Recovery of Stimuli Encoded with a Hodgkin-Huxley Neuron Using Conditional PRCs

Anmo J. Kim and Aurel A. Lazar*

Abstract Understanding neural encoding/decoding mechanisms is one of the most fundamental problems in the field of sensory neuroscience. The Hodgkin-Huxley equations provide an explicit description of an encoding mechanism. However, the daunting complexity of the Hodgkin-Huxley equations makes the task of recovery of stimuli encoded with a Hodgkin-Huxley neuron particularly challenging. A highly effective strategy calls for reducing the Hodgkin-Huxley neuron to a project-integrate-and-fire (PIF) neuron. Using the reduced PIF model, we present three different recovery algorithms for stimuli encoded with a Hodgkin-Huxley neuron. All algorithms reconstruct the stimuli from the neuron's output spike train. The first algorithm is based on the assumption that the Hodgkin-Huxley neuron has a known PRC. The second algorithm assumes that the PRC is conditionally known on each inter-spike time interval. Finally, the third algorithm operates under the assumption that the conditional PRC is unknown and has to be estimated. We establish an estimate of the conditional PRC based upon the readily observable inter-spike time interval. We compare the performance of these algorithms for a wide range of input stimuli.

Anmo J. Kim
Department of Electrical Engineering,
Columbia University, New York, NY 10027 e-mail: ak2263@columbia.edu

Aurel A. Lazar
Department of Electrical Engineering,
Columbia University, New York, NY 10027 e-mail: aurel@ee.columbia.edu

* Phase Response Curves in Neuroscience, N. W. Schultheiss, A. Prinz and R. Butera, eds., Springer, 2011.

1 Introduction

Neural circuits in sensory systems encode continuous-time stimuli, such as the concentration of an odor, the intensity of an acoustic event or the brightness of a light source. Information is represented in these circuits as a sequence of action potentials, or a spike train. Yet information can be reliably transmitted and received between neurons. Thus, a spike train may faithfully encode information, and postsynaptic neurons readily decode it.

Generation of an action potential across a cell membrane fascinated the earliest neuroscientists and opened the dawn of electrophysiological studies in neuroscience (Schuetze, 1983). Hodgkin and Huxley measured the change of ionic conductances for different membrane potential values by employing the voltage-clamp method, and their results led to a mathematical description of the action potential generation, the so called Hodgkin-Huxley neuron (Hodgkin & Huxley, 1952). From an information theory perspective, the Hodgkin-Huxley neuron provides an explicit description of encoding a continuous-time stimulus into a discrete time sequence, or spike train.

The problem of reconstruction of continuous-time stimuli from a spike train has been investigated in the literature primarily using phenomenological rather than mechanistic models. For example, one of the most popular phenomenological encoding models is the linear-nonlinear-Poisson (LNP) model. In this model, the Poisson spike generation is preceded by linear or non-linear processing blocks. For these models, a spike rate function is first recovered by constructing a Peri-Stimulus Time Histogram (PSTH) or by Gaussian filtering (Dayan & Abbott, 2001; Perkel, Gerstein, & Moore, 1967). The stimulus reconstruction is completed by taking the inverse of the processing blocks acting on the spike rate function. (Bialek & Rieke, 1991; Arcas, Fairhall, & Bialek, 2003).

However, these methods generally require many repeated trials to achieve an adequate temporal resolution and thereby average out the spike train variability across trials (Medina & Lisberger, 2007). Moreover, the validity of the phenomenological models is limited and often strongly dependent on statistical parameters of stimuli (Naka, Chan, & Yasui, 1979; Kim, Lazar, & Slutskiy, 2010). Such phenomenological model limitations call for the derivation of stimulus reconstruction algorithms from mechanistic models of the stimulus encoding process. In this chapter, we derive stimulus recovery algorithms using the explicit Hodgkin-Huxley encoding model.

The dynamics of Hodgkin-Huxley neurons for simple injected current waveforms such as steps and ramps has been extensively investigated in the literature (Izhikevich, 2007). A non-linear perturbation analysis showed that a Hodgkin-Huxley neuron with deterministic gating variables is I/O-equivalent with a project-integrate-and-fire (PIF) neuron with a variable threshold sequence (Lazar, 2007, 2010). The PIF neuron integrates a projection of the stimulus onto the phase response curve that is, in turn, modulated by a phase shift process. The phase shift process is described by a differential equation that is stimulus driven. In the absence of the small perturbation term, the stimulus is tangentially coupled into the limit

cycle of the Hodgkin-Huxley neuron and the PIF neuron reduces to an integrate-and-fire neuron.

Based on the I/O-equivalent PIF neuron a stimulus recovery algorithm was given in (Lazar, 2007, 2010). The recovery works well provided that the stimulus restricts the PRC to a small parameter set. If, however, the Hodgkin-Huxley neuron sweeps across a broad set of PRCs, different recovery algorithms are needed for improved performance. The key limitation of the recovery algorithm in (Lazar, 2007, 2010) is due to the static nature of the PRC. By introducing an adaptive sampling kernel, the PIF can provide a more faithful I/O description of the Hodgkin-Huxley neuron.

In order to achieve an accurate I/O description, the PIF is endowed with a sampling kernel conditioned on the stimulus on each inter-spike time interval. The kernel is, in effect, a conditional PRC that is parameterized by the inter-spike time interval. We shall also consider the case when the conditional sampling kernel is unknown and devise a simple kernel estimator. The proposed recovery algorithms are very flexible; they only require the inversion of a matrix. They can be used for the recovery of stimuli encoded by a large class of neurons in the limit cycle region. It remains an open problem how the dendritic tree of a postsynaptic neuron is capable of recovering the information from its incoming spike sequence.

This paper is organized as follows. In section 2, a Hodgkin-Huxley neuron is shown to be I/O-equivalent with the Project-Integrate-and-Fire neuron. The recovery of stimuli encoded with a Hodgkin-Huxley neuron is presented in section 3. Three cases are considered. In section 3.1, the stimulus is assumed to consist of a known base term and a weak bandlimited signal. In section 3.2, the problem is extended to strong stimuli by exploiting the known base terms. In section 3.3, the values of the base terms are unknown but, as shown, can be estimated. For all three cases a recovery algorithm is presented. Examples are given in section 4 and a brief discussion in section 5.

2 Reduction of the Hodgkin-Huxley Neuron to the I/O-Equivalent Project-Integrate-and-Fire Neuron

Using a vector notation, the Hodgkin-Huxley neuron model (see also the Appendix) is described by the system of differential equations (Hodgkin & Huxley, 1952)

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}) + \frac{1}{C}[I_{\mathrm{ext}}\ 0\ 0\ 0]^T,$$

where $\mathbf{x}$ and $\mathbf{f}$ are vectors of appropriate dimensions, and $\mathbf{x}(0) = \mathbf{x}_0$ is the initial condition. The state vectors can easily be identified from the set of Hodgkin-Huxley equations as $\mathbf{x} = (x_1, x_2, x_3, x_4) = (V, m, h, n)$. The expression for $\mathbf{f} = (f_1, f_2, f_3, f_4)$ can also be easily derived from the same set of equations (see the Appendix).

In what follows we shall assume that, if the (step) injected current I_{ext} is in the appropriate range, the essential dynamics of this set of equations are de-

scribed by a limit cycle (Izhikevich, 2007) (see Fig. 1(A)). In other words, when $I_{\text{ext}} = b$ (μA/cm^2), these equations have a periodic solution $\mathbf{x}^o(t,b)$ (also called an orbital solution) with period $T = T(b)$ (see Fig. 1(B)).

The dynamics of the Hodgkin-Huxley neuron are rather complex. In order to devise reconstruction algorithms, it is advantageous to find an I/O-equivalent model of the Hodgkin-Huxley neuron that is mathematically tractable. For weak stimuli, the I/O-equivalent model is the project-integrate-and-fire neuron. We shall first derive the phase equation of the Hodgkin-Huxley neuron. Since the state of the phase equation is not observable, we shall reduce it to the PIF neuron model whose spikes are observable. The Hodgkin-Huxley and PIF neuron models have the property that they are first order I/O equivalent. That means that, given the same stimulus, the spike times of these neurons are, to the first order, the same.

Because of the periodicity of the orbital solution (see Fig. 1(A)), the behavior of the Hodgkin-Huxley neuron can simply be described by its phase rather than the four state variables of the Hodgkin-Huxley equations. In this chapter, the phase of the system $\theta(t) \in [0, T(b)]$ is defined as

$$\theta(t) = t - t_k \bmod T(b),$$

where $T(b)$ is the period of the oscillation, and t_k is the last spike event ($t \geq t_k$). Note that the phase is in units of time and thus it retains the dependency on the base current b.

If a weak perturbation current $u = u(t), t \in \Re$, is injected in addition to the base current b, the system deviates from the orbit and results in either an advancement or a delay of its phase.

The phase shift in response to a current perturbation is determined by the vector space in the neighborhood of the limit cycle. The mapping of the delta-pulse current perturbation into the value of the phase shift of the orbital solution is known as the *infinitesimal Phase Response Curve (iPRC)*, or simply the PRC hereafter (Hastings & Sweeney, 1958; Winfree, 1967; Kuramoto, 1984; Ermentrout, 1996; Izhikevich, 2007). Informally, the tangential component of the perturbation to the limit cycle contributes to a phase shift, i.e., the value of the phase response curve. In contrast, the normal component of the perturbation dies out over time and does not contribute to the value of the phase of the orbital solution. This simple analysis is valid provided that the stimulus is weak.

The PRC, denoted by $\psi(t,b)$ (see Figure 2), is parameterized by b since the limit cycle $\mathbf{x}^o(t,b)$ depends on the base current. It can be obtained via a simulations method in which the change of the phase of the system is evaluated with a Dirac-delta pulse as its input. Based on Malkin's theorem (Hoppensteadt & Izhikevich, 1997), one can also obtain the PRC by solving the adjoint equation

$$\dot{\psi}(t,b) = (J(\mathbf{f}(\mathbf{x}^o(t,b))))^T \psi(t,b),$$

where $(J(\mathbf{f}(\mathbf{x}^o(t,b))))^T$ is the transposed Jacobian of $\mathbf{f}$, with the boundary condition $\psi(0,b)(\mathbf{f}(\mathbf{x}^o(0,b)) + [b\ 0\ 0\ 0]^T) = 1$. When using the simulation method, the result-

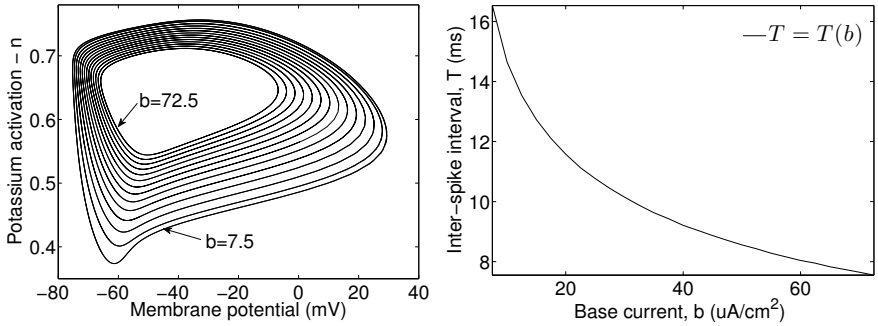


Fig. 1 The limit cycle and inter-spike time interval as a function of the base current. (A) The size of the trajectory of the limit cycle depends on the base current; it shrinks as the base current increases (step= $5\,\mu\mathrm{A}/\mathrm{cm}^2$). (B) inter-spike time interval with respect to the base current.

ing PRC is normalized by the amplitude and width of the Dirac-delta pulse. In what follows, all examples are based on PRCs evaluated by the direct simulation method.

When the dynamical system describing the Hodgkin-Huxley neuron runs along a limit cycle, its phase is, by definition, moving at a fixed velocity of 1, i.e.,

$$\dot{\theta}(t) = 1. \tag{1}$$

A weak perturbation $u(t)$ introduces an additive term to the right hand side of the equation above. This term consists of the perturbation weighted by the corresponding PRC amplitude (Winfree, 1967; Izhikevich, 2007; Kuramoto, 1984; Yoshimura & Arai, 2008), i.e.,

$$\dot{\theta} = 1 + \psi(\theta, b)u(t) + \varepsilon(t). \tag{2}$$

Here $\varepsilon(t)$ represents the model reduction error due to the finite amplitude of $u(t)$ that renders the linear approximation of the local vector field inaccurate. Note, however, that $\varepsilon(t)$ becomes arbitrarily small if $u(t)$ is arbitrarily small.

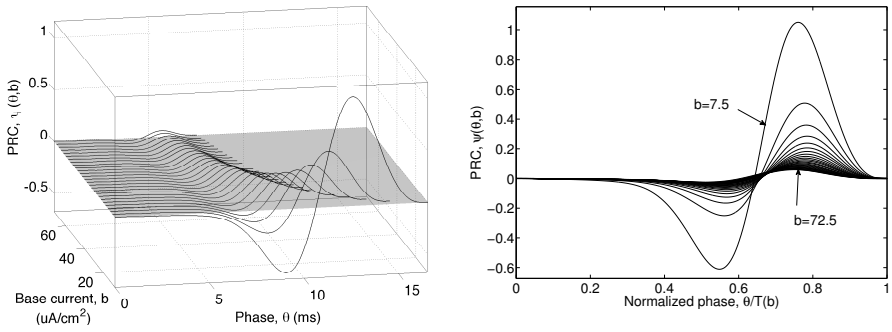
Equation (2), also called a *phase equation*, is the result of the reduction of a multi-dimensional nonlinear dynamical system to a single-dimensional differential equation. When $\varepsilon(t)$ is small enough, the phase equation is approximated by

$$\dot{\theta} = 1 + \psi(\theta, b)u(t). \tag{3}$$

The state of the phase equation is not observable. However, spike times are observable. Integrating both sides of equation (3) on an inter-spike time interval $[t_k, t_{k+1}]$ yields

$$\theta(t_{k+1}) - \theta(t_k) = \int_{t_k}^{t_{k+1}} [1 + \psi(\theta(s), b)u(s)]\, ds.$$

Since, to a first order, $\theta(t_{k+1}) - \theta(t_k) = T(b)$ (Lazar, 2007, 2010),



(A) (B)

Fig. 2 Phase response curves of the Hodgkin-Huxley neuron for different base currents ($b \in [7.5, 72.5]\,\mu\text{A}/\text{cm}^2$).) The shape of the phase response curve depends on the base current; its amplitude decreases as the base current increases (step$= 2.5\,\mu\text{A}/\text{cm}^2$). (A) PRCs depicted in 3D. (B) Projection of the PRCs onto the b axis.

$$\int_{t_k}^{t_{k+1}} [1 + \psi(\theta(s), b)u(s)]\, ds = T(b). \tag{4}$$

This equation represents the t-transform of a *project-integrate-and-fire neuron* (Lazar, 2007, 2010). The t-transform provides a measurement or sampling of the input $u = u(t)$ on the time interval between two consecutive spikes. The input stimulus is multiplied with an encoding kernel $\psi(\theta, b)$, and the next spike is generated when the integration of these values reaches the threshold value $T(b) - (t_{k+1} - t_k)$. The project-integrate-and-fire neuron is input/output equivalent to the Hodgkin-Huxley neuron under weak perturbation (Lazar, 2007, 2010). Since the evaluation of $\theta(t)$ is mathematically intractable, the PIF neuron will be further simplified.

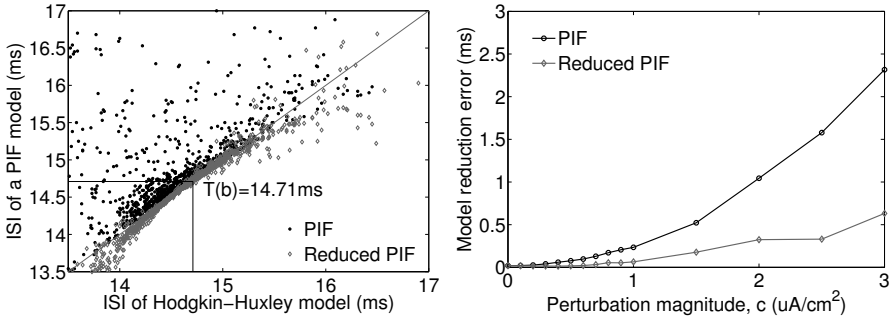
If $u(t)$ is small enough, the phase shift $\theta(t) - t$ becomes negligibly small within $[t_k, t_{k+1}]$, and thus $\psi(\theta(t), b)$ can be replaced with $\psi(t, b)$ (Lazar, 2007, 2010; Ota, Omori, & Aonishi, 2009). Equation (4) can then be rewritten as

$$\int_{t_k}^{t_{k+1}} \psi(s - t_k, b)u(s)ds = T(b) - (t_{k+1} - t_k). \tag{5}$$

A neuron model whose t-transform is given by equation (5) is called *a reduced PIF neuron*.

Note that numerical solutions of equation (4) require the evaluation of $\theta(t)$. The latter is often unstable due to the accumulation of numerical errors. In contrast, the numerical computation of the reduced PIF model is a simple dot product of two vectors and it is not prone to the accumulation of errors.

In order to assess the approximations introduced by the I/O-equivalent neurons above, we ran numerical simulations of three neural encoders: the Hodgkin-Huxley neuron, the PIF neuron, and the reduced PIF neuron. Assuming the same initial condition, the occurrences of the first two spikes were recorded and the inter-spike time intervals compared when the latter were generated by the PIF neurons and



(A) (B)

Fig. 3 Model reduction error of the project-integrate-and-fire neuron with respect to the Hodgkin-Huxley neuron (A) Inter-spike time intervals (ISI) of the PIF models with respect to the ISI of Hodgkin-Huxley neuron. (B) The model reduction error of the PIF models with respect to the perturbation magnitude. The model reduction error is defined as the difference of ISIs between the Hodgkin-Huxley neuron and the PIF models
($b = 10\,\mu\text{A/cm}^2$, $n = 50$, and $\Omega = 2\pi 50$).

the Hodgkin-Huxley neuron, respectively. The perturbation signal is bandlimited ($\Omega = 2\pi \cdot 50$ Hz) by construction

$$u(t) = c \sum_{k \in \mathbf{Z}} u(\frac{k\pi}{\Omega}) \frac{\sin(\Omega t - k\pi)}{\Omega t - k\pi}, \tag{6}$$

where $(u(k\pi/\Omega))_{k\in\mathbf{Z}}$ is drawn randomly from a uniform distribution ranging between $[-1,1]$, c represents the maximum magnitude of the perturbation.

Figure 3(A) depicts the inter-spike time intervals of the PIF models with respect to Hodgkin-Huxley model with $b = 10\mu\text{A/cm}^2$. The steady state inter-spike time interval $T(b)$ is 14.71 ms. The samples along the diagonal have zero model reduction error, and the vertically projected distance of each sample to the diagonal line is the amount of error for a particular perturbation signal. 3(B) depicts the mean absolute errors with respect to different perturbation magnitudes. In general, increasing the magnitude of the perturbation increases the model reduction error; the reduced PIF model exhibits lower errors than the (standard) PIF neuron.

3 Recovery of Stimuli Encoded with Hodgkin-Huxley Neurons

In this section we present algorithms for recovery of stimuli encoded with a Hodgkin-Huxley neuron. All algorithms recover the stimuli from the neuron's output spike train. As shown in the previous section and following (Lazar, 2007, 2010), the multi-dimensional nonlinear Hodgkin-Huxley neuron can be reduced to an I/O-equivalent PIF neuron.

In what follows we shall assume that the stimulus is a bandlimited signal of the form $u(t) = \sum_{k\in\mathbf{Z}} b_k 1_{[t_k,t_{k+1}]} + v(t)$, where b_k is the mean stimulus amplitude during the kth inter-spike time interval and $v = v(t)$ is a weak perturbation signal with zero average on each inter-spike time interval. We shall distinguish three cases: (i) $b_k = b$ is known to take a constant value (section 3.1), (ii) b_k is constant between two consecutive spike times $[t_k, t_{k+1}]$ and is known (section 3.2) and (iii) b_k is unknown (section 3.3). In all cases the range of values for b_k leads to limit cycle oscillations of the Hodgkin-Huxley neuron.

In section 3.1, we assume weak stimuli with known base current b. Since the perturbation is weak, the encoding mechanism of the system can be described by a known PRC. In section 3.2, we deal with stimuli with known base current b_k on $[t_k, t_{k+1}]$. Since the neuron oscillates on different limit cycles depending on the level of the base current, the conditional PRC characterizes the operation of the neuron on each inter-spike time interval. Finally in section 3.3, the b_k's are unknown and have to be estimated on each inter-spike time interval.

3.1 Recovery of Stimuli Encoded with Known PRCs

Following the established decoding practice of stimuli encoded with the integrate-and-fire neuron (Lazar, 2004; Lazar & Pnevmatikakis, 2008; Lazar & Tóth, 2004), the derivation of the recovery algorithm begins with the formulation of the t-transform that describes the mapping (or encoding) of the stimulus u into the spike time sequence (t_k).

As already mentioned, the t-transform of the reduced PIF is given by

$$\int_{t_k}^{t_{k+1}} \psi(s-t_k, b)u(s)ds = T(b) - (t_{k+1} - t_k). \tag{7}$$

If the stimulus $u = u(t), t \in \Re$, lives in the space of bandlimited functions with frequency support $[-\Omega, \Omega]$ then (Lazar, 2004),

$$\hat{u}(t) = \sum_{k\in\mathbf{Z}} c_k g(t - s_k), \tag{8}$$

where $g(t) = \frac{\sin \Omega t}{\pi t}$ and $s_k = \frac{t_k + t_{k+1}}{2}$. The above equality holds if the Nyquist-type rate condition (Lazar & Tóth, 2004)

$$t_{k+1} - t_k \leq \frac{\pi}{\Omega}$$

is satisfied. Replacing $u(t)$ in equation (7) with the representation in (8), we obtain

$$\sum_{l\in\mathbf{Z}} c_l \int_{t_k}^{t_{k+1}} \psi(s-t_k, b)g(s-s_l)ds = T(b) - (t_{k+1} - t_k). \tag{9}$$

Thus the coefficients $[\mathbf{c}]_l = c_l$ satisfy the matrix equation

$$\mathbf{Gc} = \mathbf{q},$$

where $[\mathbf{G}]_{kl} = \int_{t_k}^{t_{k+1}} \psi(s-t_k,b)g(s-s_l)ds$, and $[\mathbf{q}]_k = T(b) - (t_{k+1}-t_k)$. Note that, entries of the matrix $\mathbf{G}$ can be obtained from the PRC $\psi(t,b)$ and the reconstruction kernel $g(t), t \in \Re$, and the vector $\mathbf{q}$ from the period $T(b)$ of the oscillation and the spike times $(t_k), k \in \mathbf{Z}$, of the neuron. Therefore, $\mathbf{c}$ can be computed by taking the Moore-Penrose pseudo-inverse of the $\mathbf{G}$ matrix (Penrose, 1955), i.e.,

$$\mathbf{c} = \mathbf{G}^{+}\mathbf{q}.$$

Finally, the reconstruction of $u(t)$ is completed by evaluating equation (8) with the values of the vector $\mathbf{c}$.

We arrived, therefore, at the following result (Lazar, 2007, 2010)

Theorem 1. *Let $(t_k), k \in \mathbf{Z}$, be the sequence of spike times generated by a Hodgkin-Huxley neuron with input $b + u(t)$. If b is known, the stimulus can be reconstructed using*

$$\hat{u}(t) = \sum_{k \in \mathbf{Z}} c_k g(t - s_k),$$

where $g(t) = \frac{\sin(\Omega t)}{\pi t}$, $s_k = \frac{t_k + t_{k+1}}{2}$, and the coefficients $[\mathbf{c}]_k = c_k$, $k \in \mathbf{Z}$, are given by

$$\mathbf{c} = \mathbf{G}^{+}\mathbf{q},$$

where $[\mathbf{G}]_{kl} = \int_{t_k}^{t_{k+1}} \psi(s-t_k,b)g(s-s_l)ds$ and $[\mathbf{q}]_k = T(b) - (t_{k+1}-t_k)$.

An implementation of the encoding and decoding algorithms described in this section is shown in block diagram form in Fig. 4.

Remark 1. While the input to the Hodgkin-Huxley neuron is $b + u(t)$, we incorporated b into the PRC $\psi(t,b)$. Consequently, we only had to recover the weak perturbation signal $u(t), t \in \Re$.

3.2 Recovery of Stimuli Encoded with Known Conditional PRCs

In this section we consider the recovery of a bandlimited stimulus of the form $u(t) = \sum_{k \in \mathbf{Z}} b_k 1_{[t_k, t_{k+1}]} + v(t)$, where b_k is the mean stimulus amplitude during the kth inter-spike time interval and $v(t)$ is the perturbation signal. Note that although $u(t)$ is a bandlimited stimulus, $v(t) = u(t) - \sum_{k \in \mathbf{Z}} b_k 1_{[t_k, t_{k+1}]}$ is, in general, not bandlimited, since $\sum_{k \in \mathbf{Z}} b_k 1_{[t_k, t_{k+1}]}$ typically has infinite bandwidth.

Since the mean stimulus value b_k may take a distinct value on each inter-spike time interval, the t-transform has to be established separately for every interval. This can be achieved by using a formulation akin to the one in equation (7) and requires a

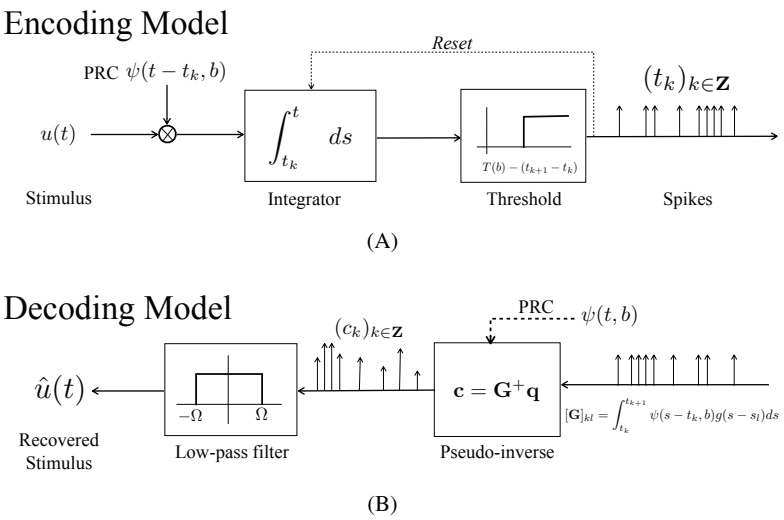


Fig. 4 Model of stimulus encoding/decoding using the project-integrate-and-fire neuron with known PRC. (A) The stimulus $u(t)$ is multiplied by the PRC determined by the known base current, and a spike is generated when the output of the integrator reaches the threshold value. (B) The known PRC is used to construct the $\mathbf{G}$ matrix, which generates an irregular spikes train, $(c_k), k \in \mathbf{Z}$, whose amplitudes are the solution of a system of linear equations. The stimulus is finally recovered by passing these spikes through a low-pass filter.

selection of the corresponding PRC for every inter-spike time interval. We formulate the t-transform as

$$\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)v(s)ds = T(b_k)-(t_{k+1}-t_k), \tag{10}$$

where b_k, is the mean stimulus amplitude for the kth inter-spike interval, $k \in \mathbf{Z}$. $\psi(t,b_k), t \in \mathfrak{R}$, for a given $k, k \in \mathbf{Z}$, is called the *conditional PRC* (see Figure 2).

As we shall require the recovered signal to live in a bandlimited space, we will replace in equation (10) $v(t)$ with $u(t)-b_k$ on each inter-spike time interval $[t_k, t_{k+1}]$, and thereby obtain the equivalent t-transform

$$\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)u(s)ds = T(b_k)-(t_{k+1}-t_k)+b_k\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)ds. \tag{11}$$

As in the known PRC case we have the following recovery algorithm.

Theorem 2. *Let $(t_k), k \in \mathbf{Z}$, be the sequence of spike times generated by a Hodgkin-Huxley neuron with bandlimited input $u(t) = v(t) + \sum_{k\in\mathbf{Z}} b_k 1_{[t_k,t_{k+1}]}$. If the mean stimulus amplitudes $(b_k)_{k\in\mathbf{Z}}$ are known, the stimulus can be reconstructed using*

$$\hat{u}(t) = \sum_{k\in\mathbf{Z}} c_k g(t-s_k),$$

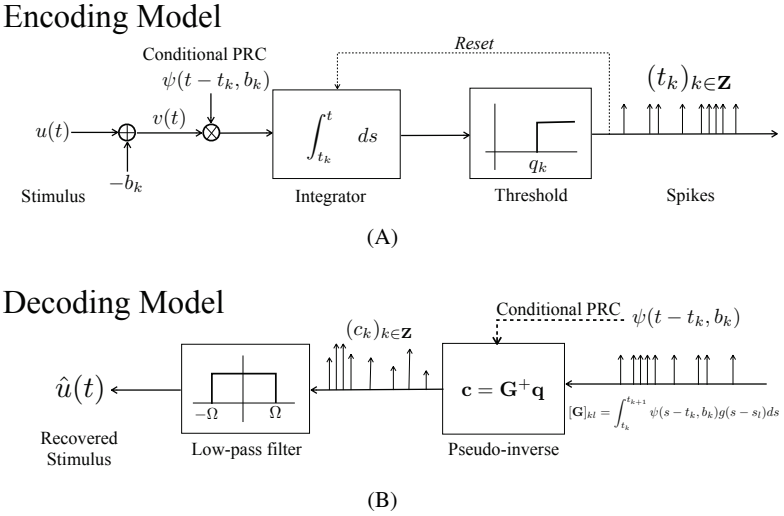


Fig. 5 Model of stimulus encoding/decoding using the project-integrate-and-fire neuron with known conditional PRC. (A) The stimulus $u(t)$ is weighted by the conditional PRC determined by the base current. (B) The sampling kernel (the known conditional PRC) is used to construct the $\mathbf{G}$ matrix. The stimulus is recovered as the response of a low-pass filter to the observed spike train.

where $g(t)=\frac{\sin(\Omega t)}{\pi t}$ *and* $s_k=\frac{t_k+t_{k+1}}{2}$, $t\in\Re, k\in\mathbf{Z}$. *The coefficients* $[\mathbf{c}]_k=c_k$, $k\in\mathbf{Z}$, *are given by*

$$\mathbf{c}=\mathbf{G}^{+}\mathbf{q},$$

where $[\mathbf{G}]_{kl}=\int_{t_k}^{t_{k+1}}\psi(s-t_k,b_k)g(s-s_l)ds$ *and* $[\mathbf{q}]_k=q_k=T(b_k)-(t_{k+1}-t_k)+b_k\int_{t_k}^{t_{k+1}}\psi(s-t_k,b_k)ds$.

An implementation of the encoding and decoding algorithms described in this section is shown in block diagram form in Fig. 5.

Remark 2. Note that in this section the recovered signal is the sum of a perturbation $v(t)$ and a variable base level current, whereas, in the previous section, the recovered signal is only the perturbation signal.

3.3 Recovery of Stimuli Encoded with Unknown Conditional PRCs

As in the previous section we consider the recovery of a bandlimited stimulus of the form $u(t)=\sum_{k\in\mathbf{Z}}b_k 1_{[t_k,t_{k+1}]}+v(t)$, where b_k is the mean stimulus amplitude during the kth inter-spike time interval. However, we shall assume here that the b_k's are unknown.

On each inter-spike time interval $[t_k,t_{k+1}]$

$$\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)u(s)ds = T(b_k)-(t_{k+1}-t_k)+b_k\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)ds. \quad (12)$$

Since for large values of b_k,

$$T(b_k)-(t_{k+1}-t_k) \ll b_k\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)ds,$$

equation (12) becomes

$$\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)u(s)ds = b_k\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)ds.$$

Normalizing by the right hand side of the above equation,

$$\int_{t_k}^{t_{k+1}} \chi(s-t_k,b_k,t_{k+1}-t_k)u(s)ds = 1, \quad (13)$$

where $\chi(t-t_k,b_k,t_{k+1}-t_k) = \psi(t-t_k,b_k)/\left(b_k\int_{t_k}^{t_{k+1}} \psi(s-t_k,b_k)ds\right)$. Since the value of b_k is unknown, we shall use on each inter-spike time interval $[t_k,t_{k+1}]$ the estimate $\hat{b}_k = T^{-1}(t_{k+1}-t_k)$ instead. Therefore, by abuse of notation,

$$\chi(t-t_k,t_{k+1}-t_k) = \psi(t-t_k,\hat{b}_k)/\left(\hat{b}_k\int_{t_k}^{t_{k+1}} \psi(s-t_k,\hat{b}_k)ds\right)$$

with $\hat{b}_k = T^{-1}(t_{k+1}-t_k)$ for $t\in[t_k,t_{k+1}]$ and zero, otherwise. Equation (13) is the t-transform of the PIF neuron and $\chi(t-t_k,t_{k+1}-t_k)$ is called the *estimated conditional sampling kernel* (eCSK) (see Figure 6). For the recovery algorithm, the same matrix formulation for the evaluation of the vector of coefficients $\mathbf{c}$ is used, i.e.,

$$\mathbf{c} = \mathbf{G}^+\mathbf{q},$$

where $[\mathbf{c}]_l = c_l$, $[\mathbf{G}]_{kl} = \int_{t_k}^{t_{k+1}} \chi(s-t_k,t_{k+1}-t_k)g(s-s_l)ds$, and $[\mathbf{q}]_k = 1$. Evaluation of the right hand side of equation (8) with the resulting values for $\mathbf{c}$ completes the recovery of $u(t)$ (see also Figure 7(B)).

Theorem 3. *Let $(t_k), k\in\mathbf{Z}$, be the sequence of spike times generated by a Hodgkin-Huxley neuron with bandlimited input $u(t) = v(t)+\sum_{k\in\mathbf{Z}} b_k 1_{[t_k,t_{k+1}]}$. If the mean stimulus amplitudes $b_k, k\in\mathbf{Z}$, are unknown, the stimulus can be reconstructed using*

$$\hat{u}(t) = \sum_{k\in\mathbf{Z}} c_k g(t-s_k),$$

where $g(t) = \frac{\sin(\Omega t)}{\pi t}$ and the coefficients $[\mathbf{c}]_k = c_k$, $k\in\mathbf{Z}$, are given by

$$\mathbf{c} = \mathbf{G}^+\mathbf{q},$$

where $[\mathbf{G}]_{kl} = \int_{t_k}^{t_{k+1}} \chi(s-t_k,t_{k+1}-t_k)g(s-s_l)ds$, and $[\mathbf{q}]_k = 1$. Finally,

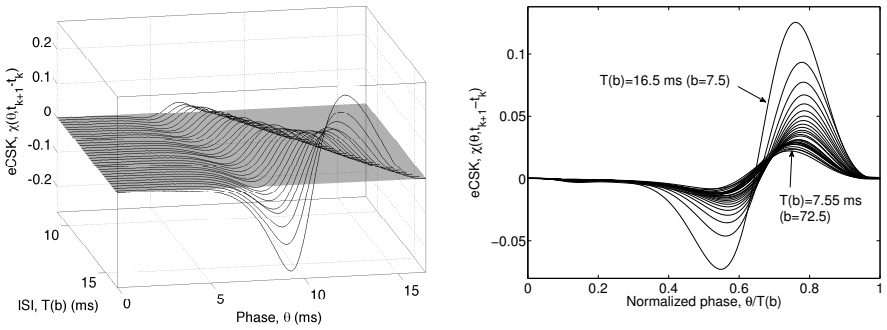


Fig. 6 The estimated conditional sampling kernels (eCSKs) for different inter-spike time intervals. The amplitude of eCSKs are monotonically decreasing with shorter inter-spike time intervals. (A) eCSKs in 3D space for varying inter-spike time intervals (ISIs). (B) 2D projection of (A) along the ISI axis.

$$\chi(t-t_k, t_{k+1}-t_k) = \psi(t-t_k, \hat{b}_k) / \left(\hat{b}_k \int_{t_k}^{t_{k+1}} \psi(s-t_k, \hat{b}_k) ds \right)$$

with $\hat{b}_k = T^{-1}(t_{k+1}-t_k)$ *for* $t \in [t_k, t_{k+1}]$ *and zero, otherwise.*

eCSKs are plotted in Fig. 6 for different inter-spike time intervals. Each eCSK is a scaled version of a conditional PRC with a variable scale factor, and this makes the amplitude ratio of the lowest eCSK ($b = 72.5\mu\text{A/cm}^2$ and $T(b) = 7.55$ms) to the tallest eCSK ($b = 7.5\mu\text{A/cm}^2$ and $T(b) = 16.5$ms) smaller than that of conditional PRCs.

An implementation of the encoding and decoding algorithms described in this section is shown in block diagram form in Fig. 7.

4 Examples

The recovery algorithms for stimuli encoded with a Hodgkin-Huxley neuron were extensively tested. All three proposed algorithms were evaluated with the same set of randomly generated stimuli by visually comparing the sample paths of - as well as by measuring the root-mean-square error between - the recovered signal and the original signal.

The input stimulus is a bandlimited signal generated using equation (6) with $\Omega = 2\pi \cdot 20$ Hz and $b = 25\mu\text{A/cm}^2$. In order to test the recovery algorithms with various waveform patterns, the samples $(u(k\pi/\Omega)), k \in \mathbf{Z}$, were generated from a uniform distribution between $[-1,1]$. Thus, the parameter c in equation (6) determines the magnitude of the perturbation.

Our simulation results are depicted in Figure 8. In each plot in Figure 8 both the original stimulus and the recovered stimulus as a function of time are shown. The

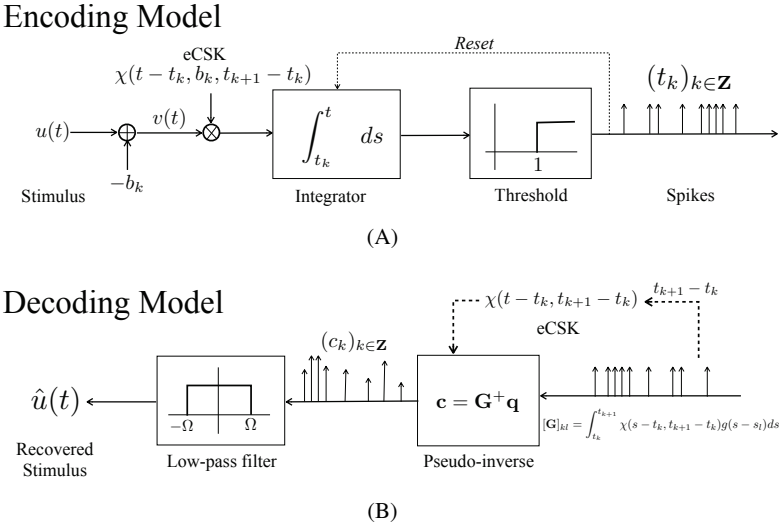


Fig. 7 Model of stimulus encoding/decoding using the project-integrate-and-fire neuron with unknown conditional PRC. (A) The stimulus $u(t)$ is weighted by the estimated conditional sampling kernel (eCSK) determined by an estimate of the base current. (B) The estimated conditional sampling kernel (eCSK) is used to construct the $\mathbf{G}$ matrix. The stimulus is recovered as the response of a low-pass filter to the weighted spike train.

three rows show the sample paths of the recovery results of the three algorithms, i.e., for stimulus encoding with a known PRC, a known conditional PRC and an unknown conditional PRC. In the left column a weak stimulus with $c = 0.5\mu\text{A/cm}^2$ was employed. In the right column a strong stimulus was used with $c = 15\mu\text{A/cm}^2$.

The three algorithms show different levels of error recovery for the weak and the strong stimuli, respectively. Algorithm 1, based on Theorem 1, can recover a weak perturbation with a low recovery error (Figure 8(A)). For strong perturbations, however, the error is markedly larger (Figure 8(B)). Algorithm 3, based on Theorem 3, exhibits a low recovery error for strong stimuli (Figure 8(F)) but fails to recover weak perturbations (Figure 8(E)). Finally Algorithm 2, based on Theorem 2, displays median performance for both stimuli.

The differences in performance of the three recovery algorithms can be traced back to the assumptions about the three sampling kernels: known PRC, known conditional PRC, and unknown conditional PRC, respectively. The three sampling kernels encoding the strong stimuli in Figure 8(B)(D)(F) are shown in Figure 9. The dashed vertical lines in Figure 9 designate the spike times of the Hodgkin-Huxley neuron. Note that the sampling kernels act independently during each inter-spike time interval. In Figure 9 they were stitched together at spike times and, consequently, appear as continuous-time signals. The base current $b = 25\mu\text{A/cm}^2$ was chosen throughout. While the overall pattern is quite similar, minute differences are

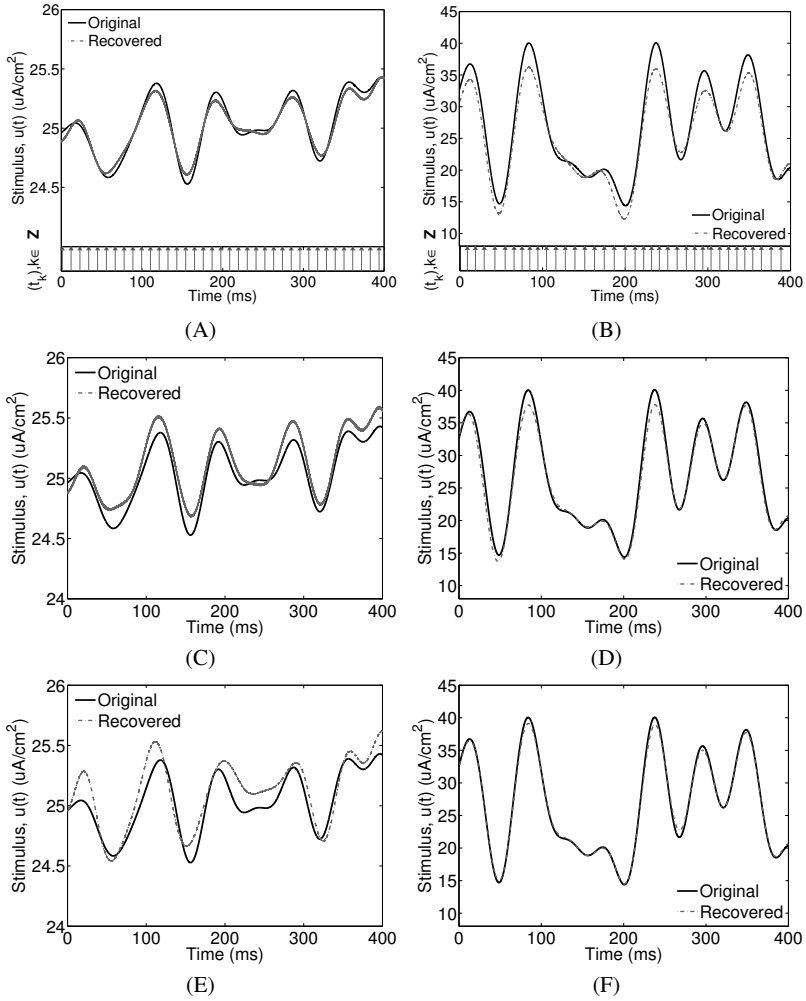


Fig. 8 Recovery of weak (left column) and strong (right column) stimuli. (A)(B) Recovery of a weak stimulus ($c = 0.5\mu\text{A/cm}^2$) and of a strong stimulus ($c = 15\mu\text{A/cm}^2$), respectively, encoded with a known PRC. (C)(D) Recovery of the weak and the strong stimulus, respectively, encoded with known conditional PRC. (E)(F) Recovery of the weak and the strong stimulus, respectively, encoded with an unknown conditional PRC.

observed in Figure 9(B)(C) (marked by arrows). The estimation of the base current level for a weak and a strong stimulus is shown in Fig. 11 and 10, respectively.

An extensive simulation of the three recovery algorithms was carried out in response to randomly generated signals of various magnitudes. Each stimulus was generated as in Figure 8, and c was varied from $0.1\mu\text{A/cm}$ to $25\mu\text{A/cm}^2$ with $b = 25\mu\text{A/cm}^2$ and $\Omega = 2\pi \cdot 20$ Hz. 30 trials were run for each value of c. The root-mean-square (RMS) error defined as

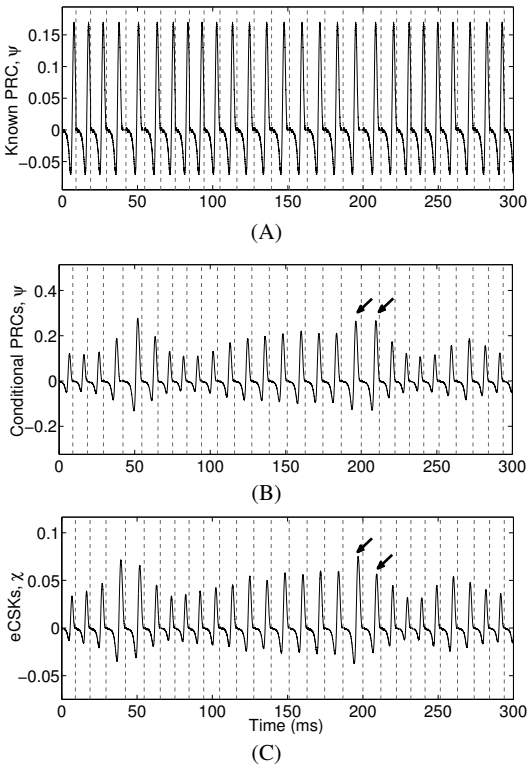
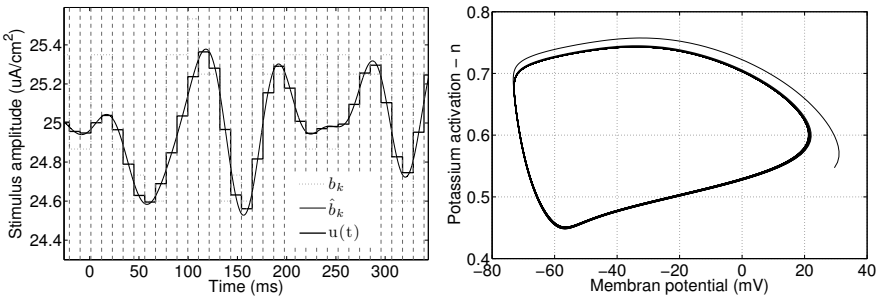


Fig. 9 Train of sampling kernels for the Hodgkin-Huxley neuron. (A) Train of sampling kernels for known PRC. (B) Train of sampling kernels for known conditional PRC. (C) Train of estimated conditional sampling kernels for unknown conditional PRC. The envelope of the sampling kernels in (B) and (C) are similar, but small differences are observed as well (arrows).

$$\sqrt{\frac{1}{S}\int_0^S [u(s)-\hat{u}(s)]^2\,ds},$$

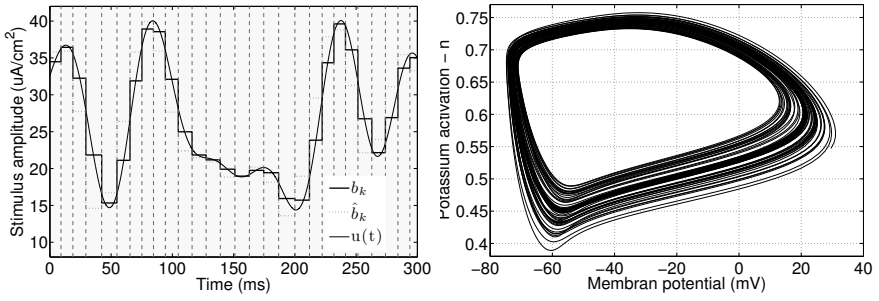
where $\hat{u}(t)$ is the recovered stimulus and S is the time interval of the simulation (400 ms), was used to characterize the distance between the stimulus and its estimate. In Figure 12, the recovery errors of three different algorithms are compared with respect to the perturbation magnitude c. When the perturbation magnitude is weak (Figure 12(B)), recovery Algorithm 1 and Algorithm 2 exhibit lower RMS than Algorithm 3. In the derivation of the Algorithm 3 in section 3.3, we essentially assumed that $T(b_k)-t_{k+1}-t_k=0$. This assumption does not hold for weak perturbations (say, below $2\mu\mathrm{A/cm}^2$ in Figure 12(B)).

The RMS of Algorithm 1 increases rapidly with the perturbation magnitude and the associated changes of the phase portrait of the Hodgkin-Huxley neuron. The



(A) (B)

Fig. 10 The bandlimited stimulus and the base current for a weak stimulus. (A) The base current and the stimulus $u(t)$ are in a close range. The estimated base current does not faithfully estimate the base current. (B) Characterization of the stimulus in the phase plane. A small number of bundled limit cycles corresponds to a weak stimulus.



(A) (B)

Fig. 11 The bandlimited stimulus and the base current for a strong stimulus. (A) The base current, the estimated base current and the stimulus $u(t)$ are in a close range. (B) Characterization of the stimulus in the phase plane. A large number of limit cycles corresponds to a strong stimulus.

performance of Algorithm 2 matches the one of Algorithm 1 for weak stimuli and the one of Algorithm 3 for strong stimuli.

5 Discussion

We investigated the recovery of stimuli encoded with a Hodgkin-Huxley neuron. The daunting complexity of the Hodgkin-Huxley equations makes this task particularly challenging. A highly effective strategy pioneered in (Lazar, 2007, 2010) calls for reducing the Hodgkin-Huxley neuron to a one-dimensional PIF neuron. The latter projects the input stimulus onto the phase response curve and compares the result with a threshold value. The PIF model is I/O-equivalent with the Hodgkin-Huxley

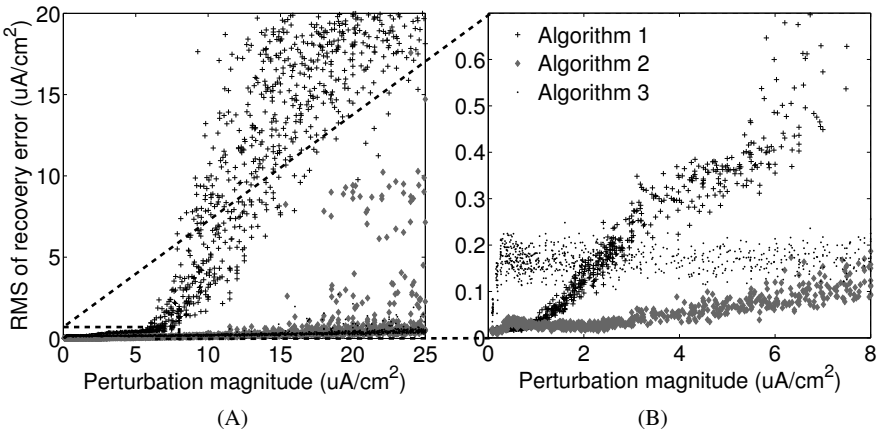


Fig. 12 Recovery errors of the three different algorithms. (B) zooms into the dashed box in (A). Algorithm 1 rapidly increases its recovery error as the perturbation magnitude increases; it shows the lowest level of error when the perturbation magnitude is below 1 $\mu A/cm^2$. Algorithm 2 shows the lowest level of errors on average across the weak/strong perturbation range. The recovery results with Algorithm 3 show the highest error level for weak perturbation levels; beyond 8 $\mu A/cm^2$ the least level of error is recorded.

neuron; given the same input and the same initial condition, the two neurons generate, to the first order, the same spike trains.

Using the reduced PIF model, we presented three different recovery algorithms for stimuli encoded with a Hodgkin-Huxley neuron. All algorithms reconstruct the stimuli from the neuron's output spike train.

The first algorithm is based on the assumption that the Hodgkin-Huxley neuron has a known PRC. It performs well for weak stimuli. However, if the base level of the stimulus changes, the recovery algorithm rapidly increases its error. The second algorithm assumes that the PRC is known on each inter-spike time interval. Since the base level is known, the sampling kernels of the t-transform that best describe the Hodgkin-Huxley neuron are conditionally known on each inter-spike time interval. The algorithm recovers both the strong drift of the base level and the weak perturbation of the stimulus. Finally, the third algorithm operates under the assumption that the conditional PRC is unknown and has to be estimated. We established an estimate of the conditional PRC based upon the readily observable inter-spike time interval. The third recovery algorithm shows excellent performance for strong base level signals.

The proposed algorithms are easy to implement and are very flexible; they only require the inversion of a matrix. These algorithms can also be used for the recovery of stimuli encoded by a large class of spiking neuron models in the limit cycle region. How the dendritic tree of a postsynaptic neuron recovers the information contained in the incoming spike sequence remains an open problem.

Appendix

The Hodgkin-Huxley Equations

The Hodgkin-Huxley equations are as follows. They differ from the original formulation in that the membrane potential is shifted by -65mV.

$$
\begin{aligned}
C\frac{dV}{dt} &= -\left[\bar{g}_{Na}m^3h(V-E_{Na}) + \bar{g}_K n^4(V-E_K) + \bar{g}_L(V-E_L)\right] + I_{ext} \\
\frac{dn}{dt} &= \alpha_n(V)(1-n) - \beta_n(V)n \\
\frac{dh}{dt} &= \alpha_h(V)(1-h) - \beta_h(V)h \\
\frac{dm}{dt} &= \alpha_m(V)(1-m) - \beta_m(V)m,
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_n(V) &= -0.01(V+55)/\left(e^{-(V+55)/10}-1\right) \\
\beta_n(V) &= 0.125e^{-(V+65)/80} \\
\alpha_h(V) &= 0.07e^{-(V+65)/20} \\
\beta_h(V) &= 1/\left(e^{-(V+35)/10)}+1\right) \\
\alpha_m(V) &= -0.1(V+40)/\left(e^{-(V+40)/10}-1\right) \\
\beta_m(V) &= 4e^{-(V+65)/18}
\end{aligned}
$$

and

$\bar{g}_{Na}$ = 120mS/cm^2 $\bar{g}_K$ = 36 mS/cm^2 $\bar{g}_L$ = 0.3 mS/cm^2
E_{Na} = 50 mV E_K = -77 mV E_L = -54.387 mV
$C = 1$ uF/cm^2.

References

Arcas, B. A. y, Fairhall, A. L., & Bialek, W. (2003). Computation in a single neuron: Hodgkin and Huxley revisited. *Neural Computation*, *15*(8), 1715–49.

Bialek, W., & Rieke, F. (1991). Reading a neural code. *Science*, *252*(5014), 1854–1857.

Dayan, P., & Abbott, L. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. MIT Press.

Ermentrout, B. (1996). Type I membranes, phase resetting curves, and synchrony. *Neural Computation*, *8*(5), 979–1001.

Hastings, J., & Sweeney, B. (1958). A Persistent Diurnal Rhythm of Luminescence in Gonyaulax Polyedra. *The Biological Bulletin*, *115*(3), 440–458.

Hodgkin, A. L., & Huxley, A. F. (1952). A Quantitative Description of Membrane Current and its Application to Conduction and Excitation. *J. Physiol*, *117*, 500–557.

Hoppensteadt, F., & Izhikevich, E. (1997). *Weakly Connected Neural Networks*. Springer.

Izhikevich, E. M. (2007). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press.

Kim, A. J., Lazar, A. A., & Slutskiy, Y. B. (2010, Aug). System identification of drosophila olfactory sensory neurons. *Journal of Computational Neuroscience*. (published online doi:10.1007/s10827-010- 0265-0)

Kuramoto, R. (1984). *Chemical Oscillations, Waves, and Turbulence*. Springer-Verlag.

Lazar, A. A. (2004). Time Encoding with an Integrate-and-Fire Neuron with a Refractory Period. *Neurocomputing*, 58–60.

Lazar, A. A. (2007). Recovery of Stimuli Encoded with Hodgkin-Huxley Neurons. In *Cosyne'07* (p. III-94).

Lazar, A. A. (2010). Population encoding with hodgkin-huxley neurons. *IEEE Transactions on Information Theory*, *56*(2), 821-837. Special Issue on Molecular Biology and Neuroscience.

Lazar, A. A., & Pnevmatikakis, E. A. (2008). Faithful Representation of Stimuli with a Population of Integrate-and-Fire Neurons. *Neural Computation*, *20*(11).

Lazar, A. A., & Tóth, L. T. (2004). Perfect Recovery and Sensitivity Analysis of Time Encoded Bandlimited Signals. *Circuits and Systems I: Regular Paper, IEEE Transactions on*, *51*(10), 2060–2073.

Medina, J. F., & Lisberger, S. G. (2007). Variation, Signal, and Noise in Cerebellar Sensory-Motor Processing for Smooth-Pursuit Eye Movements. *Journal of Neuroscience*, *27*(25), 6832.

Naka, K. I., Chan, R. Y., & Yasui, S. (1979). Adaptation in catfish retina. *Journal of Neurophysiology*, *42*(2), 441–54.

Ota, K., Omori, T., & Aonishi, T. (2009). MAP Estimation Algorithm for Phase Response Curves based on Analysis of the Observation Process. *Journal of Computational Neuroscience*, *26*(2), 185–202.

Penrose, R. (1955). A Generalized Inverse for Matrices. In *Proceedings of the Cambridge Philosophical Society* (Vol. 51, pp. 406–413).

Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967). Neuronal spike trains and stochastic point processes. I. The single spike train. *Biophysical Journal*, *7*(4), 391–418.

Schuetze, S. M. (1983). The discovery of the action potential. *Trends in Neurosciences*, *6*, 164–168.

Winfree, A. (1967). Biological Rhythms and the Behavior of Populations of Coupled Oscillators. *Journal of Theoretical Biology*, *16*(1), 15–42.

Yoshimura, K., & Arai, K. (2008, Jan). Phase Reduction of Stochastic Limit Cycle Oscillators. *Physical Review Letters*.