. This follows, for example, from continuity

$$S(u_s+\tau|u_1,\ldots,u_s) = 1-\tau p(u_s+\vartheta\tau|u_1,\ldots,u_s) \geq 1-\tau c \overset{\tau\to 0}{\longrightarrow} 1\,,$$

where $\vartheta\in[0,1]$. Thus,

$$\int\limits_{u_s}^{t}\mathrm{d}\tau\,\lambda(\tau|u_1,\ldots,u_s) = -\ln S(t|u_1,\ldots,u_s)$$

and so

$$S(t|u_1,\ldots,u_s) = \exp\left[-\int\limits_{u_s}^{t}\mathrm{d}\tau\,\lambda(\tau|u_1,\ldots,u_s)\right]\,.$$

Therefore, the ISI distribution is given by

$$\begin{aligned}p(t|u_1,\ldots,u_s) &= -\frac{\mathrm{d}}{\mathrm{d}t}S(t|u_1,\ldots,u_s)\\ &= \lambda(t|u_1,\ldots,u_s)\exp\left[-\int\limits_{u_s}^{t}\mathrm{d}\tau\,\lambda(\tau|u_1,\ldots,u_s)\right]\,. \qquad (6)\end{aligned}$$

Of course, S is always a nonincreasing function, $p,\lambda\geq 0$.

A point process is called stationary when

$$\begin{aligned}\lambda &= \lambda(t-u_s)\,,\\ p &= p(t-u_s)\,,\\ S &= S(t-u_s)\,.\end{aligned}$$

4. Poisson process

Poisson process is the simplest point process where the conditional probability depends only on time: $\lambda(t|u_1,\ldots,u_s)\equiv\lambda(t)$. Further, if λ is constant, we have a homogeneous Poisson process. Due to its simplicity, majority of objects we discussed before can be computed which we shall do in this section as an illustration of the general theory.

Let us divide the time of experiment $(0,T]$ into M intervals δt, Fig. 2.

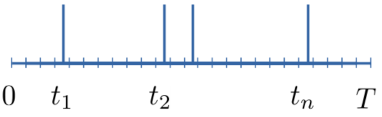

Fig. 2. Discrete representation of a spike train.

Then

$$\Pr[\text{there were spikes at times } u_1,u_2,\ldots,u_s] = \frac{\displaystyle\prod_{j=1}^{s}(\lambda(u_j)\delta t)\prod_{n=1}^{M}\left[1-\lambda\left(\left(n-\tfrac{1}{2}\right)\delta t\right)\delta t\right]}{\displaystyle\prod_{j=1}^{s}(1-\lambda(u_j)\delta t)}\,.$$

Thus, the probability density to observe specific spiking history $N_{0:t}$ is

$$\begin{aligned} p(N_{0:T}) &= \lim_{M\to\infty}\frac{\Pr[\text{there were spikes at times } u_1,u_2,\ldots,u_s]}{(\delta t)^s} \\ &= \prod_{j=1}^{s}\lambda(u_j)\lim_{M\to\infty}\prod_{n=1}^{M}\left[1-\lambda\left(\left(n-\tfrac{1}{2}\right)\delta t\right)\delta t\right]\,. \end{aligned}$$

Taking logarithm of the second factor, we obtain

$$\begin{aligned} \ln\prod_{n=1}^{M}\left[1-\lambda\left(\left(n-\tfrac{1}{2}\right)\delta t\right)\right] &= \sum_{n=1}^{M}\ln\left[1-\lambda\left(\left(n-\tfrac{1}{2}\right)\delta t\right)\delta t\right] \\ &\approx \sum_{n=1}^{M}\left[-\lambda\left(\left(n-\tfrac{1}{2}\right)\delta t\right)\delta t+o\left(\delta t^2\right)\right] \\ &\approx -\delta t\sum_{n=1}^{M}\lambda\left(\left(n-\tfrac{1}{2}\right)\delta t\right) \\ &\xrightarrow{M\to\infty} -\int_0^T\lambda(t)\mathrm{d}t\,. \end{aligned}$$

Thus, we obtain the fundamental formula

$$p(N_{0:T}) = \left(\prod_{j=1}^{s} \lambda(u_j)\right) \exp\left[-\int\limits_0^T \lambda(t)\mathrm{d}t\right] . \tag{7}$$

For homogeneous Poisson process of intensity $\lambda(t) \equiv \lambda$, this simplifies to

$$p(N_{0:T}) = \lambda^s e^{-\lambda T} .$$

Using the fundamental formula above, we can compute many properties of the Poisson process. For example, the probability to observe exactly n spikes during the time of the experiment $(0, T]$ is

$$\begin{aligned} P_{(0,T]}[n] &= \int\limits_0^T \mathrm{d}u_1 \int\limits_{u_1}^T \mathrm{d}u_2 \ldots \int\limits_{u_{n-1}}^T \mathrm{d}u_n \lambda^n e^{-\lambda T} \\ &= \lambda^n e^{-\lambda T} \int\limits_0^T \mathrm{d}u_1 \int\limits_{u_1}^T \mathrm{d}u_2 \ldots \int\limits_{u_{n-1}}^T \mathrm{d}u_n \\ &= \lambda^n e^{-\lambda T} \frac{T^n}{n!} . \end{aligned}$$

This is the Poisson distribution which gives name to the whole process

$$P_{(0,T]}[n] = \frac{(\lambda T)^n}{n!} e^{-\lambda T} .$$

For an inhomogeneous process, one can show [10]

$$P_{(0,T]}[n] = \frac{1}{n!} \left(\int\limits_0^T \lambda(t)\mathrm{d}t\right)^n \exp\left[-\int\limits_0^T \lambda(t)\mathrm{d}t\right] = \frac{(\bar{\lambda} T)^n}{n!} e^{-\bar{\lambda} T} .$$

A disadvantage of the Poisson process as a model for real spike trains is that it formally allows for arbitrarily close spikes. This cannot happen in real neurons because of their refractory properties, which are consequences of nonlinear properties of ion channel functioning. However, in many cases, for relatively low frequencies, this is not a serious limitation. On the other hand, Poisson process is very easy to estimate from data and so it is commonly used in analysis of experimental data.

5. More complex examples

The next level of complexity in point process theory is provided by renewal processes. They are defined by a fixed distribution of interspike intervals, $P(\tau)$. In this case, as we computed above in the general case, the hazard function is given by

$$\lambda(\tau) = \frac{P(\tau)}{1 - \int_0^\tau \mathrm{d}s P(s)}\,,$$

and the survival function by

$$S(\tau) = \int_\tau^\infty \mathrm{d}s\, P(s)\,.$$

Kass and Ventura [7] proposed a model which can be considered a generalization of renewal processes which they called inhomogeneous Markov interval model. They considered two-parameter hazard function, depending on the time from the start of the experiment, t, to address the need to describe response to the stimulus, and time since the last spike, τ, to take into account refractory properties of the cell

$$\lambda(t,\tau) = \lambda_1(t)\lambda_2(\tau)\,.$$

To estimate the model from data, one can use generalized additive models [7], but in experiments containing at least occasionally constant stimuli, a more direct method was proposed by Wójcik *et al.* [5]. Here, the distribution of interspike intervals in responses to stationary parts of stimuli is used to estimated $\lambda_2(\tau)$. This, together with experimental firing rate, is used to estimate $\lambda_1(t)$.

Figure 3 shows an example of artificial data simulated using Poisson, parametric IMI (where a gamma distribution was fitted to ISI distibution), and non-parametric IMI models, fitted to experimental data. A natural

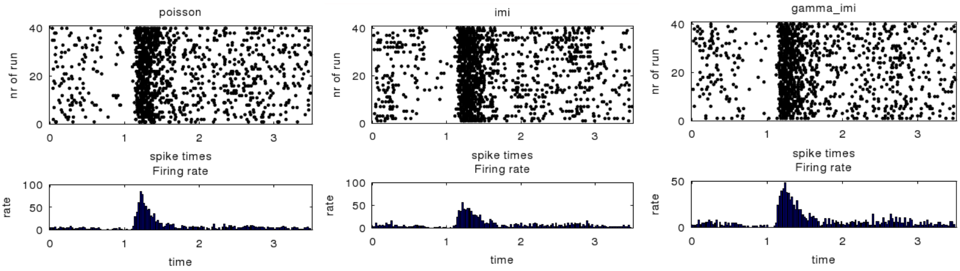

Fig. 3. Artificial spike train data modeling observed experimental activity.

question appears if we can somehow select between the different models. Should any of them be selected over others? To address this, we can use a goodness of fit test based on time-rescaling theorem.

6. Time-rescaling theorem

The time-rescaling theorem [11] says that if we know the true hazard function of the process, we can use it to rescale spike times so that they become a realization of the homogeneous Poisson process with unit rate. More formally:

Theorem *(Brown et al. 2002). Let* $0 < u_1 < u_2 < \ldots < u_n < T$ *be a realization of a point process with conditional intensity* $\lambda(t|N_t)$*. Define a transformation*

$$\Lambda(u_k) = \int_0^{u_k} \lambda(u|N_u)\,\mathrm{d}u\,,$$

for $k = 1, \ldots, n$*. Then* $\Lambda(u_k)$ *give a homogeneous Poisson process of unit rate.*

For proof see [11].

This theorem can be used as a basis for the following goodness-of-fit test:

1. Compute rescaled ISI: $\tau_k = \Lambda(u_k) - \Lambda(u_{k-1})$.
2. Transform τ_k to a new variable, $z_k = 1 - \exp(-\tau_k)$.
3. Then z_k are independent uniform variables on the interval.
4. Order z_k from smallest to largest and plot cumulative values of uniform density against the ordered z_ks.
5. If the model is correct, resulting curve will be diagonal.

Figure 4 illustrates this test on example data. We can see that the Poisson model strays from diagonal the most. In particular, it fails the most for the shortest ISIs. This is consistent with our expectations and the problems of addressing refraction which we mentioned before. Both IMI models fare much better, and we see that non-parametric IMI model is more consistent with the data than the gamma IMI model. This should be expected, as the spline-based non-parametric approach offers more flexibility in data modeling. However, it does require more data for fitting. For smaller data sets, parametric approach might be more reliable.

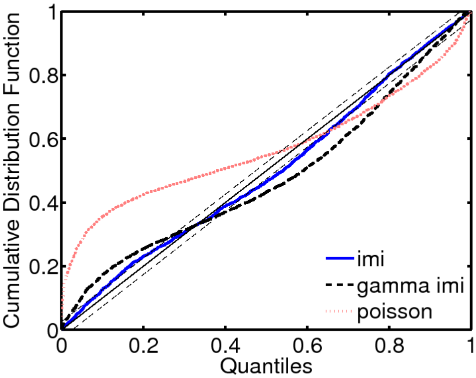

Fig. 4. Test of the model quality based on the time rescaling theorem.

7. Summary

In this tutorial, we presented the rudiments of the theory of the point processes as it is applied to the description of the spike trains in neuroscience. It is worth mentioning that like every mathematical theory its utility, even in neuroscience, is not limited to the description of spike trains. For example, point processes form a natural description of mice drinking behavior in IntelliCages, which was used, for example, by Kiryk *et al.* [12].

We have not been able to give justice to all aspects of this theory. In particular, simultaneous activity of multiple cells demands further extension of the theory [13]. The major challenges of the field today are consequences of availability of massive spiking data from thousands of channels [14, 15]. Reduction, processing and understanding of such data are open challenges in the field.

The author is grateful to the organizers of the LVIII Cracow School of Theoretical Physics in Zakopane, 2018, for invitation and for giving him opportunity to present this material.

REFERENCES

[1] E.D. Adrian, *Basis of Sensation*, Christophers, 1928.

[2] F. Rieke, D. Warland, R. de Ruyter van Steveninck, W. Bialek, *Spikes: Exploring the Neural Code*, MIT Press, 1999.

[3] R. Quian Quiroga, S. Panzeri (Eds.), *Principles of Neural Coding*, CRC Press, Taylor & Francis Group, 2013.

[4] G. Mochol *et al.*, *J. Neurosci.* **30**, 3199 (2010).

[5] D.K. Wójcik *et al.*, *Neural Comput.* **21**, 2105 (2009).
[6] G.L. Gerstein, *Science (New York)* **131**, 1811 (1960).
[7] R.E. Kass, V. Ventura, *Neural Comput.* **13**, 1713 (2001).
[8] R.E. Kass, V. Ventura, E.N. Brown, *J. Neurophysiol.* **94**, 8 (2005).
[9] W. Truccolo *et al.*, *J. Neurophysiol.* **93**, 1074 (2005).
[10] D.H. Johnson, *J. Comput. Neurosci.* **3**, 275 (1996).
[11] E.N. Brown *et al.*, *Neural Comput.* **14**, 325 (2002).
[12] A. Kiryk *et al.*, *Curr. Alzheimer Res.* **8**, 883 (2011).
[13] E.N. Brown, R.E. Kass, P.P. Mitra, *Nature Neurosci.* **7**, 456 (2004).
[14] L. Berdondini *et al.*, *Biosens. Bioelectron.* **21**, 167 (2005).
[15] J.J. Jun *et al.*, *Nature* **551**, 232 (2017).