LETTER — Communicated by Terrence Sejnowski

The Spike-Triggered Average of the Integrate-and-Fire Cell Driven by Gaussian White Noise

Liam Paninski
liam@stat.columbia.edu
Department of Statistics, New York University, New York, NY 10027, U.S.A.

We compute the exact spike-triggered average (STA) of the voltage for the nonleaky integrate-and-fire (IF) cell in continuous time, driven by gaussian white noise. The computation is based on techniques from the theory of renewal processes and continuous-time hidden Markov processes (e.g., the backward and forward Fokker-Planck partial differential equations associated with first-passage time densities). From the STA voltage, it is straightforward to derive the STA input current. The theory also gives an explicit asymptotic approximation for the STA of the leaky IF cell, valid in the low-noise regime $\sigma \to 0$. We consider both the STA and the conditional average voltage given an observed spike "doublet" event, that is, two spikes separated by some fixed period of silence. In each case, we find that the STA as a function of time-preceding-spike, τ, has a square root singularity as τ approaches zero from below and scales linearly with the scale of injected noise current. We close by briefly examining the discrete-time case, where similar phenomena are observed.

1 Introduction

The spike-triggered average (STA) (de Boer & Kuyper, 1968; Bryant & Segundo, 1976; Chichilnisky, 2001) is an easily measured experimental quantity defined as the conditional average stimulus to a cell, given that the cell has just emitted an action potential. Thus, this average quantity summarizes, in a sense, what stimulus led to a spike, and as such has taken on some importance in studies of neural coding (Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997; Simoncelli, Paninski, Pillow, & Schwartz, 2004) and Hebbian models of short-term synaptic plasticity (Dayan & Abbott, 2001).

Computing this quantity for model neurons, in turn, has led to some insight into the coding properties of these models. For example, for the linear-nonlinear-Poisson (LNP) cascade model (Simoncelli et al., 2004) the STA turns out to be closely associated with the linear filter of the cell (the "L" stage of the model) (Bussgang, 1952; Chichilnisky, 2001; Paninski, 2003, 2004), allowing for straightforward estimation of the model parameters via simple STA-based computations. For the linear-nonlinear model with multiplicative history effects (Berry & Meister, 1998), the STA is perturbed in an easily characterized fashion by these history terms (Paninski, 2003; Aguera y Arcas & Fairhall, 2003).

Here we consider the linear integrate-and-fire (IF) neuron driven by white gaussian noise of scale σ and mean μ; in this model, the voltage V satisfies the stochastic differential equation,

$$dV_t = \mu dt + I_t, \tag{1.1}$$

with I_t a white noise process of scale σ (that is, $I_t = \sigma dB_t$, with B_t a standard Brownian motion); the cell spikes and V is reset to some value V_r upon each threshold crossing, $V(t) = V_{th}$, where $V_{th} > V_r$. This is the most common base model for stochastic neuronal responses and has proven useful in a wide variety of contexts (Koch, 1999; Gerstner & Kistler, 2002; Paninski, Lau, & Reyes, 2003; Pillow, Paninski, Uzzell, Simoncelli, & Chichilnisky, 2005). Thus, it is worthwhile to examine its properties in analytical detail where possible.

Computing the STA for the integrate-and-fire cell has proven somewhat more complex than in the simpler LNP case described above, basically because the IF cell has a more complex history dependence. Some preliminary approximate analysis of this problem appeared in Gerstner (2001) and Kanev, Wenning, and Obermayer (2004). More recently, Badel, Richardson, and Gerstner (2005) presented some exact asymptotic results based in part on large-deviations approximations (Freidlin & Wentzell, 1984; Kautz, 1988); see also Paninski, in press, for a recent application to the IF model) and in part on the theory of partial differential (Fokker-Planck) equations associated with Brownian motion.

Here we show that it is possible to give relatively simple exact (nonasymptotic) formulas for the STA of the nonleaky IF cell. In addition, these nonasymptotic results lead fairly naturally to exact asymptotic results that hold slightly more generally. Our results thus complement those of Badel et al. (2005), who considered more general versions of the basic IF model but give only asymptotic results.

This article is organized as follows. Section 2 contains our main result: here we explicitly calculate the conditional average voltage and input current of the nonleaky IF cell in continuous time given an observed spike "doublet" event, that is, two spikes separated by some fixed period of silence. This result has a natural extension to an approximate solution for the leaky case (see section 2.1); this approximation may be shown to be exact in the small noise limit $\sigma \to 0$, via comparison with the large-deviation results (Paninski, in press; Badel et al., 2005). In section 3 we use this doublet-triggered average and some basic renewal theory to compute the exact STA. We discuss an alternate approach in the discrete-time setting in section 4 and summarize a few salient points about the form of the STA, and generalizations to other IF-based models, in section 5.

2 The Doublet-Triggered Average

To compute the STA, we first compute the doublet-triggered density,

$$P(V(t)|s[t_1, t_2] = \{t_1, t_2\}),$$

the probability density of $V(t)$ given that $s[t_1, t_2]$, the observed spike data in the interval $[t_1, t_2]$, consisted of spikes at times t_1 and t_2, with no spikes observed at times $t \in (t_1, t_2)$. Once we have calculated the corresponding doublet-triggered expected voltage,

$$E(V(t)|s[t_1, t_2] = \{t_1, t_2\}),$$

we will use equation 1.1 to recover the doublet-triggered expected current. (Of course, the doublet-triggered voltage density $P(V(t)|s[t_1, t_2] = \{t_1, t_2\})$ is of independent interest; de Ruyter & Bialek, 1988; Paninski, in press; Badel et al., 2005.) We emphasize that $V(t)$ here is assumed to follow the nonleaky noisy dynamics 1.1; we address the leaky case below, in section 2.1. To save on notation, we will fix $t_1 = 0$ in this section (without loss of generality).

The doublet-triggered density $P(V(t)|s[0, t_2] = \{0, t_2\})$ is given by a normalized product of two terms,

$$P(V(t)|s[0, t_2] = \{0, t_2\}) = \frac{1}{Z(t)} P_f(V, t) P_b(V, t), \qquad t \in (0, t_2). \tag{2.1}$$

This follows from the fact that the integrate-and-fire cell is a special case of a continuous-time hidden Markov model (HMM): V acts as the "hidden" variable, which evolves according to Markovian dynamics, and the presence or absence of a spike in time bin t is the observed variable, which is dependent (in this case deterministically) only on $V(t)$ at the single time point t. Thus, we may adapt the existing methods for computing (and sampling from) the conditional density of the hidden variable of an HMM, conditioned on its beginning and end states to this special IF model case. (See, e.g., Rabiner, 1989, and Harvey, 1991, for further detail.)

We start by defining P_f and P_b; this is a simple matter of some manipulations with Bayes' rule. For all $0 < t < t_2$, we have

$$P(V(t)|s[0, t_2] = \{0, t_2\}) = \frac{P(s[0, t_2] = \{0, t_2\}|V(t))P(V(t))}{P(s[0, t_2] = \{0, t_2\})}$$

$$= \frac{P(s[0, t] = \{0\}|V(t))P(s[t, t_2] = \{t_2\}|V(t))P(V(t))}{P(s[0, t_2] = \{0, t_2\})}$$

$$= \frac{P(s[0,t]=\{0\})}{P(s[0,t_2]=\{0,t_2\})} P(V(t)|s[0,t]=\{0\}) P(s[t,t_2]=\{t_2\}|V(t))$$
$$\equiv \frac{1}{Z(t)} P_f(V,t) P_b(V,t),$$

where the second equality reflects the conditional independence of $s([0,t])$ and $s((t,T))$ given $V(t)$ and the last equality is a definition. The ratio $P(s[0,t]=\{0\})/P(s[0,t_2]=\{0,t_2\})$, which is constant in V, may be taken as a normalization factor that ensures the conditional V-probability integrates to one.

It is well known that the forward term solves the Fokker-Planck (forward) equation (Karlin & Taylor, 1981; Tuckwell, 1988; Risken, 1996; Brunel & Hakim, 1999; Haskell, Nykamp, & Tranchina, 2001),

$$\frac{\partial P_f(V,t)}{\partial t} = \frac{\sigma^2}{2}\frac{\partial^2 P_f(V,t)}{\partial V^2} - \mu \frac{\partial P_f(V,t)}{\partial V},$$

with boundary conditions

$$P_f(V_{th},t) = 0 \qquad \forall t \in [0,t_2]$$

and

$$P_f(V,0) = \delta(V - V_r).$$

This may be solved explicitly via the method of images (Daniels, 1982) as

$$P_f(V,t) = \mathcal{N}(V_r + \mu t, \sigma^2 t) - e^{2\mu(V_{th}-V_r)/\sigma^2}\mathcal{N}(2V_{th} - V_r + \mu t, \sigma^2 t),$$
$$V \le V_{th},\ t > 0,$$

where $\mathcal{N}(\mu,\sigma^2) = \mathcal{N}(\mu,\sigma^2)(V)$ denotes the gaussian kernel of mean μ and scale σ.

The backward term, on the other hand, solves the Kolmogorov backward equation (Karlin & Taylor, 1981),

$$\frac{\partial P_b(V,t)}{\partial t} = -\frac{\sigma^2}{2}\frac{\partial^2 P_b(V,t)}{\partial V^2} - \mu \frac{\partial P_b(V,t)}{\partial V},$$

with boundary conditions

$$P_b(V_{th},t) = 0 \qquad \forall t \in [0,t_2]$$

and

$$P_b(V, t_2) = \delta(V - V_{th}).$$

Solving for P_b is slightly more delicate; if we try to solve the backward equation directly, starting at $t = t_2$ and using the method of images to propagate the solution backward in time, all the mass is absorbed at V_{th} immediately. Thus, a more indirect, limiting argument is required. We start with the exact solution to the backward equation started at $V(t_2) = V_{th} - \epsilon$, $\epsilon > 0$,

$$\begin{aligned} P_b^\epsilon(V,t) =& \mathcal{N}(V_{th} - \epsilon - \mu(t_2 - t), \sigma^2(t_2 - t)) \\ & - e^{-2\mu\epsilon/\sigma^2} \mathcal{N}(V_{th} + \epsilon - \mu(t_2 - t), \sigma^2(t_2 - t)), \qquad V \le V_{th}, t < t_2, \end{aligned}$$

and then take the (normalized) limit as $\epsilon \to 0$. Abbreviating $t_2 - t = w$, we have

$$\begin{aligned} P_b^\epsilon(V,t) =& c(t)\Bigg[\exp\left(-\frac{(V_{th} - \epsilon - \mu w - V)^2}{2\sigma^2 w}\right) - e^{-2\mu\epsilon/\sigma^2} \\ & \times \exp\left(-\frac{(V_{th} + \epsilon - \mu w - V)^2}{2\sigma^2 w}\right)\Bigg] \\ =& c(t)\exp\left(-\frac{(V_{th} - \mu w - V)^2}{2\sigma^2 w}\right)\Bigg[\exp\left(\frac{2\epsilon(V_{th} - \mu w - V)}{2\sigma^2 w}\right) \\ & - \exp\left(\frac{-2\mu\epsilon}{\sigma^2} - \frac{2\epsilon(V_{th} - \mu w - V)}{2\sigma^2 w}\right)\Bigg] + o(\epsilon) \\ =& c(t)\exp\left(-\frac{(V_{th} - \mu w - V)^2}{2\sigma^2 w}\right)\Bigg[\frac{2\epsilon(V_{th} - \mu w - V)}{2w} + 2\mu\epsilon \\ & + \frac{2\epsilon(V_{th} - \mu w - V)}{2w}\Bigg] + o(\epsilon) \\ =& c(t)\exp\left(-\frac{(V_{th} - \mu w - V)^2}{2\sigma^2 w}\right)[V_{th} - \mu w - V + \mu w]\epsilon + o(\epsilon) \\ =& c(t)\exp\left(-\frac{(V_{th} - \mu w - V)^2}{2\sigma^2 w}\right)[V_{th} - V]\epsilon + o(\epsilon), \end{aligned}$$

where the $c(t)$ above represents an irrelevant normalization factor, constant in V; thus, we arrive at

$$P_b(V,t) = c(t)[V_{th} - V]\mathcal{N}(V_{th} - \mu(t_2 - t), \sigma^2(t_2 - t)), \qquad V \le V_{th}.$$

Now we may solve for $P(V(t)|s[0, t_2] = \{0, t_2\})$ (and therefore any conditional moment of $V(t)$, such as the mean voltage given the observed data $s[0, t_2]$) simply by plugging the above formulas for P_b and P_f into equation 2.1 and normalizing. This normalization, in turn, requires that we compute integrals of the truncated gaussian distribution,

$$\mathcal{N}^+(m, V)(x) = e(-m/\sqrt{V})^{-1}\mathcal{N}(m, V)(x), \qquad x > 0,$$

with

$$e(x) = \int_x^\infty \mathcal{N}(0, 1)(u)du.$$

We will need the first and second moments of this distribution,

$$\langle x \rangle_{m,V} = m + \frac{\sqrt{V}\sqrt{2/\pi}}{\text{erfcx}(-m/(\sqrt{V}\sqrt{2}))}$$

and

$$\langle x^2 \rangle_{m,V} = m^2 + V + \frac{m\sqrt{V}\sqrt{2/\pi}}{\text{erfcx}(-m/(\sqrt{V}\sqrt{2}))},$$

with erfcx(.) denoting the scaled complementary error function,

$$\text{erfcx}(x) = \frac{2}{\sqrt{\pi}}e^{x^2}\int_x^\infty e^{-t^2}dt.$$

After one last change of variables, we have

$$P(V(t)|s[0, t_2] = \{0, t_2\}) = \frac{V_{th} - V}{z(t)}\left[\mathcal{N}\left(\frac{(V_{th} - V_r)t}{t_2} + V_r, q(t)\right)\right.$$
$$\left. - \mathcal{N}\left(\frac{(V_r - V_{th})t}{t_2} + 2V_{th} - V_r, q(t)\right)\right]$$

for $V < V_{th}$, with the variance term

$$q(t) = \sigma^2(t^{-1} + (t_2 - t)^{-1})^{-1}$$

and the normalization

$$z(t) = e(-y(t)/\sqrt{q(t)}) < x >_{y(t),q(t)} - e(y(t)/\sqrt{q(t)}) < x >_{y(t),q(t)},$$

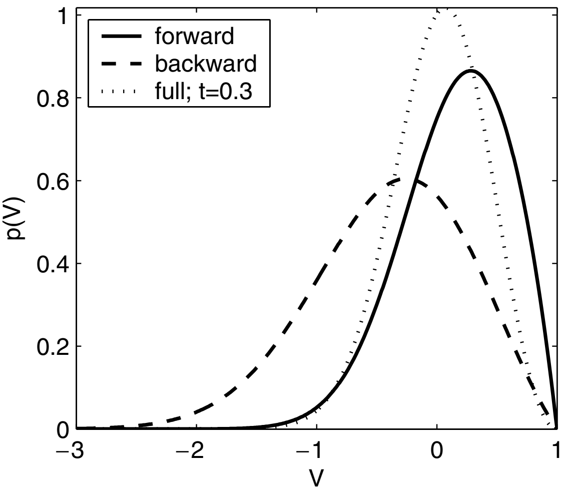

Figure 1: The densities $P_f(V, t)$, $P_b(V, t)$, and $P(V(t)|s[0, t_2] = \{0, t_2\}) = P_f(V, t)P_b(V, t)/z(t)$, for $t = 0.3$; $V_{th} = 1$, $V_r = 0$, $\sigma^2 = 1$, $\mu = 1$, $t_1 = 0$, $t_2 = 1$.

where we have abbreviated

$$y(t) = (V_{th} - V_r)(1 - t/t_2).$$

See Figures 1 and 2 for an illustration of the three densities P_f, P_b, and $P = P_f P_b/z$. Now, finally, we may read off our main result:

$$E(V(t)|s[0, t_2] = \{0, t_2\}) = V_{th} - z(t)^{-1}[e(-y(t)/\sqrt{q(t)}) < x^2 >_{y(t),q(t)}$$
$$-e(y(t)/\sqrt{q(t)}) < x^2 >_{-y(t),q(t)}],$$

for $t \in [0, t_2]$. We will abbreviate this solution as $S_{t_1,t_2}(t)$, that is,

$$S_{t_1,t_2}(t) \equiv E(V(t)|s[t_1, t_2] = \{t_1, t_2\}), \qquad t \in [t_1, t_2]$$

for use below.

At this point, it is worth pausing to note a few salient properties of this doublet-triggered average. First, $S_{t_1,t_2}(t)$ behaves as

$$S_{t_1,t_2}(t) = V_{th} - \sigma\sqrt{8/\pi}\sqrt{t_2 - t} + o(\sqrt{t_2 - t})$$

as $t \nearrow t_2$. (See Figures 2 and 3.) This square-root behavior at t_2 was also noted by Badel et al. (2005).

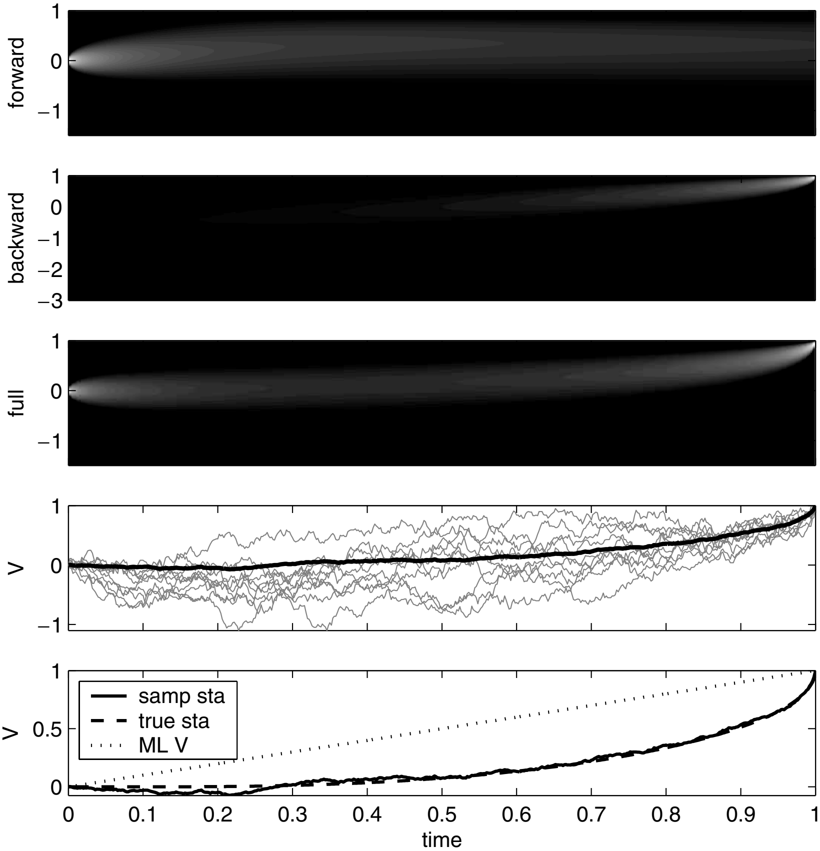

Figure 2: The doublet-triggered average, $S_{0,1}(t)$. Parameters as in Figure 1. Panels 1–3 show the evolution of densities $P_f(t)$, $P_b(t)$, and $P(V(t)|s[0, t_2] = \{0, t_2\})$, for $t \in [0, 1]$; grayscale level indicates height of density. Panel 4 shows some samples (gray traces) from the conditional voltage path distribution given spikes at $t_1 = 0$ and $t_2 = 1$ (see the appendix for a brief description of the exact sampling procedure), with the empirical mean given 100 samples shown in black. The bottom panel shows the most likely path (dotted trace), the analytical doublet-triggered average (dashed), and the empirical doublet-triggered average (solid).

Second, in the low-noise regime, $\sigma \to 0$, the doublet-triggered average converges uniformly to the most likely (ML) voltage path (see Figure 3), which in this case takes the simple linear form (Paninski, in press)

$$V_{ML}(t) = \frac{1}{t_2 - t_1}(V_r(t_2 - t) + V_{th}(t - t_1)), t \in [t_1, t_2].$$

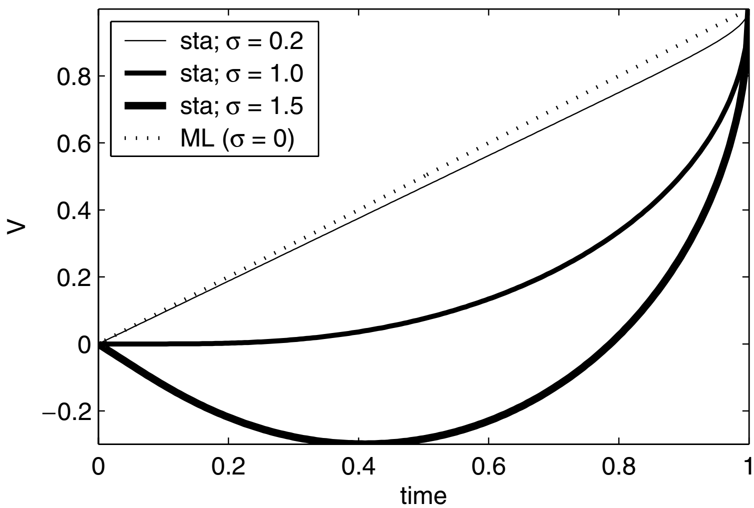

Figure 3: Effects of varying σ on $S_{t_1,t_2}(t)$. Note the convergence to the most likely voltage path (the linear dotted voltage trace) as $\sigma \to 0$. Also note that at sufficiently high noise levels, the doublet-triggered average voltage is actually hyperpolarized below V_r due to the "killing" effect of the absorbing boundary at V_{th}, as described in the text.

This result is consistent with basic results from the theory of large deviations (Freidlin & Wentzell, 1984; Paninski, in press), which indicates that this most likely path will dominate expectations as $\sigma \to 0$.

Thus, the doublet-triggered average may be roughly described as a straight line between V_r and V_{th}, minus a sag of size proportional to σ; this sag, in turn, behaves as a square root as $t \to t_2$, and is due to the fact that voltage paths that happen to be depolarized by noise above threshold are "killed" by the absorbing boundary at V_{th}.

Finally, due to some cancellations, μ does not appear in the above expressions; thus, the doublet-triggered average is (somewhat surprisingly) independent of the mean input current μ. (As we will see in section 3, the STA is dependent on μ; this dependence enters strictly through the μ-dependence of the IF interspike interval density.) To get a better sense of why μ drops out here, it is enlightening to take an alternate approach, based on the Brownian bridge, which is defined (Karlin & Taylor, 1981; Karatzas & Shreve, 1997) as the stochastic process formed by conditioning Brownian motion (with or without drift) on its start and end points. To see the relevance of the Brownian bridge here, we may consider the doublet-triggered density in two steps. First, we condition $V(t)$ to end at $V(t_2) = V_{th}$. This gives us a Brownian bridge started at $(0, V_r)$ and ended at (t_2, V_{th}); note, importantly, that this is the point at which the dependence on μ drops out (since the Brownian bridge has no dependence on the drift μ of the original Brownian motion). Then we condition further, imposing the inequality constraints $V(t) < V_{th}, 0 < t < t_2$, to obtain the doublet-triggered distribution; since μ plays no role in the Brownian bridge process, μ can play no role in this further conditioned process either. (The relevant computations for this conditioned Brownian bridge may be carried out explicitly, using a form of the method of images for the Brownian bridge; the final result is the same, so we omit the details.) This alternate approach also explains the form of the variance term $q(t)$, which exactly matches the variance of the Brownian bridge (Karlin & Taylor, 1981).

With the doublet-triggered average voltage in hand, it is straightforward to derive the doublet-triggered average current from the IF dynamics, equation 1.1; we may simply write

$$E(I(t)|s[t_1, t_2] = \{t_1, t_2\}) = \frac{\partial}{\partial t} S_{t_1,t_2}(t) - \mu.$$

(The careful reader will note that we have been rather blithe about an interchange between a derivative and an expectation here; we will discuss this further in section 4.) This doublet-triggered average current diverges as $(t_2 - t)^{-1/2}$ as $t \to t_2$; this has an interesting effect on the discretized STA for current, which appears not to converge to any physically reasonable limit as the time discretization dt goes to zero (see section 4 below).

2.1 The Leaky Case. The above results suggest a simple approximation for the leaky case,

$$dV_t = (\mu - gV_t)dt + I_t.$$

The forward equation in this case becomes

$$\frac{\partial P_f(V,t)}{\partial t} = \frac{\sigma^2}{2}\frac{\partial^2 P_f(V,t)}{\partial V^2} - \mu\frac{\partial P_f(V,t)}{\partial V} + g\frac{\partial[P_f(V,t)V]}{\partial V},$$

with boundary conditions as above,

$$P_f(V_{th}, t) = 0 \qquad \forall t \in [0, t_2]$$

and

$$P_f(V, 0) = \delta(V - V_r).$$

In this case, P_f satisfies the renewal equation (Karlin & Taylor, 1981; Plesser & Tanaka, 1997; Burkitt & Clark, 1999; Paninski, Haith, Pillow, & Williams, 2005),

$$P_f(V,t) = P_{V_r,0}(V,t) - \int_0^t p_1(s) P_{V_{th},s}(V,t) ds,$$

with $p_1(t)$ denoting the first-passage density,

$$p_1(t) = \frac{\partial}{\partial t}\left(1 - \int_{-\infty}^{V_{th}} P_f(V,t)dV\right) = -\frac{\partial}{\partial t}\int_{-\infty}^{V_{th}} P_f(V,t)dV,$$

and $P_{x,s}(V,t)$ denoting the "free" solution to the forward equation, that is, the (uniquely well-behaved) solution to the forward equation in the absence of the threshold boundary condition, for example,

$$P_{V_r,0}(V,t) = \mathcal{N}\left(V_r + \left(\frac{\mu}{g} - V_r\right)(1 - e^{-gt}), \frac{\sigma^2}{2g}(1 - e^{-2gt})\right).$$

No elementary analytical solution for P_f is available in the leaky case to our knowledge; instead, we simply neglect the second term in the above renewal expression for P_f and approximate

$$P_f(V,t) \approx P_{V_r,0}(V,t).$$

This is accurate as $\sigma \to 0$, if $\mu/g < V_{th}$.

For the backward equation, we replace our analytical solution above with

$$P_b(V,t) \approx c(t)[V_{th} - V]\mathcal{N}\left(\left(V_{th} - \frac{\mu}{g}\right)e^{g(t_2-t)} + \frac{\mu}{g}, \frac{\sigma^2}{2g}(e^{2g(t_2-t)} - 1)\right), \quad V \le V_{th},$$

which again makes use of the free solution to the backward equation.

The corresponding approximation to the doublet-triggered average, formed by plugging the above approximations to P_f and P_b into equation 2.1, is crude but nonetheless asymptotically correct as $\sigma \to 0$: this approximate doublet-triggered average, $\tilde{S}_{0,t_2}(t)$, behaves like

$$\tilde{S}_{0,t_2}^{\sigma\to 0}(t) = \frac{\frac{\sigma^2}{2g}(1 - e^{-2gt})((V_{th} - \frac{\mu}{g})e^{g(t_2-t)} + \frac{\mu}{g}) + \frac{\sigma^2}{2g}(e^{2g(t_2-t)} - 1)\left(V_r + \left(\frac{\mu}{g} - V_r\right)(1 - e^{-gt})\right)}{\frac{\sigma^2}{2g}(1 - e^{-2gt}) + \frac{\sigma^2}{2g}(e^{2g(t_2-t)} - 1)}$$

as $\sigma \to 0$. Some algebra reduces this to

$$\tilde{S}_{0,t_2}^{\sigma \to 0}(t) = \frac{\mu}{g} + ae^{-gt} + be^{gt},$$

with a, b suitably chosen constants in t. Since $\tilde{S}_{0,t_2}^{\sigma \to 0}(t)$ uniquely satisfies the second-order differential equation,

$$\frac{\partial^2}{\partial t^2} \tilde{S}_{0,t_2}^{\sigma \to 0}(t) = -g(-g\tilde{S}_{0,t_2}^{\sigma \to 0}(t) + \mu) = g^2\left(\tilde{S}_{0,t_2}^{\sigma \to 0}(t) - \frac{\mu}{g}\right),$$

with boundary conditions $\tilde{S}_{0,t_2}^{\sigma \to 0}(0) = V_r$ and $\tilde{S}_{0,t_2}^{\sigma \to 0}(t_2) = V_{th}$, $\tilde{S}_{0,t_2}^{\sigma \to 0}(t)$ corresponds exactly to the ML voltage path, which dominates the true doublet-triggered average in the limit $\sigma \to 0$ (Paninski, in press), as discussed above. See Figure 4 for some examples of this approximation.

The doublet-triggered average current is only slightly more complex in this case, since we have to include the effect of the leak, in particular, the mean leak current,

$$E(-gV(t)|s[0, t_2] = \{0, t_2\}) = -gE(V(t)|s[0, t_2] = \{0, t_2\}) = -gS_{t_1,t_2}(t).$$

Our approximation in this case thus takes the form

$$E(I(t)|s[t_1, t_2] = \{t_1, t_2\}) \approx \frac{\partial}{\partial t}\tilde{S}_{t_1,t_2}(t) - \mu + g\tilde{S}_{t_1,t_2}(t).$$

3 The Spike-Triggered Average

Given the doublet-triggered distributions $P(V(t)|s[t_1, t_2])$, it is straightforward to obtain the full STA (and more generally, the full distribution of $V(s)$ at any time s, given a spike at time t).

We make use of the renewal representation of the spike times in this model: note that for $\mu > 0$, the IF model represents a stationary stochastic process. Moreover, since $V(t)$ is strong Markov, the sequence of spike times is a renewal process; as is well known, it is straightforward to calculate the interspike interval density for every order (e.g., via the reflection principle for Brownian motion, coupled with the Girsanov formula; Karatzas & Shreve, 1997). The density of the interval between a given spike and the ith following spike is given by the inverse gaussian density (Seshadri, 1993):

$$p_i(t) = \frac{i(V_{th} - V_r)}{\sqrt{2\pi\sigma^2 t^3}} e^{-(i(V_{th}-V_r)-\mu t)^2/2\sigma^2 t} = p_1 *_{i-1} p_1,$$

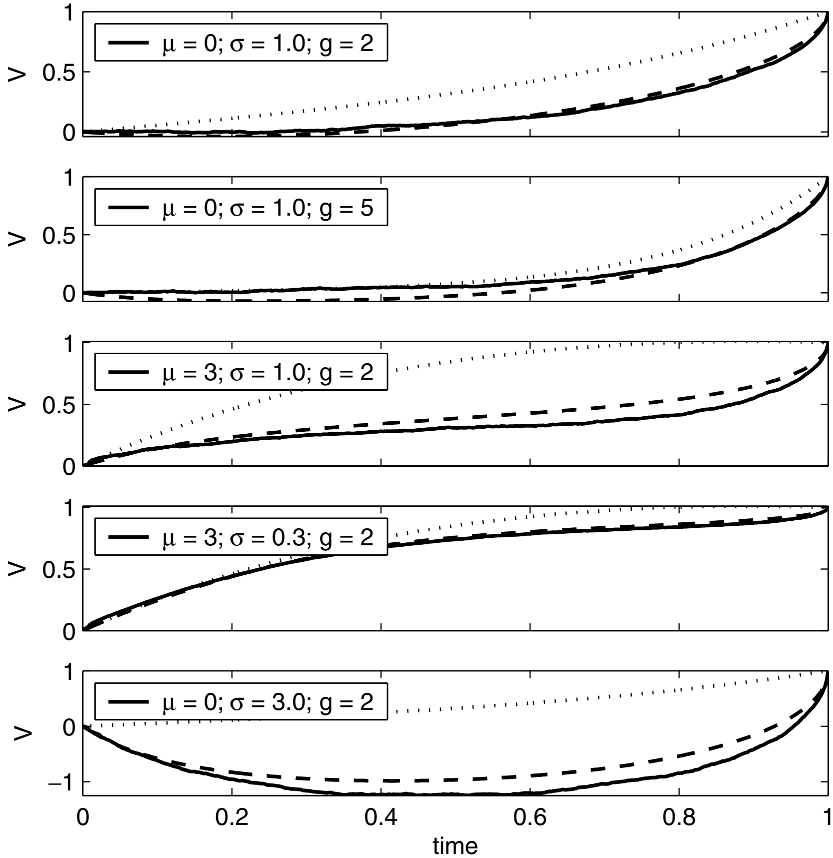

Figure 4: Some approximate doublet-triggered averages in the leaky case. The approximation is fairly accurate in general and is exact as $\sigma \to 0$. The approximation fails in the σ large case, when the free approximation to the forward solution fails (that is, when a nonnegligible amount of probability mass crosses threshold before t_2). Conventions are as in the bottom panel of Figure 2: dashed curves are analytical approximations to the doublet-triggered averages, dotted curves are ML paths, and solid curves are empirical doublet-triggered averages (based on 1000 samples from true conditional voltage distribution given $s[0, t_2]$).

where $*_i$ denotes the i-fold convolution.

Define

$$S_1^+(t) = \int_t^\infty p_1(s) S_{0,s}(t) ds, \qquad t > 0,$$

and

$$S_1^-(t) = \int_{-\infty}^{-t} p_1(-s) S_{s,0}(t) ds, \qquad t > 0.$$

These are the doublet-triggered averages S_{0,t_2} averaged over the next and previous interspike intervals, respectively. Denote the firing rate function (Rudd & Brown, 1997) of the IF cell as

$$f(t) = \sum_{i=1}^{\infty} p_i(t);$$

this is the expected firing rate of the cell at time t given a spike at time 0. Then the spike-triggered average voltage for positive times t is given by the convolution

$$STA(t) = (f(t) + \delta(t)) * S_1^+(t) = S_1^+(t) + \int_0^s f(t-s) S_1^+(s) ds, \qquad t > 0,$$

and similarly for negative times,

$$STA(-t) = (f(t) + \delta(t)) * S_1^-(t) = S_1^-(t) + \int_0^s f(t-s) S_1^-(s) ds, \qquad t > 0.$$

The spike-triggered average current and distributions $P(V(t)|s[0] = \{0\})$ follow similarly. See Figure 5 for a few examples of the STA.

The leaky case follows exactly the same route, with the exception (again) that no explicit analytical solution is known for $p_1(t)$ or $f(t)$ in the leaky case. But the spikes from the leaky IF (LIF) cell are still a renewal process, and the STA can still be written in the convolution form given above.

It is worth pointing out that a simpler approach suffices for positive times. In this case, we can form the spike-triggered distributions $P(V(t)|s[0] = 0)$ directly, without going through the intermediate step of computing the doublet-triggered distributions. By the usual renewal argument, we have that

$$P(V(t)|s[0] = 0) = P_f(V, t) + \int_0^t P_f(V, t-s) f(s) ds, \qquad t > 0,$$

from which we may read off the expectation to obtain the STA. From this, it is easy to see that the STA may be approximated for small, positive times t with the simple linear form

$$STA(t) \approx V_r + (\mu - g V_r)t + o(t)$$

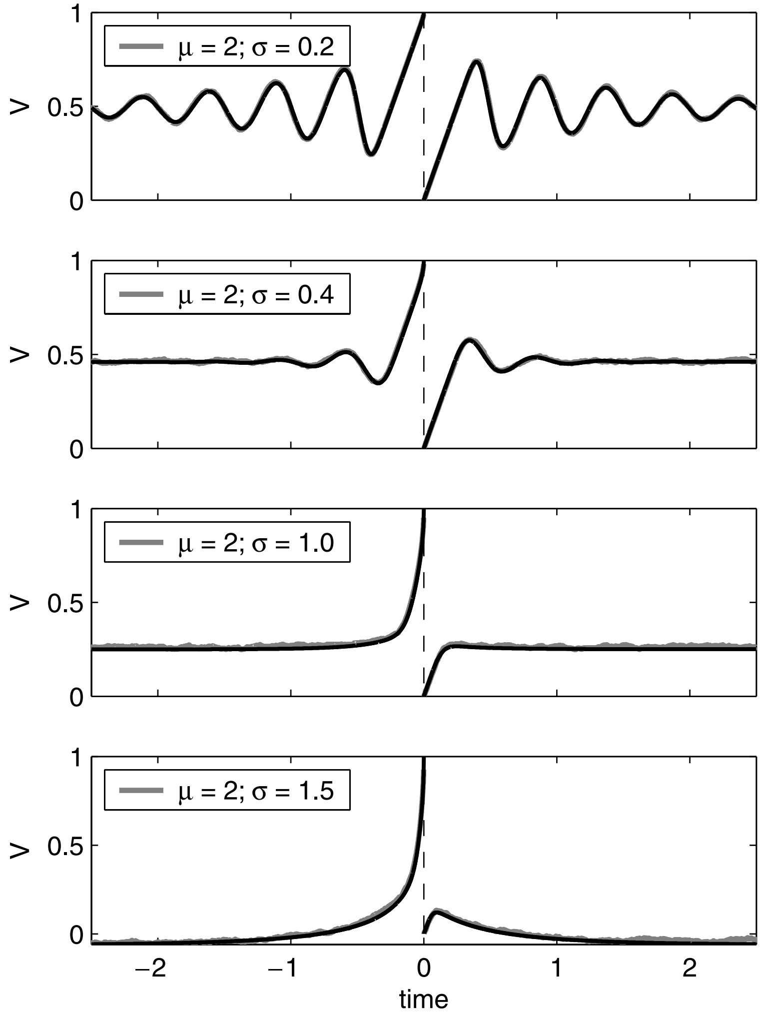

Figure 5: A few examples of the spike-triggered average voltage, for different values of σ (in each case, the leak $g = 0$). Black trace is analytical STA; gray trace (mostly obscured by the black trace) is empirical STA, given 2000 seconds of simulated data. In the low-noise regime $\sigma \to 0$ (top panel), an oscillatory ringing is visible, at a frequency approaching the firing rate of the cell in the absence of noise $\mu/(V_{th} - V_r)$, 2 Hz in this case; the square root singularity at $t \to 0^-$ becomes more visible as σ increases.

(in contrast with the square root singularity as t approaches zero from the left).

4 Discrete-Time Approach

In the previous sections, we computed the spike-triggered average voltage in continuous time. Deriving the spike-triggered average current from the average voltage may be done formally by a simple interchange of the derivative in equation 1.1 and the expectation taken when forming the STA. However, justifying this interchange rigorously runs into the usual issues associated with white noise in continuous time (specifically, the fact that white noise is only defined as a measure on a space of generalized functions; Hida, 1980) and would take us slightly afield. Instead, we discuss here a more direct, rigorous calculation of the spike-triggered average current in discrete time and point out some similarities to the formal continuous-time approach taken above.

We begin by writing the LIF model in discrete time:

$$V(t+dt) = V(t) + (\mu - gV)dt + I_t,$$

where I_t, the input current, is discrete gaussian white noise with mean zero and scale $\sigma\sqrt{dt}$, and the voltage is reset to V_r at each threshold crossing. Note that the $\sqrt{dt}$ scaling on the input noise current ensures the existence of a continuum limit of the above process, as $dt \to 0$; this limit process is equivalent to an Ornstein-Uhlenbeck process that resets at each crossing of the threshold V_{th}. We will focus on this $dt \to 0$ limit below.

In the following, we suppress the argument of $I(t - \tau dt)$ and abbreviate the event $s[t] = \{t\}$ as s. Thus, we write the spike-triggered average current as

$$\begin{aligned} E(I(t-\tau dt)|s[t] = \{t\}) &= \int I P(I|s) dI \\ &= \int I \frac{P(s|I)}{P(s)} P(I) dI. \end{aligned}$$

We analyze each term in the above expression in turn.

First, by definition,

$$P(I) = \mathcal{N}(0, \sigma^2 dt).$$

Next,

$$P(s) = F dt + o(dt),$$

where F denotes the invariant (steady-state) firing rate of the cell. F, in turn, can be computed as follows (Karlin & Taylor, 1981; Brunel & Hakim, 1999; Haskell, Nykamp, & Tranchina, 2001; Paninski et al., 2003):

$$F = -\frac{\sigma^2}{2} \frac{\partial P_\infty(V)}{\partial V}\bigg|_{V=V_{th}},$$

where $P_\infty(V)$ is the invariant density on voltage. It turns out that we will not need to know anything about this invariant distribution beyond the fact that it exists uniquely, is differentiable from below at V_{th}, and is zero above V_{th}. Somewhat surprisingly, we will not even need to compute F.

The hard part is $P(s|I)$. We condition on the voltage at time t, as follows:

$$\begin{aligned} P(s|I) &= \int \frac{P(s, I, V)}{P(I)} dV \\ &= \int \frac{P(s|I, V)P(I, V)}{P(I)} dV \\ &= \int P(s|I, V)P(V|I)dV \\ &= \int P(s|I, V)P(V)dV, \end{aligned}$$

where we have used the independence of the current and the voltage at a given time step in the final line.

To compute $p(s|I, V)$, we need to introduce some limiting arguments—the fact that $dt \to 0$ will allow us to compute $p(s|I, V)$ exactly asymptotically, and this does not appear to be possible for arbitrary dt. First, we write out the definition more carefully:

$$\begin{aligned} P(s|I, V) &= P(V(t) > V_{th}|V(t-(\tau+1)dt), I(t-\tau dt)) \\ &= P(V(t) > V_{th}|V(t-\tau dt) = V(t-(\tau+1)dt) + (\mu - gV)dt + I). \end{aligned}$$

To put it more simply, this is the probability that the stochastic process V will cross the boundary V_{th} on the τth time step; clearly, any such crossing must be from below, given the definition of the LIF model.

The first limiting argument is simple: as $dt \to 0$, the probability that more than one such crossing will occur in the interval $(t - \tau dt, t)$ decreases exponentially. Thus, we have a relatively simple gaussian first passage time problem:

$$P(s|I, V) = q_{D-OU}(\tau, V_{th}, V + I, \sigma\sqrt{dt}, g, \mu/g, dt) + O(\sqrt{dt}),$$

where q_{D-OU} is the probability that a discrete Ornstein-Uhlenbeck process, starting at $V + I$, with leak parameter g, equilibrium potential μ/g, time step dt, and scale $\sigma\sqrt{dt}$, will first cross the threshold V_{th} at τ time steps. It will be helpful to rescale q_{D-OU} as follows:

$$q_{D-OU}(\tau, V_{th}, V + I, \sigma\sqrt{dt}, g, \mu/g, dt)$$
$$= q_{D-OU}(\tau, 0, (V + I - V_{th})/\sigma\sqrt{dt}, 1, gdt, \mu/g, 1).$$

Putting all the pieces together, we have

$$\begin{aligned} STA(\tau) &= \frac{1}{P(s)} \iint I P(I)P(V)P(s|I, V)dIdV \\ &= -\frac{2}{\sigma^2 dt \dot{P}_\infty(V_{th})} \iint \frac{I}{\sigma\sqrt{dt}} G\left(\frac{I}{\sigma\sqrt{dt}}\right) P_\infty(V) \\ &\quad \times(q_{D-OU}(\tau, 0, (V+I - V_{th})/\sigma\sqrt{dt}, 1, gdt, \mu/g, 1) + O(\sqrt{dt}))dIdV, \end{aligned}$$

where G denotes the standard gaussian density. Now we change variables:

$$a = I/\sigma\sqrt{dt},$$

and

$$b = (V_{th} - V)/(\sigma\sqrt{dt}),$$

to simplify our integral to

$$\begin{aligned} &\frac{2}{\dot{P}_\infty(V_{th})} \iint aG(a)P_\infty(V_{th} - b\sigma\sqrt{dt}) \\ &\quad \times(q_{D-OU}(\tau, 0, a - b, 1, gdt, \mu/g, 1) + O(\sqrt{dt}))dadb. \end{aligned}$$

Finally, by L'Hopital and a simple dominated convergence argument and the fact that the Ornstein-Uhlenbeck mean and covariance matrix converge to that of discrete Brownian motion in this limit, we have

$$STA(\tau) = 2\sigma\sqrt{dt} \int aG(a)da \int_0^\infty bq_{DB}(\tau, b - a)db + O(dt), \qquad dt \to 0,$$

with q_{DB} denoting the probability that a standard discrete Brownian motion (that is, a cumulative sum of independently and identically standard normal variables) will first cross the threshold $b - a$ at time τ.

Unfortunately, q_{DB} does not seem to have a simple analytical expression, although we can compute this quantity fairly explicitly for small τ, and the large τ asymptotics can be computed by appealing to known results on the corresponding quantity for Brownian motion. For example, we can compute

$$q_{DB}(1, u) = \Theta(-u)e(u);$$

in general, q_{DB} is given by a similar τ-dimensional gaussian integral over an orthant (Paninski, Pillow, & Simoncelli, 2004), or alternately by a repeated convolution of error functions.

As discussed in the previous section, the corresponding crossing probabilities for continuous-time Brownian motion can be computed exactly:

$$q_B(\tau, u) = \int_{\tau}^{\tau+1} p_1^u(t)dt = 2\int_{(\tau+1)^{-1/2}u}^{\tau^{-1/2}u} G(x)dx,$$

with $p_1^u(t)$ denoting the first passage time of a standard Brownian motion to the threshold u, and it is fairly easy to show that $q_B \sim q_{DB}$ as $\tau \to \infty$. Since

$$q_B(\tau, u) \approx G(\tau^{-1/2}u)\tau^{-3/2}u$$

for large τ, we have

$$STA(\tau) \approx 2\sigma\sqrt{dt}\tau^{-3/2}\int da \int_0^{\infty} ab(b-a)G(a)G(\tau^{-1/2}(b-a))db.$$

Another change of variable demonstrates that the spike-triggered average current behaves like

$$STA(\tau) = A\sigma\sqrt{dt}\tau^{-1/2} + o(\tau^{-1/2})$$

(with the prefactor A given by a gaussian polynomial integral over a half-space) as $\tau \to \infty$, locally at the spike time (i.e., for $\tau dt \to 0$). This matches the result established by the continuous-time argument in the preceding section. See Figure 6 for an illustration of the singular behavior of the discrete STA as $\tau \to 0^-$ and of the dependence of the STA on dt.

5 Conclusions and Extensions

We can derive three somewhat surprising conclusions from the above results. First, the spike-triggered average current of the LIF cell in discrete time does not, in fact, have a continuum limit as $dt \to 0$ in the usual sense. We might have expected that the STA would live on a timescale of $\sim 1/g$—roughly, that the cell would integrate over about a membrane

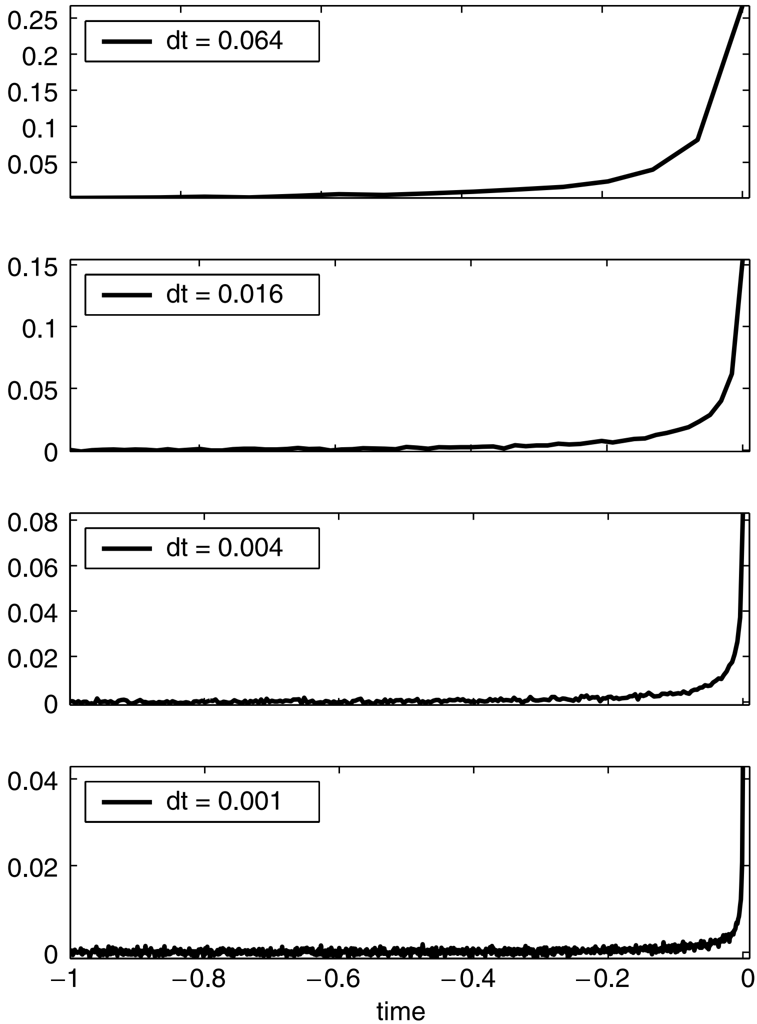

Figure 6: The discrete spike-triggered average current, as a function of dt. Note that the STA becomes sharper with decreasing dt, with a horizontal scale proportional to dt and a vertical scale proportional to $\sqrt{dt}$. Parameters: $\sigma = 1, \mu = 2, g = 0$.

time constant's worth of input before "deciding" whether to spike. In fact, the STA is effectively supported on a dt timescale, and therefore the width of the spike-triggered average vanishes as $dt \to 0$. This illustrates the danger of thinking of the STA too glibly as the "linear prefilter" of the cell, applied to the input before some nonlinear probabilistic spiking step (a similar point is made, in Aguera y Arcas & Fairhall, 2003).

From a more physical point of view, of course, this "degeneracy" of the STA is perhaps less surprising (in retrospect, at least), since decreasing dt corresponds to increasing the bandwidth of the current, and this should increase the "bandwidth" of the cross-correlation between the current input and the spike output (namely, the STA), as well. This intuition is supported by numerical experiments in which the white noise current input is preceded by some fixed prefilter of limited bandwidth (Pillow & Simoncelli, 2003; Paninski et al., 2004); for this prefiltered input, the STA does indeed have a nonvanishing limit as $dt \to 0$.

Second, on a related note, $STA(\tau)$ displays a square root singularity as $\tau \to 0^-$, which is due to the interaction of the Brownian motion term in the IF stochastic differential equation with the absorbing threshold at V_{th} (see also Badel et al., 2005, for a discussion of this point).

Finally, perhaps most surprising, the STA is basically parameter independent for τ close enough to zero. The STA scales linearly in σ, but all the other model parameters—μ, g, v_L, and v_{reset}—become irrelevant in the $\tau \to 0^-$ limit, due to the $\sqrt{dt}$ relationship between the scale and drift of a diffusion with bounded coefficients. Loosely speaking, the noise term dominates the leak terms on small timescales; diffusion processes with bounded parameters can be locally approximated by (zero-drift) Brownian motion. The linear scaling of the STA in σ, on the other hand, has interesting implications for the "adaptive" properties of the LIF cell, as discussed in more detail in Rudd & Brown (1997), Paninski et al. (2003), and Yu and Lee (2003). More globally (that is, if we do not confine our attention to times very near the spike), as emphasized in Paninski (in press) and Badel et al. (2005), the most likely voltage path dominates the STA for σ sufficiently small.

5.1 Directions. We briefly indicate a few possible ways to generalize the above results. As mentioned above, we could replace our linear LIF model with a more general stochastic differential equation:

$$dV = f(V)dt + a(V)I_t,$$

with $f(V), a(V)$ some fixed, uniformly smooth functions of voltage V (Brunel & Latham, 2003). Again, though, while this will change the firing statistics of the model (perhaps drastically), our results on the STA in the $\tau \to 0^-$ regime remain unchanged. (We are assuming, of course, that the Fokker-Planck equation corresponding to this model has a unique, differentiable invariant density $P_\infty(V)$; without such a unique invariant $P_\infty(V)$, it is typically not possible even to define the STA. (See, e.g., Karlin & Taylor, 1981, for conditions ensuring the existence of a $P_\infty(V)$ with the required properties.)

One interesting application of this idea involves invertible rescalings of the voltage axis, $V \to g(V)$, for $g(.)$ a smooth, invertible function. For example, taking $U(t) = \exp(V(t))$ gives us a geometric Brownian motion, which serves the same fundamental role in, for example, financial applications that the Ornstein-Uhlenbeck process serves in neural applications. Since invertible rescalings preserve the Markov property, all of our results go through unchanged after applying the usual change-of-measure formula.

Perhaps the fundamental step of our analysis is the Markov assumption. Thus, generalizations that would be worth exploring include the extension to more general spike-response models (Gerstner & Kistler, 2002), as defined by

$$dV = (f(V) + \eta(t - t_s))dt + a(V, t - t_s)I_t,$$

where η is a smooth function of $t - t_s$, the time since the last spike (this class of models allows for more interesting interspike interactions, since the subthreshold dynamics are no longer Markovian), and perhaps more importantly, to colored or conductance noise input I (Badel et al., 2005).

Appendix: Sampling

Sampling from the unconditioned stochastic differential equation 1.1 is straightforward, and will not be discussed further here (see, e.g., Risken, 1996 and Karatzas & Shreve, 1997, for a discussion). Conditional sampling, on the other hand—drawing samples from model 1.1, given the observed spike data $s[t_1, t_2]$—is not quite so obvious. The exact sampling method used here is a variant of the forward-backward algorithm described above for computing conditional densities and is again inherited from methods for sampling from hidden Markov models (Rabiner, 1989). An identical procedure is used in Paninski (in press) but will be described here for completeness. We initialize $V(t_2) = V_{th}$. (This initial condition is due to the data that a spike occurred at time t_2, as above.)

Now, for $t_2 > t > t_1$, sample backward:

$$\begin{aligned}
V(t) &\sim P(V(t)|\{V(u)\}_{t<u<t_2}, s[t_1, t_2] = \{t_1, t_2\}) \\
&= P(V(t)|V(t + dt), s[t_1, t] = \{t_1\}) \\
&= \frac{1}{Z} P(V(t + dt), s[t_1, t] = \{t_1\}|V(t))P(V(t)) \\
&= \frac{1}{Z} P(V(t + dt)|V(t))P(s[t_1, t] = \{t_1\}|V(t))P(V(t))
\end{aligned}$$

$$= \frac{1}{Z} P(V(t+dt)|V(t)) P(V(t)|s[t_1, t] = \{t_1\})$$
$$= \frac{1}{Z} P(V(t+dt)|V(t)) P_f(V, t).$$

Thus, sampling on each time step simply requires that we draw independently from a one-dimensional density, proportional to the product in the last line. Once this product has been computed, this sampling can be done using standard methods (namely, the inverse cumulative probability transform (Press, Teukolsky, Vetterling, & Flannery, 1992)). (Of course, $P_f(V, t)$ need only be computed once for $t_1 < t < t_2$, no matter how many samples are required.) The second term is computed directly from the gaussian stochastic dynamics, equation 1.1, given each $V(t + dt)$. Putting the samples together, for $0 < t < t_2$ clearly gives a sample from $P(\{V(t)\}_{0<t<t_2}|s[t_1, t_2] = \{t_1, t_2\})$, as desired.

Acknowledgments

This work was supported by a predoctoral fellowship from HHMI and an International Research Fellowship from the Royal Society. We thank Z. Ghahramani, P. Latham, B. Lau, J. Pillow, and E. Simoncelli for many interesting discussions, and especially L. Badel, W. Gerstner, and M. Richardson for discussions about their work (Badel et al., 2005). Part of this work (specifically, the material on the discrete-time STA) has appeared previously in unpublished preprint form.

References

Aguera y Arcas, B., & Fairhall, A. (2003). What causes a neuron to spike? *Neural Computation, 15*, 1789–1807.

Badel, L., Richardson, M., & Gerstner, W. (2006). Dependence of the spike-triggered average voltage on membrane response properties. *Neurocomputing, 69*, 1062–1065.

Berry, M., & Meister, M. (1998). Refractoriness and neural precision. *Journal of Neuroscience, 18*, 2200–2211.

Brunel, N., & Hakim, V. (1999). Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Computation, 11*, 1621–1671.

Brunel, N., & Latham, P. (2003). Firing rate of the noisy quadratic integrate-and-fire neuron. *Neural Computation, 15*, 2281–2306.

Bryant, H., & Segundo, J. (1976). Spike initiation by transmembrane current: A white-noise analysis. *Journal of Physiology, 260*, 279–314.

Burkitt, A., & Clark, G. (1999). Analysis of integrate-and-fire neurons: Synchronization of synaptic input and spike output. *Neural Computation, 11*, 871–901.

Bussgang, J. (1952). *Crosscorrelation functions of amplitude-distorted gaussian signals* (RLE Tech. Rep., 216). Cambridge, MA: MIT Research Laboratory of Electronics.

Chichilnisky, E. (2001). A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems, 12*, 199–213.

Daniels, H. (1982). Sequential tests contructed from images. *Annals of Statistics, 10*, 394–400.

Dayan, P., & Abbott, L. (2001). *Theoretical neuroscience*. Cambridge, MA: MIT Press.

de Boer, E., & Kuyper, P. (1968). Triggered correlation. *IEEE Transactions on Biomedical Engineering, 15*, 159–179.

de Ruyter, R., & Bialek, W. (1988). Real-time performance of a movement-sensitive neuron in the blowfly visual system: Coding and information transmission in short spike sequences. *Proc. R. Soc. Lond. B, 234*, 379–414.

Freidlin, M., & Wentzell, A. (1984). *Random perturbations of dynamical systems*. Berlin: Springer-Verlag.

Gerstner, W. (2001). Coding properties of spiking neurons: Reverse and cross-correlations. *Neural Networks, 14*, 599–610.

Gerstner, W., & Kistler, W. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge: Cambridge University Press.

Harvey, A. (1991). *Forecasting: Structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.

Haskell, E., Nykamp, D., & Tranchina, D. (2001). Population density methods for large-scale modelling of neuronal networks with realistic synaptic kinetics. *Network, 12*, 141–174.

Hida, T. (1980). *Brownian motion*. New York: Springer-Verlag.

Kanev, J., Wenning, G., & Obermayer, K. (2004). Approximating the response-stimulus correlation for the integrate-and-fire neuron. *Neurocomputing, 58*, 47–52.

Karatzas, I., & Shreve, S. (1997). *Brownian motion and stochastic calculus*. New York: Springer.

Karlin, S., & Taylor, H. (1981). *A second course in stochastic processes*. New York: Academic Press.

Kautz, R. (1988). Thermally induced escape: The principle of minimum available noise energy. *Physical Review A, 38*, 2066–2080.

Koch, C. (1999). *Biophysics of computation*. New York: Oxford University Press.

Paninski, L. (2003). Convergence properties of some spike-triggered analysis techniques. *Network: Computation in Neural Systems, 14*, 437–464.

Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems, 15*, 243–262.

Paninski, L. (in press). The most likely voltage path and large deviations approximations for integrate-and-fire neurons. *Journal of Computational Neuroscience.*

Paninski, L., Haith, A., Pillow, J., & Williams, C. (2005). Improved numerical methods for computing likelihoods in the stochastic integrate-and-fire model. *Comp. Sys. Neur. '05*. Salt Lake City, UT.

Paninski, L., Lau, B., & Reyes, A. (2003). Noise-driven adaptation: In vitro and mathematical analysis. *Neurocomputing, 52*, 877–883.

Paninski, L., Pillow, J., & Simoncelli, E. (2004). Maximum likelihood estimation of a stochastic integrate-and-fire neural model. *Neural Computation, 16*, 2533–2561.

Pillow, J., Paninski, L., Uzzell, V., Simoncelli, E., & Chichilnisky, E. (2005). Accounting for timing and variability of retinal ganglion cell light responses with a stochastic integrate-and-fire model. *Journal of Neuroscience, 25*, 11003–11013.

Pillow, J., & Simoncelli, E. (2003). Biases in white noise analysis due to non-Poisson spike generation. *Neurocomputing, 52*, 109–115.

Plesser, H., & Tanaka, S. (1997). Stochastic resonance in a model neuron with reset. *Physics Letters A, 225*, 228–234.

Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (1992). *Numerical recipes in C*. Cambridge: Cambridge University Press.

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE, 77*, 257–286.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Risken, H. (1996). *The Fokker-Planck equation*. New York: Springer.

Rudd, M., & Brown, L. (1997). Noise adaptation in integrate-and-fire neurons. *Neural Computation, 9*, 1047–1069.

Seshadri, V. (1993). *The inverse gaussian distribution*. Oxford: Clarendon.

Simoncelli, E., Paninski, L., Pillow, J., & Schwartz, O. (2004). Characterization of neural responses with stochastic stimuli. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (3rd ed.). Cambridge, MA: MIT Press.

Tuckwell, H. (1988). *Introduction to theoretical neurobiology*. Cambridge: Cambridge University Press.

Yu, Y., & Lee, T. (2003). Dynamical mechanisms underlying contrast gain control in single neurons. *Physical Review E, 68*, 011901.

Received September 12, 2005; accepted March 27, 2006.