A new EC–PC threshold estimation method for neural spike detection

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2012 J. Neural Eng. 9 046017

(http://iopscience.iop.org/1741-2552/9/4/046017)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 192.122.131.20
The article was downloaded on 03/08/2012 at 09:39

Please note that terms and conditions apply.
A new EC–PC threshold estimation method for in vivo neural spike detection

Zhi Yang\(^1,5\), Wentai Liu\(^2\), Mohammad Reza Keshtkaran\(^1\), Yin Zhou\(^1\), Jian Xu\(^1\), Victor Pikov\(^3\), Cuntai Guan\(^4\) and Yong Lian\(^1\)

\(^1\) Department of Electrical and Computer Engineering, National University of Singapore, Singapore 119077, Singapore
\(^2\) Department of Bioengineering and California Nanosystems Institute, UCLA, Los Angeles, CA 90095-1600, USA
\(^3\) Huntington Medical Research Institutes, Pasadena, CA 91105, USA
\(^4\) Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore 138632, Singapore

E-mail: eleyangz@nus.edu.sg

Received 14 December 2011
Accepted for publication 8 June 2012
Published 13 July 2012
Online at stacks.iop.org/JNE/9/046017

Abstract

This paper models in vivo neural signals and noise for extracellular spike detection. Although the recorded data approximately follow Gaussian distribution, they clearly deviate from white Gaussian noise due to neuronal synchronization and sparse distribution of spike energy. Our study predicts the coexistence of two components embedded in neural data dynamics, one in the exponential form (noise) and the other in the power form (neural spikes). The prediction of the two components has been confirmed in experiments of in vivo sequences recorded from the hippocampus, cortex surface, and spinal cord; both acute and long-term recordings; and sleep and awake states. These two components are further used as references for threshold estimation. Different from the conventional wisdom of setting a threshold at \(3 \times \text{RMS}\), the estimated threshold exhibits a significant variation. When our algorithm was tested on synthesized sequences with a different signal to noise ratio and on/off firing dynamics, inferred threshold statistics track the benchmarks well. We envision that this work may be applied to a wide range of experiments as a front-end data analysis tool.

\(^{5}\) Author to whom any correspondence should be addressed.

1. Introduction

Understanding a neural code responsible for information representation, generation and transmission is one of the important problems in neuroscience toward demystifying the neuronal substrate of intelligence, cognitive function and memory [1–4]. Intensive neurophysiological experiments [5–7], neuroprothetic studies [8, 9] and computational modeling effort [10, 11] have been reported to infer a potential neural code, and code candidates include ‘firing-rate’, ‘synchrony propagation’, or a combination of the two [8, 12–15]. For all the candidates, they take spiking activities from neural ensembles as input. In this sense, reliably detecting spikes from neurophysiological recordings is a prerequisite for information decoding. In general, the detected spike rate is very sensitive to detection threshold, implying a great deal of uncertainty in harvested information regardless of the choice of neural code [16]. After a few decades of electrophysiology experiments and algorithm development [11, 17–20, 20–29], experimentalists and theoreticians still raise the question in various occasions: how do you set the detection threshold? A trade-off is that if one lowers the detection threshold too much, a large portion of threshold crossing activities become attributed to noise; raising the threshold, on the other hand, gets rid of these false detections at a cost of nonlinearly reducing information content. This is further complicated by several imposed constraints. For example,
In this work, we propose a new approach to choose detection threshold. The framework is compatible with many detection algorithms, e.g., spike magnitude, nonlinear energy operator [18, 30], instantaneous energy and waveform slope (derivative). Our focus is to make the algorithm of practical use in different experiment setups and robust to recording imperfections such as waveform variation, spike overlapping, recording artifact and interference, and microglia cells that may severely degenerate the signal to noise ratio (SNR).

We first discuss the main sources of electronic noise. In this section, we show that electronic noise is a significant source of noise in neural recording. Figure 1 shows a lumped circuit model of the electronic noise. The electronic noise is given by

\[ I_{\text{electronic noise}} = 4kT / R_b \]

where \( k \) is Boltzmann constant, \( T \) is temperature, and \( R_b \) is the input referred electronics thermal noise by transistors, switches and resistors. \( I_{\text{electronic noise}} \) is the input referred electronics flicker noise produced by transistors [36].

Detection errors are not accessible in most experiments, making the topic unsupervised and lacking a benchmark for objective evaluation; neurons and neural networks exhibit nonlinearity and non-stationarity; some spike detection methods rely on waveform shapes [22, 28, 29] and require fine-tuned templates that are difficult to obtain especially in the presence of waveform variation, spike overlapping, recording artifact and interference, and microglia cells that may severely degenerate the signal to noise ratio (SNR).

In this work, we propose a new approach to choose detection threshold. The framework is compatible with many detection algorithms, e.g., spike magnitude, nonlinear energy operator [18, 30], instantaneous energy and waveform slope (derivative). Our focus is to make the algorithm of practical use in different experiment setups and robust to recording imperfections such as waveform variation, spike overlapping, recording artifact and interference, and low SNR. The rest of this paper is organized as follows. Section 2 focuses on neural interface noise characterization. Section 3 presents algorithms for spike detection. Section 4 presents experimental results. Section 5 gives concluding remarks.

2. In vivo neural interface noise characterization

Recorded neural spikes are superimposed with neural interface noise that exhibits non-stationary and non-white-Gaussian characteristics. It can be approximately fitted as \( 1/f \) noise, where \( f \) is frequency and \( x \) is a positive number less than 3. The frequency dependence is contributed by multiple sources including neuron noise [31–34], electrode–electrolyte interface noise [35], tissue thermal noise and electronic noise [36], which are illustrated in figure 1 using a lumped circuit model. Except the tissue thermal noise \( 4kTR_b \) in figure 1 that has a flattened spectrum, the rest decrease with frequency [37].

To verify the neural interface noise contribution from different sources, an in vivo experiment is performed that uses two sharp tungsten electrodes separated by 125 \( \mu \)m to record the hippocampus neuronal activities of a rat. One of the electrodes is coated with carbon nanotube (CNT), while the other is uncoated. After the electrodes have been placed, a euthanizing drug is injected. After 5 s of drug injection, the recording of the two electrodes start and last until the time of death. The noise analysis results are summarized and presented in figure 2. In figure 2(a), a 5 min segment that captures the decaying of background activities is plotted. In figure 2(b), the estimated neural interface noise from 600 Hz to 6 KHz for both recording sites are plotted, where noise dramatically reduces (>80%) after the drug takes effect. Initially, the CNT electrode records a comparatively larger noise (697 \( \mu \)V^2) compared with the uncoated electrode (610 \( \mu \)V^2). After a few minutes, the background noise recorded by the CNT electrode quickly reduces eventually reaching 37 \( \mu \)V^2 that is about 1/3 of noise recorded by its counterpart (112 \( \mu \)V^2), suggesting that the noise floor of using the uncoated tungsten electrode (112 \( \mu \)V^2) is set by the electrode. From these two plots, we can estimate that the neuron noise is around 500–600 \( \mu \)V^2, electrode interface noise is ~80 \( \mu \)V, while the sum of electronic noise and electrolyte bulk noise is less than 37 \( \mu \)V^2 (only ~5% of the neural interface noise). Figure 2(c) displays the 1/f noise spectrum recorded from the uncoated tungsten electrode \( (x = 1.8, 1.4, 1.0, 0.9, \text{estimated at } 0, 15, 30, 45 \text{ min after drug injection}) \). Figure 2(d) displays 1/f noise spectrum recorded from the CNT-coated electrode \( (x = 2.1, 1.3, 0.9, 0.8, \text{estimated at } 0, 15, 30, 45 \text{ min after drug injection}) \).

Similar experiments have been repeated several times and the results are consistent. Summarized from the experiments, the dominant noise source (>80% reduction) is neuron noise followed by electrode noise. As a follow up to our previous work to characterize contributions from different noise sources [37], in the next section, we study the in vivo neural interface noise (the total noise) and its deviation from Gaussian. Since both electrode noise (regardless of the thermal noise model or shot noise model [35]) and circuit noise [36] are Gaussian and they are not the dominant noise sources, without loss of generality, we sometimes use neuron noise to replace neural interface noise (the summation of neuron noise, electrode noise and circuit noise) in the rest of the paper.

2.1. Neuron noise and central limit theorem

In a generic form, recorded data \( V(t) \) are superimposed from spikes, field potentials and neuron noise:

\[
V(t) = \sum_{i=1,2,...,J} V_{1,i}(t) + \sum_{j=1,2,...,J} V_{2,j}(t) + V_{\text{syn}}(t) + N_0(t),
\]

where \( V_{1,i}(t) \) are the activities of a neuron within the recording radius \( r_1 \) (spike power is much larger than noise power), \( V_{2,j}(t) \) are the activities of a neuron in an extended radius \( r_2 \) (spike power is smaller than or comparable to noise power), \( V_{\text{syn}}(t) \) is the field potential and \( N_0(t) \) is the neural interface noise.

To study neural interface noise and its deviation from Gaussian, histograms of recorded data are shown in figure 3. For broadband data of 1 Hz–8 KHz, the histograms exhibit significant fluctuations as plotted in figure 3(a). The histogram cannot be smoothed by increasing the data length, suggesting data non-stationarity. After applying a high-pass filter at 300 Hz to remove low-frequency activities, data histograms become smoothed and can be approximately fitted by a...
Figure 2. In vivo recording for identifying noise sources. (a) Recording segment of 5 min capturing the decay of background activities. (b) Traces of the estimated neural interface noise versus time are plotted. Black ■ curve represents the noise recorded from a custom tungsten electrode; red ▲ curve represents the noise recorded from a CNT-coated electrodes with the same size. (c), (d) Noise power spectrums estimated at the 0, 15, 30, 45 min after the drug injection. In (c), a conventional tungsten electrode is used. A CNT-coated tungsten electrode of equal size is used for comparison (reproduced from [37]) in (d).

Figure 3. (a) A histogram of broadband neural data at 1 Hz–8 kHz. (b) A histogram of bandpass filtered neural data at 300 Hz–8 kHz. (c) Re-plot of (b) in the log scale.

Gaussian distribution as shown in figure 3(b). The fitting curve excellently matches the histogram at its central region and deviates at regions beyond the data standard deviation (SD) as re-plotted in figure 3(c) in the log scale.

To quantitatively study high-frequency neuron noise, its Gaussian signature and deviations from Gaussian, we refer to the central limit theorem (CLT) [38], which states that the sum of a sufficiently large number of independent random variables, each with identical distribution, finite mean and variance, follows a Gaussian distribution. The conditions of ‘independent variables’ and ‘identical distribution’ are strong and more relaxed conditions have been worked out in the literature, e.g., Lyapunov’s condition [39] that allows variables of ‘different distributions’.

Definition. Let \( X_i \), \( i = 1, 2, 3, \ldots, K \), be a sequence of independent random variables. Suppose that each \( X_i \) has a finite expected value \( E[X_i] = \mu_i \) and a finite variance \( E[(X_i - \mu_i)^2] = \sigma_i^2 \). If for some \( \delta > 0 \), the expected values \( E[|X_i|^{2+2\delta}] \) are finite and for every \( 1 \leq i \leq K \),

\[
\lim_{K \to \infty} \frac{1}{(\sum_{i=1}^{K} \sigma_i^2)^{1+\delta}} \sum_{i=1}^{K} E[|X_i - \mu_i|^{2+2\delta}] = 0 \tag{2}
\]

is satisfied, then the convergence to Gaussian holds: \( \sum_{i=1}^{K} X_i \) follows a Gaussian distribution with the mean \( \sum_{i=1}^{K} \mu_i \) and variance \( \sum_{i=1}^{K} \sigma_i^2 \).
If the voltage fluctuation induced by each neuron is treated as a random variable ($X_i$), then the recorded voltage $V$ is the sum of random variables $V = \sum X_i$ under a framework similar to the CLT. However, there are two issues regarding the application of Lyapunov’s condition here. First, mesoscopic-level synchronization positively correlates with the cortical volume enclosed [3], producing long-lasting firing rhythms against the CLT, which explain the deviations from Gaussian in figure 3(a). The synchronized activities appear at less than 200 Hz and can be removed by high-pass filtering. In other words, only a moderate number of neurons located in a proximity region (e.g. a few hundred $\mu$m) of the recording site could be treated as the pool for Lyapunov’s condition, making the large number hypothesis questionable; it is further added that the convergence to Gaussian is slow: assuming $N$ the number of neurons and a homogeneous neuron density function, we derive that equation (2) converges at a rate of $O(N^{1/3})$. Second, the strength of pairwise synchronization is inversely proportional to the distance between the neurons [1], implying that the random variables to be summed over are not completely independent. A relaxation on the requirement of independence is possible for the CLT [40, 41]; however, it still lacks a sufficient and necessary condition. Alternatively, we have analyzed sequences collected from different animal preparations to investigate the validity of CLT. As an example shown in figure 4, after being high-pass filtered at 300 Hz, typical autocorrelation curves of neural interface noise have their first zero-crossing point around a few hundred $\mu$s. The waveform ripples decay very fast initially and approach a slowly decaying state with the variance $10^{-6} - 10^{-4}$, suggesting a weak but a long-lasting correlation among samples. The data histograms, on the other hand, exhibit small deviations from Gaussian. In section 3, we reason that the small deviations are caused by sparse distribution of spike energy: neuronal synchronization breaks the hypothesis of independence (field potentials, figure 3(a)), while a number of nearby, high firing-rate neurons (neural spikes, figures 3(b) and (c)) violate Lyapunov’s conditions.

3. Unsupervised near-optimal neural spike detection

In this section, we propose a framework for estimating the spike detection threshold. As shown in sections 3.1 and 3.2, a transformation of neural data to its analytic form reveals two statistical components: an exponential term caused by the regulation of CLT (neuron noise) and a power term caused by the violation of Lyapunov’s condition (spikes). In sections 3.3 and 3.4, a threshold estimation method is proposed based on these two components.

3.1. Neural interface noise—exponential component

**Definition.** A real sequence $V(t)$ (neural data) and its Hilbert transform $HV(t)$ are related to each other that they together form a strong analytic signal $V_a(t)$ [42–44]:

$$V_a(t) = V(t) + iHV(t) = V(t) + i\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{V(\tau)}{t - \tau} d\tau, \quad (3)$$

where $P$ in front of the integral denotes the Cauchy principal value and $H$ denotes the Hilbert transform.

We choose to detect neural spikes based on the analytic signal $|V_a(t)| = \sqrt{V(t)^2 + HV(t)^2}$ rather than $V(t)$ (although they are approximately equivalent) for two reasons. First, extracellular spike could have significant variation in shape; sometimes they may require multiple detection windows of different thresholds. As a comparison, the corresponding analytic signal (square root of instantaneous energy) has less variation in shape and only require a single threshold for different shaped spikes, as illustrated in figure 5. Second, as to be derived here, background noise has a simple representation in Hilbert space. Denote the discrete versions of $V(t)$ and $HV(t)$ as $V(m\Delta T)$ and $HV(m\Delta T)$, where $m = \ldots, -1, 0, 1, 2, \ldots$ and $\Delta T$ is the sampling interval. By definition, $HV(m\Delta T)$ is a weighted sum of a series of correlated Gaussian random variables that approximately converges to Gaussian. The dependence between $V(m\Delta T)$ and $HV(m\Delta T)$ can be quantified through a modified cosine similarity $Y_V$ (from completely independent to dependent, $Y_V$ increases from 0 to 1):

$$Y_V = 1 - \frac{\int_{-\infty}^{\infty} f_V(x_1) f_{HV}(x_2) V_{tHV}(x_1, x_2) dx_1 dx_2}{\int_{-\infty}^{\infty} f_{HV}(x_1) f_{HV}(x_2) dx_1 dx_2}, \quad (4)$$

where $f_V$ and $f_{HV}$ are the density functions of $V(m\Delta T)$ and $HV(m\Delta T)$, respectively; $f_{tHV}$ is the joint density function. Measured results based on in vivo data show that $Y_V$ is in a range from 0.0007 (associated with inactive states, e.g.,...
of analytic signal following property of neuron noise: define Hilbert transform does not change the signal power, it gives the are nearly independent Gaussian variables. Added that the SD of background noise.

Figure 5. Neural spike data (black) and their analytic waveforms (brown). (a) A burst of spikes with 1–2 ms width and 100 µV magnitude. (b) A burst of spikes with 2–4 ms width and 1 mV magnitude. Waveforms in both (a) and (b) are picked from a same recorded sequence with 3 s of separation.

As discussed in section 3.2, spikes are large in magnitude and low in frequency, making Θ in equation (6) small. Consequently, Z follows an exponential distribution \( f(Z) = 10^{-1.7} e^{-0.503Z} \). As a comparison shown in figure 6(c), spikes have a large magnitude and high firing rates, which degenerate Lyapunov’s condition; as a result, \( f(Z) \) deviates from an exponential distribution as plotted in figure 6(d). In the following, we quantitatively investigate this deviation.

Following equation (1), the analytic signal of recorded neural data \( V_s(mΔT) \) is

\[
V_s(mΔT) = \sum_{i=1}^{A} V_{si}(mΔT) + N(mΔT),
\]

where \( Θ \) is a parameter to quantify the validity of Lyapunov’s condition (a smaller \( Θ \) suggests Lyapunov’s condition holds better); \( L \) is used to identify the variable of the largest variance (large spikes), \( |T_i| \) is the spike width, \( r_L \) is the firing rate, \( f_s \) is the sampling frequency, \( σ^2 \) is neural data variance and \( σ_L^2 \) represents the spike variance within the waveform window. If we assume a homogeneous neuron density and independent firings, \( Θ \) converges to 0 as \( N \) increases. However, the convergence rate is slower than linear (~\( N^{1/3} \)). The slow convergence makes the validity of Lyapunov’s condition to rely on a hypothesis of small variance of each individual variable, which is proportional to its firing rate and induced spike power. As an example shown in figure 6(a), spikes are large in magnitude and low in frequency, making Θ in equation (6) small. Consequently, Z follows an exponential distribution \( f(Z) = 10^{-1.7} e^{-0.503Z} \). As a comparison shown in figure 6(c), spikes have a large magnitude and high firing rates, which degenerate Lyapunov’s condition; as a result, \( f(Z) \) deviates from an exponential distribution as plotted in figure 6(d). In the following, we quantitatively investigate this deviation.

Following equation (1), the analytic signal of recorded neural data \( V_s(mΔT) \) is

\[
V_s(mΔT) = \sum_{i=1}^{A} V_{si}(mΔT) + N(mΔT),
\]

where \( Θ \) is a parameter to quantify the validity of Lyapunov’s condition (a smaller \( Θ \) suggests Lyapunov’s condition holds better); \( L \) is used to identify the variable of the largest variance (large spikes), \( |T_i| \) is the spike width, \( r_L \) is the firing rate, \( f_s \) is the sampling frequency, \( σ^2 \) is neural data variance and \( σ_L^2 \) represents the spike variance within the waveform window. If we assume a homogeneous neuron density and independent firings, \( Θ \) converges to 0 as \( N \) increases. However, the convergence rate is slower than linear (~\( N^{1/3} \)). The slow convergence makes the validity of Lyapunov’s condition to rely on a hypothesis of small variance of each individual variable, which is proportional to its firing rate and induced spike power. As an example shown in figure 6(a), spikes are large in magnitude and low in frequency, making Θ in equation (6) small. Consequently, Z follows an exponential distribution \( f(Z) = 10^{-1.7} e^{-0.503Z} \). As a comparison shown in figure 6(c), spikes have a large magnitude and high firing rates, which degenerate Lyapunov’s condition; as a result, \( f(Z) \) deviates from an exponential distribution as plotted in figure 6(d). In the following, we quantitatively investigate this deviation.

Following equation (1), the analytic signal of recorded neural data \( V_s(mΔT) \) is

\[
V_s(mΔT) = \sum_{i=1}^{A} V_{si}(mΔT) + N(mΔT),
\]

where \( Θ \) is a parameter to quantify the validity of Lyapunov’s condition (a smaller \( Θ \) suggests Lyapunov’s condition holds better); \( L \) is used to identify the variable of the largest variance (large spikes), \( |T_i| \) is the spike width, \( r_L \) is the firing rate, \( f_s \) is the sampling frequency, \( σ^2 \) is neural data variance and \( σ_L^2 \) represents the spike variance within the waveform window. If we assume a homogeneous neuron density and independent firings, \( Θ \) converges to 0 as \( N \) increases. However, the convergence rate is slower than linear (~\( N^{1/3} \)). The slow convergence makes the validity of Lyapunov’s condition to rely on a hypothesis of small variance of each individual variable, which is proportional to its firing rate and induced spike power. As an example shown in figure 6(a), spikes are large in magnitude and low in frequency, making Θ in equation (6) small. Consequently, Z follows an exponential distribution \( f(Z) = 10^{-1.7} e^{-0.503Z} \). As a comparison shown in figure 6(c), spikes have a large magnitude and high firing rates, which degenerate Lyapunov’s condition; as a result, \( f(Z) \) deviates from an exponential distribution as plotted in figure 6(d). In the following, we quantitatively investigate this deviation.

Following equation (1), the analytic signal of recorded neural data \( V_s(mΔT) \) is

\[
V_s(mΔT) = \sum_{i=1}^{A} V_{si}(mΔT) + N(mΔT),
\]

where \( Θ \) is a parameter to quantify the validity of Lyapunov’s condition (a smaller \( Θ \) suggests Lyapunov’s condition holds better); \( L \) is used to identify the variable of the largest variance (large spikes), \( |T_i| \) is the spike width, \( r_L \) is the firing rate, \( f_s \) is the sampling frequency, \( σ^2 \) is neural data variance and \( σ_L^2 \) represents the spike variance within the waveform window. If we assume a homogeneous neuron density and independent firings, \( Θ \) converges to 0 as \( N \) increases. However, the convergence rate is slower than linear (~\( N^{1/3} \)). The slow convergence makes the validity of Lyapunov’s condition to rely on a hypothesis of small variance of each individual variable, which is proportional to its firing rate and induced spike power. As an example shown in figure 6(a), spikes are large in magnitude and low in frequency, making Θ in equation (6) small. Consequently, Z follows an exponential distribution \( f(Z) = 10^{-1.7} e^{-0.503Z} \). As a comparison shown in figure 6(c), spikes have a large magnitude and high firing rates, which degenerate Lyapunov’s condition; as a result, \( f(Z) \) deviates from an exponential distribution as plotted in figure 6(d). In the following, we quantitatively investigate this deviation.

Following equation (1), the analytic signal of recorded neural data \( V_s(mΔT) \) is

\[
V_s(mΔT) = \sum_{i=1}^{A} V_{si}(mΔT) + N(mΔT),
\]

where \( Θ \) is a parameter to quantify the validity of Lyapunov’s condition (a smaller \( Θ \) suggests Lyapunov’s condition holds better); \( L \) is used to identify the variable of the largest variance (large spikes), \( |T_i| \) is the spike width, \( r_L \) is the firing rate, \( f_s \) is the sampling frequency, \( σ^2 \) is neural data variance and \( σ_L^2 \) represents the spike variance within the waveform window. If we assume a homogeneous neuron density and independent firings, \( Θ \) converges to 0 as \( N \) increases. However, the convergence rate is slower than linear (~\( N^{1/3} \)). The slow convergence makes the validity of Lyapunov’s condition to rely on a hypothesis of small variance of each individual variable, which is proportional to its firing rate and induced spike power. As an example shown in figure 6(a), spikes are large in magnitude and low in frequency, making Θ in equation (6) small. Consequently, Z follows an exponential distribution \( f(Z) = 10^{-1.7} e^{-0.503Z} \). As a comparison shown in figure 6(c), spikes have a large magnitude and high firing rates, which degenerate Lyapunov’s condition; as a result, \( f(Z) \) deviates from an exponential distribution as plotted in figure 6(d). In the following, we quantitatively investigate this deviation.

Following equation (1), the analytic signal of recorded neural data \( V_s(mΔT) \) is

\[
V_s(mΔT) = \sum_{i=1}^{A} V_{si}(mΔT) + N(mΔT),
\]

where \( Θ \) is a parameter to quantify the validity of Lyapunov’s condition (a smaller \( Θ \) suggests Lyapunov’s condition holds better); \( L \) is used to identify the variable of the largest variance (large spikes), \( |T_i| \) is the spike width, \( r_L \) is the firing rate, \( f_s \) is the sampling frequency, \( σ^2 \) is neural data variance and \( σ_L^2 \) represents the spike variance within the waveform window. If we assume a homogeneous neuron density and independent firings, \( Θ \) converges to 0 as \( N \) increases. However, the convergence rate is slower than linear (~\( N^{1/3} \)). The slow convergence makes the validity of Lyapunov’s condition to rely on a hypothesis of small variance of each individual variable, which is proportional to its firing rate and induced spike power. As an example shown in figure 6(a), spikes are large in magnitude and low in frequency, making Θ in equation (6) small. Consequently, Z follows an exponential distribution \( f(Z) = 10^{-1.7} e^{-0.503Z} \). As a comparison shown in figure 6(c), spikes have a large magnitude and high firing rates, which degenerate Lyapunov’s condition; as a result, \( f(Z) \) deviates from an exponential distribution as plotted in figure 6(d). In the following, we quantitatively investigate this deviation.

Following equation (1), the analytic signal of recorded neural data \( V_s(mΔT) \) is

\[
V_s(mΔT) = \sum_{i=1}^{A} V_{si}(mΔT) + N(mΔT),
\]
law; $M$ is inversely proportional to the distance between the source (neuron) and the measuring point (electrode). Based on this property, the pdf of the number of neurons with respect to $M$, $\rho(M)$, is

$$\rho(M) = c r(M)^2 \left| \frac{d\rho(M)}{dM} \right| \propto M^{-4}, \quad (8)$$

where $c$ is a constant relating to neuron density (number of neurons per mm$^2$), and $r(M)$ is the distance from a targeted neuron to the recording site.

Assume that an analytic spike of magnitude $M$ introduces $W$ equally spaced samples on average ($W \approx \frac{Z}{\Delta T}$), based on equation (8), the density function of spike power $f_d(Z)$ is

$$f_d(Z) \approx \int_{-\infty}^{+\infty} \rho(M) \frac{W}{M} \frac{dM}{dZ} \left| \frac{d\rho(M)}{dM} \right| \frac{dx}{dx} \frac{1}{Z^4} \propto \frac{1}{Z^3}, \quad (9)$$

where $f_d(Z)$ is the added component by neural spikes, causing $f(Z)$ to deviate from Gaussian.

Combine equations (5) and (9); $f(Z)$ is a combination of an exponential component ($e^{-\lambda_1 Z}$, EC) and a power-law component ($Z^{-\lambda_2}$, PC); $e^{-\lambda_1 Z}$ is caused by background noise regulated by the CLT, and $Z^{-\lambda_2}$ is caused by detectable spikes that violate Lyapunov’s condition of small variance of individual variables. Typical results derived from in vivo data are shown in figure 7, which clearly confirm the existence of both $e^{-\lambda_1 Z}$ and $Z^{-\lambda_2}$.

3.3. Threshold estimation

Assume $\tilde{f}_d(Z)$ and $f_d(Z)$ the exponential component and the power-law component trained in real-time, and $f(Z) = \tilde{f}_d(Z) + f_d(Z)$, $\int_0^\infty f(z) dz = 1$. Given a sample $Z(m\Delta T) = Z_0$, the probability of the presence of a spike $p_s(m\Delta T)$ can be quantitatively accessed, e.g., using the maximum-likelihood estimation:

$$p_s(m\Delta T) = \frac{\tilde{f}_d(Z_0)}{f_d(Z_0) + \tilde{f}_d(Z_0)} |Z_0 = \lambda_1 (m\Delta T)^2 + \lambda_2 (m\Delta T)^2|^{-1}. \quad (10)$$

As discussed in section 1, a quantitative evaluation of threshold requires the benchmark as a priori that is not
available. $p_s(m\Delta T)$ in equation (10) can be used as a replacement of the benchmark. The optimal threshold is the one that generates spike firing patterns matching $p_s(m\Delta T)$ most: given a segmentation scheme of neural data $w_k$, $k = 1, 2, 3, \ldots$, and a threshold, the detected spike patterns are noted as $P_6(k)$. An integration procedure over $p_s(m\Delta T)$ is applied over the same segmentation scheme; in this work, we use a simple winner-take-all strategy that captures the peak instantaneous energy regardless of the waveform width

$$P_t(k) = \max_{m\in[w_k, w_{k+1}]} [p_s(m\Delta T)],$$

(11)

where $P_t(k)$ is the probability that a spike appears in the $k$th segment and an optimal detection threshold is the one that maximizes the similarity between $P_t(k)$ and $P_6(k)$.

3.4. Step-by-step algorithm recipe

A step-by-step algorithm explanation correlated with algorithm flow shown in figure 8 are given as follows.

Data preparation: high pass filtering raw data at 300 Hz to remove low frequency activities. Filtered data are noted as $V(t)$, which are contributed by a large number of background neurons.

Step 1. Applying the Hilbert transform to the high pass filtered neural data $V(t)$ as shown in equation (3). According to equations (3)–(5), neuron noise has an exponential distribution $e^{-\sigma_2^2}$, where $\sigma_2^2$ is the noise variance and $Z$ is the instantaneous energy in Hilbert space. Because noise variance is always smaller than the data variance $\lambda^2$, thus the exponential distribution takes a general form $e^{-\lambda^2Z}$ with $\lambda^2 > \frac{1}{2\tau^2}$.

According to equations (6) to (9), detectable neural spikes generate a power-law distribution $Z^{-\lambda^2}$. It can be attributed to violations of Lyapunov’s conditions: some neurons are more close to the electrode and induce large spikes that invalidate the hypothesis of small variance of individual variables. The induced spike magnitude is inversely proportional to the distance between the neuron and the electrode, and consequently, detectable neural spikes give a power-law distribution. In the ideal case that neurons are homogeneously distributed and each with equal transmembrane current density, the distribution follows $Z^{-2.5}$.

Figure 7. (a) Analytic waveforms of recorded data with thresholds labeled for comparison. (b) Pdf of analytic signal power $f(Z)$. For this sequence, the crossing point of $e^{-\lambda_1^2Z}$ and $Z^{-\lambda_2^2}$ (EC/PC) suggests over 30 times increased spike events compared with using $3\times$ RMS threshold. (c) When $Z/\sigma^2 < 1$, $e^{-\lambda_1^2Z}$ dominates $Z^{-\lambda_2^2}$; therefore, $f(Z)$ is fitted by an exponential trace in the log scale. (d) As $Z$ increases, $Z^{-\lambda^2}$ dominates $e^{-\lambda_1^2Z}$, and $f(Z)$ is fitted by a power component in the log–log scale.
Figure 8. Step-by-step flow of the proposed spike detection algorithm.

Figure 9. (left) EC/PC decomposition. The x-axis is signal power normalized data variance; the y-axis is the probability in log_{10}. Red (circle) line is an inferred EC (noise) and blue (square) line is an inferred PC (signal). (right) Superimposed EC+PC in comparison with neural data energy distribution (f(Z)).

Step 2. Applying a linear decomposition algorithm to separate $f(Z)$ into two components, $\tilde{f}_n(Z) \sim e^{-\lambda_1 Z}$ (noise) and $\tilde{f}_d(Z) \sim \frac{1}{\sigma^2_n} Z^{-\lambda_2}$ (signal). For $\tilde{f}_d(Z)$, the regulation at small $Z$ is because the noise variance $\sigma^2_n$ implicitly sets a lower boundary for $s$ in equation (9). In other words, if the magnitude of induced spikes is too small, spikes contribute to background noise and cannot be reliably detected. A snapshot of typical components decomposition results is shown in figure 9.

Step 3. Generating a ‘spiking probability map’ $P_s$ based on equations (10) and (11). The estimated $\tilde{f}_n$ and $\tilde{f}_d(Z)$ are the input. The ‘spiking probability’ refers to the chance of the presence of a spike at a given time, as shown in figure 10.

Step 4. Choosing a spike detection threshold based on the generated ‘spiking probability map’. The user can choose detection threshold based on the spiking probability, e.g., 80%. (On average, 80 out 100 detected spikes are true spikes. The rest 20 are from noise.) Or alternatively, the user can calculate an optimal threshold that gives the maximal temporal similarity between the detection results (1/0 pattern) and the spiking probability map.

4. Data analysis results

In section 4.1, we use animal data to argue that the estimated detection threshold correlates with the background spiking activities. Furthermore, we report that there are substantial variations in estimated detection threshold for (1) data recorded from the same channel but at different time slots and (2) data recorded from different channels at the same time.

In section 4.2, we use synthesized data to quantitatively evaluate the algorithm performance. We report that the proposed algorithm can predict receiver operator characteristics (ROC) curves that are almost identical to benchmarks. The results confirm that the proposed spiking probability map can be used for choosing detection threshold. Finally, quantitative comparisons between the predicted
spiking probabilities and measured probabilities are included to validate the algorithm prediction accuracy.

4.1. Validation using experimental in vivo animal data

4.1.1. Validation of two statistical components. Testing results on 109 in vivo sequences are summarized in figure 11. The data are independently recorded from hippocampus, cortex surface and spinal cord; both acute and long-term recordings; and sleep and awake states. Typical in vivo neural data illustrating different firing rates, on/off duties and SNR are shown in figure 12 along with the corresponding EC/PC decomposition results. Over ~30% sequences tested (an example is shown in figure 7), a threshold near the crossing point of $e^{-x^2}$ and $Z^{-k_2}$ (EC/PC) gives >10 times more spikes than $\lambda$RMS does, suggesting a large headroom for increasing information capacity.

On all 109 sequences tested, $\lambda_1/\sigma^2$ is averaged to be 0.7719 ± 0.5082, as shown in figure 11(a). The histogram of $\lambda_1/\sigma^2$ clearly peaks at ‘0.5’, which is consistent with the prediction by equation (5). A power component $Z^{-k_2}$ has been observed in all the sequences. As summarized in figure 11(b), $\lambda_2 = 2.516 ± 0.564$ astonishingly matches the predicted ‘2.5’ by equation (9). The model fitting of $f(Z)$ using $e^{-x^2}$ and $Z^{-k_2}$ in the log scale have passed statistical validations: the repeatability measured by $R^2$ score [45] is 99.56 ± 0.14%; the accuracy measured by the root mean square deviation (RMSD) is 0.056 ± 0.014. As summarized in figures 11(c) and (d), the estimated optimal threshold exhibits significant variations from 23 to 155 $\mu$V.

4.1.2. EC/PC crossing point variation over time. As illustrated in figure 13(a)–(c), $Z^0.5_{EC/PC}$ (the predicted threshold that corresponds to 50% correct detections and 50% false detections) versus time traces are plotted, where data are recorded from hippocampus, cortex and cortex superficial layer. We find out in some cases $Z^0.5_{EC/PC}$ does not change much (<30%) over time, e.g., figures 13(a) and (b); while in some other cases, $Z^0.5_{EC/PC}$ may change suddenly, e.g., figure 13(c). Our experiments confirm that $Z^0.5_{EC/PC}$ well tracks spiking activities: when spiking activities are frequent and differentiable from noise (a high SNR), $Z^0.5_{EC/PC}$ tends to be lower than 2.5 (Figure 13(a)). When spiking activities are infrequent, $Z^0.5_{EC/PC}$ tends to be higher (Figure 13(c) and (f)). The results here are consistent with the prediction that when individual neural spikes with variances comparable to the background data RMS, $\Theta$ in equation (10) is a small number ($\ll 1$) and neural spikes merge to background noise according to Lyapunov’s condition. In this case, an optimal detection threshold will be biased to reduce the false alarm according to the maximum-likelihood criterion, and thus a larger $Z^0.5_{EC/PC}$.

When neurons induce frequent and large magnitude spikes, $\Theta$ in equation (10) becomes larger and neural activities form a new distribution (PC in Hilbert space). In this case, the neural data RMS should be notably larger than the background noise RMS, and thus $Z^0.5_{EC/PC}$ is smaller. As shown in figures 13(c) and (f), $Z^0.5_{EC/PC}$ may change over time. If $Z^0.5_{EC/PC}$ is indeed coupled with the optimal detection
Figure 11. (a) Measured histogram of $\lambda_1/\sigma^2$. As predicted by equation (5), it peaks at 0.5. (b) Measured histogram of $\lambda_2$, astonishingly matching the predicted 2.5 by equation (9). (c) Histogram of EC/PC crossing point ($Z_{EC/PC}^{0.5}$) in $\mu$V. (d) Estimated threshold; x-axis: data SD in $\mu$V, and y-axis: estimated optimal threshold referred to as $\sigma (Z/\sigma^2)$. Data statistics are derived from 109/109 in vivo recordings contributed from independent experiments: $\lambda_1/\sigma^2 = 0.7719 \pm 0.5082, \lambda_2 = 2.516 \pm 0.564, R^2 = 99.56 \pm 0.14\%$ (repeatability), RMSD = 0.056 ± 0.014 (accuracy). $f(Z)$ is quantized at a resolution of $10^{-5}$.

To investigate $Z_{EC/PC}^{0.5}$ statistics over different recording channels, multi-channel recording experiments are performed using the Plexon system with 1–8000 Hz bandwidth on the amplifiers. Adjacent electrodes are separated by 150 $\mu$m to allow studying the spatial sensitivity of $Z_{EC/PC}^{0.5}$. Figure 15 summarizes derived $Z_{EC/PC}^{0.5}$ at different channels and different trails: each row is one trail that consists of five channels. Experiment data suggest that there exists substantial variation in $Z_{EC/PC}^{0.5}$ among different recording channels (even two adjacent channels), and thus independent threshold estimation should be performed for different channels.

4.2. Validation using synthesized data

4.2.1. Prediction 50%–50% threshold. To quantitatively evaluate the proposed detection algorithm, we have prepared a synthesized data base to mimic real recordings. As illustrated in figure 16, recorded in vivo data that contain a small number
of visually detectable spikes are used as the background noise; recorded spike waveforms with a large magnitude (>500 μV) are used to model the shapes of neural spikes. Independent neurons of varied transmembrane current magnitudes (to model variations in axonal radius and ion channel density) are randomly added within a few hundred μm radius of a point electrode. The induced spikes attenuate according to the distance to the electrode. The amplitudes of added neurons are in a wide range from 5 μV (small and distant neurons, more likely to appear) to 1.3 mV (large and close neurons, less likely to appear). The maximal number of neurons added in this simulation is 50, allowing overlapping spikes. Neurons’ firings are assumed to follow the inhomogeneous Poisson process with varied on/off firing states; the averaged firing
Figure 14. (a) $Z_{EC/PC}^{0.5}$ versus time, where there is a sudden step-down during 300–350 s time marker. (b) Bandpass filtered neural data at 300 Hz–8 kHz during 0–50 s time marker, where there are less visually detectable spikes. (c) Bandpass filtered neural data during 300–350 s time marker, where there is a sudden increase in spiking activities starting at 335 s time marker. The change in spiking rate is coupled with $Z_{EC/PC}^{0.5}$, as shown in (a). (d) Bandpass filtered neural data in 400–500 s, which have shown sustained high-rate spiking activities.

rates are randomly picked from 0.1 to 10 Hz. The rest of neurons are assumed to fire a train of spikes each time with the 2 mS refractory period.

Figure 17 displays the predicted EC/PC crossing point $Z_{EC/PC}^{0.5}$ (x-axis) compared with the measured threshold $V_{T.50\%}$ (y-axis) that incurs 50% correct detections and 50% false detections: $Z_{EC/PC}^{0.5}$ is predicted by equation (10); $V_{T.50\%}$ is obtained by sweeping the detection threshold and comparing results with the benchmark. According to equation (10), $V_{T.50\%}$ should be very close to $Z_{EC/PC}^{0.5}$ due to an exponential sensitivity of equation (10) around 0.5. The fitting results summarized from different sequences are plotted in figure 17, where the x-axis is $Z_{EC/PC}^{0.5}$ and the y-axis is measured $V_{T.50\%}$ (both normalized to data RMS), each dot is one synthesized sequence, and the solid line is fitted to minimize the RMS deviation. The best fitting is $V_{T.50\%} = 0.987 Z_{EC/PC}^{0.5} - 0.080$, which astonishingly matches the predicted one $V_{T.50\%} \approx Z_{EC/PC}^{0.5}$.

To allow a quantitative evaluation of the prediction statistics (given a threshold, we predict the numbers of correct detections and false detections using ‘ground truth’ data), we have built ROC-typed detection curves as shown in figure 18. The procedures to predict detection statistics have been described as a step-by-step algorithm recipe in section 3.4, where the output is a ‘spiking probability map’ that tells the chance of the presence of a spike at a given time slot. Given a threshold, the predicted ‘spiking probability map’ can output detected spike patterns, which have been further compared with the ‘ground truth’ data to get quantitative feedbacks of positive detections/false detections.

If the proposed theory works, ROC curves derived from the ‘spiking probability map’ should be consistent with ROC curves derived from other detection methods, especially the amplitude-based approach as detections from an analytic signal and from an original signal are more or less equivalent.

To be more objective on building the benchmark ROCs, we have tried three types of detection methods: amplitude, nonlinear energy operator and matched filter. As shown in figure 18, the ROC curves derived from predicted ‘spiking probability maps’ (labeled as EC–PC) are indeed the consistent
Five rows are derived from five preparations. Figure 16. Protocols to synthesize neural data. Neuron noise and a large number of spike templates (spikes over 200 $\mu$V) are first extracted from in vivo recordings. Background neurons are randomly added around the recording electrode. Each neuron produce spikes from a given template following a randomized inhomogeneous Poisson process. The amplitude of a spike is re-scaled according to the distance between the neuron and the recording electrode with added variation to mimic ion-channel density fluctuations.

Figure 15. $z_{EC_F}^{1.5}$ in multi-channel recording experiment. Each row of figures gives results derived from different channels of one preparation. Five rows are derived from five preparations.
while the RMS noise is 20–30 µV to 1.3 mV; each dot represents one sequence. Neurons are randomly induced spike amplitude in a wide range from 5 µV to 1.3 mV. Best fitting: $y = 0.987x - 0.080$ ($R^2 = 0.935$) compared with the prediction by equations (5), (9) and (10): $y = x$.

Figure 17. Predicted EC/PC crossing point ($x$-axis) versus measured threshold ($y$-axis) that incurs 50% correct detections and 50% false detections ($T_{50\%}$). Synthesized sequences are filtered at 300 Hz–5 kHz; each dot represents one sequence. The induced spike amplitude is in a wide range from 5 µV to 1.3 mV; while the RMS noise is 20–30 µV. Best fitting: $y = 0.987x - 0.080$ ($R^2 = 0.935$) compared with the prediction by equations (5), (9) and (10): $y = x$.

The benchmarks (labeled as Abs, NEO, MF-Abs) at different SNRs, i.e. 1.5 dB and above, which indirectly validates the spiking probability map.

Figure 19 shows the predicted spiking probability (given data segments of a few ms, the predicted probability that it contains a spike by equation (11)) and its comparison with the measured spiking probability. In this experiment, data segments that have similar predicted spiking probability ($\pm 2.5\%$) are grouped together and compared with the ‘ground truth’ data to measure the actual spiking probability. The experiment has shown that the predicted spiking probability is consistent with the measured probability over a wide range of SNRs and firing rates. Results in figures 18 and 19 together confirm that ‘spiking probability maps’ can be used as a priori to estimate an appropriate detection threshold.

5. Summary

A novel theory for modeling in vivo neural signals and noise has been presented with rigid experimental verifications. Our study has exposed two statistical components embedded in neural data and shown that they can be used for spike detection. When tested with in vivo data, the derived...
Figure 19. Spiking probability prediction and verification. The x-axis is the predicted spiking probability: given a data segment of a few ms, the predicted probability of the presence of a spike; the y-axis is the measured spiking probability: data segments that have similar predicted spiking probabilities (±2.5%) are grouped together and compared with the ‘ground truth’ data to measure the actual spiking probability. (a) 40 neurons with an averaged firing rate of 1.8 Hz and (b) 30 neurons with an averaged firing rate of 2.1 Hz are included. The data preparation protocol follows figure 16. The summed spiking rate in the experiment is set to below 100 Hz to avoid bias to detection.

threshold exhibits significant variations from 23 to 155 µV (1–6 RMS), suggesting that the conventional approach of fixing the threshold at a constant rms level is not reliable and may cause significant losses in information. In addition, $Z_{0.05}^{EC/PC}$ (the threshold corresponds to 50% positive detection and 50% false alarm) may change over time and vary from one channel to a different channel, which requires the training algorithms to be independent over different channels and adaptive over time. When tested with synthesized data, the predicted ROC curves match the benchmarks at a varied SNR and the predicted spiking probability is consistent with the measurement. We envision that this work may be applied to a wide range of experiments as a front-end data analysis tool.

6. Protocol

Rat data are provided by Dr Edward Keefer at Plexon Inc. The protocols can be found in [46]. Cat data are provided by Dr Victor Pikov at Huntington Medical Research Institute. The neural recordings were collected from the cerebral cortex in several cats, as previously described [47]. Briefly, the 16-channel (4 × 4) electrode arrays with a nominal geometric area of exposed electrode tips of 2000 µm² were purchased from Blackrock Microsystems. To decrease the electrode impedance, the electrodes were coated with the sputtered iridium oxide (SIROF) at the EIC Labs using the previously established procedure [48]. The array was chronically implanted in the sensorimotor cortex and connected to a percutaneous connector mounted in the animal’s head, in accordance with the HMRI Institutional Animal Care and Use Committee-approved protocol and in compliance with the USDA Animal Welfare Act. At 10–100 days after the array implantation, the animal was lightly anesthetized by an intramuscular injection of ketamine (11 mg kg⁻¹) and acepromazine (0.1 mg kg⁻¹), allowing the animal to breathe on its own and to limit the suppression of spontaneous cortical neuronal activity. Neural data were recorded at 16 bit and 25 ksp by a custom amplifier and a data acquisition board (USB-6259, National Instruments).

Acknowledgments

The authors acknowledge Dr Edward Keefer at Plexon for providing part of the experiment data and Harvey Wiggins at Plexon for helpful suggestions. This research was supported by National University of Singapore grant (R-263-000-619-133), A*STAR PSF grant (R-263-000-699-305) and A*STAR SERC grant (R-263-000-656-305).

References

[31] Teich M, Johnson D, Kumar A and Turcott R 1990 Rate fluctuations and fractional power-law noise recorded from cells in the lower auditory pathway of the cat Hear. Res. 46 41–52
[38] Rice J 1995 Mathematical Statistics and Data Analysis (Belmont, CA: Duxbury)