



Capturing Multiple Timescales of Adaptation to Second-Order Statistics With Generalized Linear Models: Gain Scaling and Fractional Differentiation

Kenneth W. Latimer^{1*} and Adrienne L. Fairhall²

¹ Department of Neurobiology, University of Chicago, Chicago, IL, United States, ² Department of Physiology & Biophysics, University of Washington, Seattle, WA, United States

Single neurons can dynamically change the gain of their spiking responses to take into account shifts in stimulus variance. Moreover, gain adaptation can occur across multiple timescales. Here, we examine the ability of a simple statistical model of spike trains, the generalized linear model (GLM), to account for these adaptive effects. The GLM describes spiking as a Poisson process whose rate depends on a linear combination of the stimulus and recent spike history. The GLM successfully replicates gain scaling observed in Hodgkin-Huxley simulations of cortical neurons that occurs when the ratio of spike-generating potassium and sodium conductances approaches one. Gain scaling in the GLM depends on the length and shape of the spike history filter. Additionally, the GLM captures adaptation that occurs over multiple timescales as a fractional derivative of the stimulus envelope, which has been observed in neurons that include long timescale afterhyperpolarization conductances. Fractional differentiation in GLMs requires long spike history that span several seconds. Together, these results demonstrate that the GLM provides a tractable statistical approach for examining single-neuron adaptive computations in response to changes in stimulus variance.

Keywords: adaptation, gain scaling, fractional differentiation, generalized linear model (GLM), Hodgkin and Huxley model

OPEN ACCESS

Edited by:

Colin W. G. Clifford,
University of New South Wales,
Australia

Reviewed by:

Paul Miller,
Brandeis University, United States
Joshua A. Goldberg,
Hebrew University of Jerusalem, Israel

*Correspondence:

Kenneth W. Latimer
latimerk@uchicago.edu

Received: 13 January 2020

Accepted: 23 July 2020

Published: 09 September 2020

Citation:

Latimer KW and Fairhall AL (2020)
Capturing Multiple Timescales of
Adaptation to Second-Order Statistics
With Generalized Linear Models: Gain
Scaling and Fractional Differentiation.
Front. Syst. Neurosci. 14:60.
doi: 10.3389/fnsys.2020.00060

1. INTRODUCTION

Neurons adapt their spiking responses in a number of ways to the statistics of their inputs (Fairhall, 2014; Weber and Fairhall, 2019). A particularly well-studied example is adaptation to the stimulus variance, which can provide important computational properties. First, neurons can show gain scaling, such that the input is scaled by the stimulus standard deviation (Fairhall et al., 2001a; Mease et al., 2013). Scaling of the gain by the stimulus standard deviation implies that single spikes maintain the same information about the stimulus independent of its overall amplitude. This adaptation of the “input gain” with stimulus standard deviation can occur very rapidly. Second, the mean firing rate can adapt to variations in the stimulus variance across multiple timescales (Fairhall et al., 2001b; Wark et al., 2007). This form of spike frequency adaptation can in some cases have power-law properties (Pozzorini et al., 2013) and serve to compute the fractional derivative of the variance (Anastasio, 1998; Lundstrom et al., 2008).

One approach to studying such adaptation is to use Hodgkin-Huxley style (HH) conductance based models to explore potential single-neuron mechanisms underlying these computations (Lundstrom et al., 2008; Mease et al., 2013). Although HH models can indeed capture such behavior, the mechanistic HH framework is not ideally suited to reveal its dynamical basis as HH model parameters are difficult to interpret in terms of computation and coding. Moreover, fitting HH models to intracellular data is difficult (Buhry et al., 2011; Cserscsik et al., 2012; Vavoulis et al., 2012; Lankarany et al., 2014), and only recently methods that fit HH models to spike trains alone have been gaining success (Meng et al., 2011, 2014).

In contrast, statistical point-process models based on the generalized linear model (GLM) framework have provided a tractable tool for modeling spiking responses of neurons in sensory systems (Truccolo et al., 2005; Pillow et al., 2008). Previous work has shown the utility of finding linear features that can explain the spiking behavior of HH models (Agüera y Arcas and Fairhall, 2003; Agüera y Arcas et al., 2003; Weber and Pillow, 2017). Unlike simple linear/non-linear models, GLMs also incorporate a dependence on the history of activity, potentially providing a helpful interpretative framework for adaptation (Mease et al., 2014; Latimer et al., 2019a). We therefore fit GLMs to spike trains generated from HH neurons. Here, we considered gain scaling and fractional differentiation individually by using one set of HH neurons that showed gain scaling and another that produced fractional differentiation. Our goal here is not to analyze the details of these models to explain the adaptation phenomena, but to use these models as a source of “experimental” data with which to test the ability of the GLM to capture the same phenomena. We found that the GLMs could reproduce the single-neuron adaptive computations of gain scaling or fractional differentiation. Capturing gain scaling across a range of HH active conductance parameters depended both on the choice of link function and spike history length. As the length of the spike history filter increased, the stimulus dependency of neurons changed from differentiating to integrating (Stevenson, 2018). Capturing adaptation as a fractional derivative required a history filter that could account for long timescale effects: on the order of 10 s. Together these results demonstrate that the GLM provides a tractable statistical framework for modeling adaptation that occurs at the single-neuron level.

2. MATERIALS AND METHODS

To generate spike train data that display the two types of adaptation we study, we use Hodgkin-Huxley neuron models with parameter choices and channel dynamics previously shown to reproduce effects seen in experimental data. In principle we could use a single model that incorporates both gain scaling and fractional differentiation. While both types of adaptation have been observed in fly neuron H1 (Fairhall et al., 2001b), they have not been reported to occur in the same cortical cells. We therefore choose here to separate the two effects by considering two previously proposed HH models: one that has been applied to model gain scaling (Mease et al., 2013) and a second which

exhibits fractional differentiation (Lundstrom et al., 2008). We verified (data not shown) that both effects can be obtained within the same model, but with slight quantitative differences.

2.1. Gain Scaling

Gain scaling refers to the case when for an input-output function of a neuron, the input gain is proportional to the standard deviation (SD) of the stimulus (σ). Thus, the gain depends on the recent context. If a neuron achieves perfect gain scaling, the firing rate R given a particular stimulus value, s , and input standard deviation can be written as:

$$R_\sigma(s) = \bar{R}_\sigma \hat{R}\left(\frac{s}{\sigma}\right) \tag{1}$$

where the normalized stimulus $\hat{s} = \frac{s}{\sigma}$, and the output gain, \bar{R}_σ , is constant in s .

To quantify the degree of gain scaling in a neuron’s spiking output, we measure the firing rate function in response to a white-noise input, $x(t)$, at different SDs and constant mean μ (Figure 1A). For each standard deviation, we compute the normalized spike-triggered average (STA; Figure 1B) (Rieke et al., 1999). We then compute the stimulus as the convolution $s(t) = \int_0^t \text{STA}(t')(x(t-t') - \mu)dt'$. The spike rate function is then defined probabilistically as

$$R_\sigma(s)\Delta_t = p_\sigma(sp|k|\hat{s}) = \frac{p_\sigma(\hat{s}|spk)}{p_\sigma(\hat{s})} p_\sigma(sp) \tag{2}$$

where the right side follows from Bayes’ rule. The average firing rate in time bin of width Δ_t is $p_\sigma(sp)$. Thus, we get $\bar{R}_\sigma \Delta_t = p_\sigma(sp)$ and $\hat{R}\left(\frac{s}{\sigma}\right) = \frac{p_\sigma(\hat{s}|spk)}{p_\sigma(\hat{s})}$. The spike-triggered stimulus distribution, $p_\sigma(\hat{s}|spk)$, is the probability of the stimulus given that a spike occurred in the bin. By definition the marginal stimulus distribution, $p_\sigma(\hat{s})$, is a standard normal distribution which does not depend on σ . Therefore, if $p_\sigma(\hat{s}|spk)$ is similar across different values of σ , gain scaling is achieved because $\hat{R}\left(\frac{s}{\sigma}\right)$ does not depend on σ .

We measure gain scaling in terms of the spike-triggered distribution. We do so using the 1st Wasserstein, or earth-mover’s metric (Rubner et al., 1998) (we obtained qualitatively similar results using the symmetrized Kullback-Leibler divergence and Jensen-Shannon divergence). The Wasserstein metric is a distance function between two probability distributions. Intuitively, it can be thought of as the minimum work needed to transform one distribution into the other by moving probability mass as if the distributions are piles of sand (Supplementary Figure 1). Formally, it is defined as

$$W_1(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu, \nu)} \int_{M \times M} d(x, y) d\gamma(x, y) \tag{3}$$

where ν and μ are probability measures on a metric space M with metric $d(\cdot, \cdot)$. The infimum is taken over the collection of measures, $\Gamma(\mu, \nu)$, on $M \times M$ with μ and ν marginal distributions. We compute the gain scaling score at σ as

$$D_\sigma = W_1(p_1(\hat{s}|spk), p_\sigma(\hat{s}|spk)). \tag{4}$$

A distance close to 0 indicates that the spike-triggered distributions are similar, and therefore the cell is gain scaling its input (Figures 1C,D). Larger values of D_σ indicate that the input-output function does not scale with σ (Figures 1E–G). We computed the spike-triggered distribution using a histogram with bins of width 0.1.

2.1.1. Gain Scaling in Hodgkin-Huxley Neurons

A previous study by Mease et al. (2013) found that Hodgkin-Huxley models could account for gain scaling observed in pyramidal neurons. Thus we simulated spikes from single-compartment Hodgkin-Huxley style models of pyramidal neurons, providing a source of data with which to explore the expression of this property using GLMs. In this model, the membrane voltage depends on voltage-gated sodium and potassium conductances (G_{Na} and G_K) and a passive leak conductance (G_L). The voltage and gating dynamics followed the equations (Mainen et al., 1995)

$$C \frac{dV}{dt} = I_{stim}(t) - G_{Na}m^3h(V - E_{Na}) - G_Kn(V - E_K) - G_L(V - E_L) \quad (5)$$

such that for each gate $x \in \{n, m, h\}$

$$\tau_x(V) \frac{dx}{dt} = x_\infty(V) - x, \quad \tau_x(V) = \frac{1}{\alpha_x(V) + \beta_x(V)} \quad (6)$$

$$n_\infty(V) = \alpha_n(V)\tau_n(V), \quad m_\infty(V) = \alpha_m(V)\tau_m(V), \quad h_\infty(V) = \frac{1}{1 + \exp\left(\frac{V+65}{6.2}\right)} \quad (7)$$

$$\alpha_n(V) = \frac{20(V - 20)}{1 - \exp\left(-\frac{V-20}{9}\right)}, \quad \beta_n(V) = \frac{-2(V - 20)}{1 - \exp\left(\frac{V-20}{9}\right)} \quad (8)$$

$$\alpha_m(V) = \frac{182(V + 35)}{1 - \exp\left(-\frac{V+35}{9}\right)}, \quad \beta_m(V) = \frac{-124(V + 35)}{1 - \exp\left(\frac{V+35}{9}\right)}$$

$$\alpha_h(V) = \frac{24(V + 50)}{1 - \exp\left(-\frac{V+50}{5}\right)}, \quad \beta_h(V) = \frac{-9.1(V + 75)}{1 - \exp\left(\frac{V+75}{5}\right)}$$

The parameters of the model were the same as in Mease et al. (2013). The reversal potentials were $E_{Na} = 50$, $E_K = -77$, and $E_L = -70$ mV and the capacitance was $C = 1\mu F/cm^2$. The leak conductance was set to $0.4\text{pS}/\mu m^2$ so that the resting membrane had a time constant of ~ 25 ms. As in Mease et al. (2013), we explored a range of values for the active conductances G_{Na} and G_K : from 600 to 2,000 pS/ μm^2 in increments of $100\text{pS}/\mu m^2$. Simulations were performed in MATLAB using a fourth-order Runge-Kutta method with step size 0.01ms. Spike times were defined as upward crossings of the voltage trace at -10mV separated by at least 2ms.

The input was constructed as independent Gaussian draws every 1 ms with parameters $\mathcal{N}(\mu, (4\mu\sigma)^2)$ where σ was set to 1.0, 1.3, 1.6, or 2.0. The mean was constrained to be proportional to the standard deviation similarly the current-clamp protocol used to study gain scaling in Mease et al. (2013). For each value of G_{Na} and G_K , the mean input, μ , was tuned so that at baseline,

where $\sigma = 1$, each simulation produced ~ 10 spk/s using a 100s simulation. We did not consider values of G_{Na} and G_K that spiked spontaneously (i.e., spiked when $\mu = 0$). We simulated 2,000 s of spiking activity at each stimulus level (generating $\sim 20,000$ spikes at $\sigma = 1$).

2.2. Fractional Differentiation

We next looked at periodic modulations of the stimulus standard deviation to model long timescale adaptive effects. We applied stimuli consisting of Gaussian noise with sinusoidal or square wave modulation of the variance between 1 and σ with σ again taking values of 1.3, 1.6, or 2.0, at a number of different frequencies. We analyzed simulated spike trains across seven noise modulation periods: 1, 2, 4, 8, 16, 32, and 64s. The simulations were 3,200 s for each period, giving a minimum of 50 cycles per period.

Lundstrom et al. (2008) found that pyramidal neurons can act as fractional differentiators of the stimulus amplitude envelope for this type of input. Fractional derivatives generalize the derivative operator such that, analogous to taking the first derivative of a function twice to obtain the second derivative, taking the fractional derivative of order $\alpha = 1/2$ twice results in the first derivative (Oldham and Spanier, 1974). Fractional differential filters respond to a square stimulus as an exponential-like decay with a time constant that depends on α (Figures 2A,B). Fractionally differentiating a sinusoidal stimulus produces a frequency dependent gain change (Figure 2C)

$$\text{gain} \propto f^\alpha \quad (9)$$

where f is the frequency. Additionally, fractionally differentiating the sine function gives a frequency independent phase shift, ϕ , of the stimulus (Figure 2D):

$$\phi = \alpha \frac{\pi}{2}. \quad (10)$$

These three measures can be combined to estimate approximate fractional differentiation by neurons.

To compute the fractional derivative order, we computed cycle-averaged responses obtained using 30 bins per cycle at each stimulus amplitude modulation frequency. We fit the cycle-averaged square-wave responses across all modulation frequencies as the best fitting fractional derivative of the stimulus amplitude (plus a baseline rate) using least-squares. To fit α to the phase lead of the sine-wave responses, we computed mean phase lead (ϕ) across frequencies and applied Equation (10). To fit α to the gain of the sine-wave responses, we applied Equation (9) by fitting a least-squares regression between the frequency of modulation and the logarithm of the gain.

2.2.1. Fractional Differentiation by Hodgkin-Huxley Neurons

We simulated neurons from the standard HH model with three additional afterhyperpolarization (AHP) currents with time constants ranging from 0.3 to 6s as was done by Lundstrom et al.

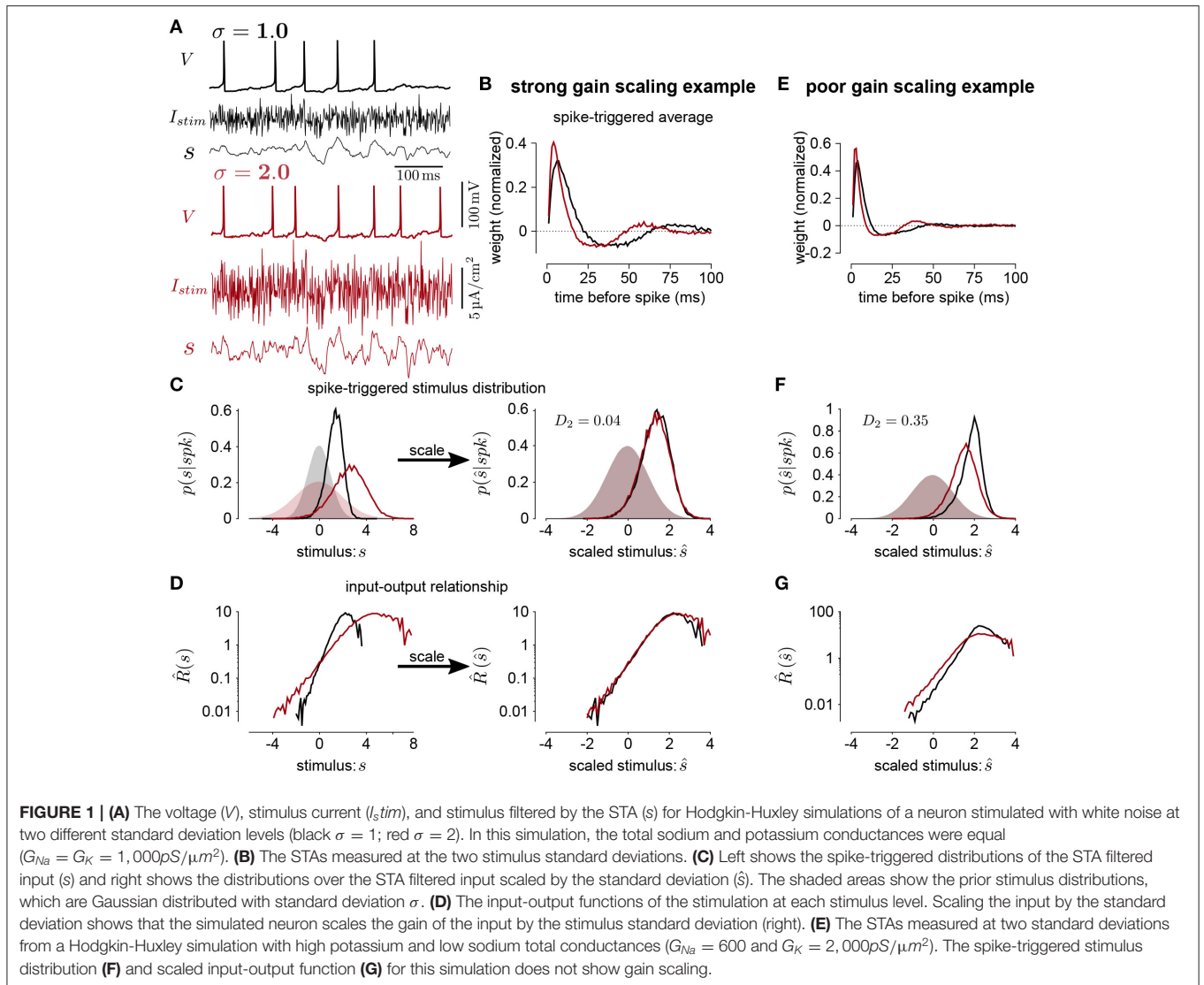


FIGURE 1 | (A) The voltage (V), stimulus current (I_{stim}), and stimulus filtered by the STA (s) for Hodgkin-Huxley simulations of a neuron stimulated with white noise at two different standard deviation levels (black $\sigma = 1$; red $\sigma = 2$). In this simulation, the total sodium and potassium conductances were equal ($G_{Na} = G_K = 1,000 \text{ pS}/\mu\text{m}^2$). **(B)** The STAs measured at the two stimulus standard deviations. **(C)** Left shows the spike-triggered distributions of the STA filtered input (s) and right shows the distributions over the STA filtered input scaled by the standard deviation (\hat{s}). The shaded areas show the prior stimulus distributions, which are Gaussian distributed with standard deviation σ . **(D)** The input-output functions of the stimulation at each stimulus level. Scaling the input by the standard deviation shows that the simulated neuron scales the gain of the input by the stimulus standard deviation (right). **(E)** The STAs measured at two standard deviations from a Hodgkin-Huxley simulation with high potassium and low sodium total conductances ($G_{Na} = 600$ and $G_K = 2,000 \text{ pS}/\mu\text{m}^2$). The spike-triggered stimulus distribution **(F)** and scaled input-output function **(G)** for this simulation does not show gain scaling.

(2008). The equations for the HH neurons were

$$C \frac{dV}{dt} = I_{stim}(t) - G_{Na} m^3 h (V - E_{Na}) - G_K n^4 (V - E_K) - G_L (V - E_L) - \sum_{i=1}^3 G_{AHP,i} a_i (V - E_{AHP}) \quad (11)$$

The gates $x \in n, m, h$ follow the dynamics

$$\tau_x(V) \frac{dn}{dt} = x_\infty(V) - x, \quad \tau_x(V) = \frac{1}{\alpha_x(V) + \beta_x(V)}, \quad x_\infty(V) = \alpha_x(V) \tau_x(V) \quad (12)$$

$$\begin{aligned} \alpha_n(V) &= \frac{0.01(V + 55)}{1 - \exp(-0.1(V + 55))}, \\ \beta_n(V) &= 0.125 \exp(-(V + 65)/80) \\ \alpha_m(V) &= \frac{0.1(V + 40)}{1 - \exp(-0.1(V + 40))}, \\ \beta_m(V) &= 4 \exp(-(V + 65)/18) \\ \alpha_h(V) &= 0.07 \exp(-(V + 65)/20), \\ \beta_h(V) &= \frac{1}{1 + \exp(-0.1(V + 35))}. \end{aligned} \quad (13)$$

The AHP currents have linear dynamics and are incremented by 1 at spike times ($t_{spk,i}$):

$$\frac{da_i}{dt} = -\frac{a_i}{\tau_i} + \sum_i \delta(t - t_{spk,i}) \quad (14)$$

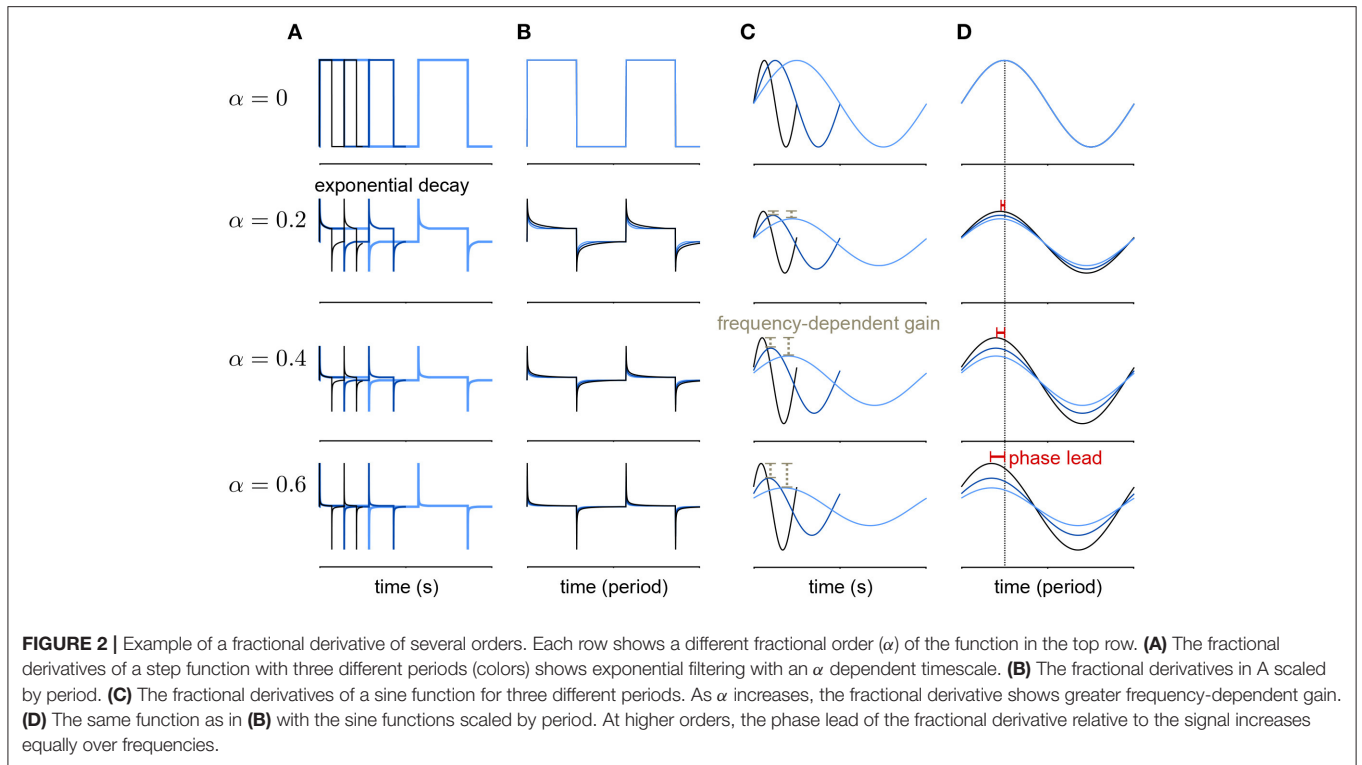


FIGURE 2 | Example of a fractional derivative of several orders. Each row shows a different fractional order (α) of the function in the top row. **(A)** The fractional derivatives of a step function with three different periods (colors) shows exponential filtering with an α dependent timescale. **(B)** The fractional derivatives in A scaled by period. **(C)** The fractional derivatives of a sine function for three different periods. As α increases, the fractional derivative shows greater frequency-dependent gain. **(D)** The same function as in **(B)** with the sine functions scaled by period. At higher orders, the phase lead of the fractional derivative relative to the signal increases equally over frequencies.

where δ is the Dirac delta function. The parameters standard were: $G_{Na} = 120$, $G_K = 36$, $G_L = 0.3mS/cm^2$; $E_{Na} = 50$, $E_K = -77$, $E_L = -54.4mv$; and $C = 1\mu F/cm^2$. The AHP conductances were set relative to the leak conductance: $G_{AHP, \cdot} = (0.05, 0.006$ and $0.004)G_L$. The AHP reversal potential was $E_{AHP} = -100mv$ and the AHP timescales were set to $\tau_i = (0.3, 1, \text{ and } 6)s$.

Similarly to the gain scaling simulations, the stimulus was sampled independently in each 1ms bin from a normal distribution with mean μ . The time-dependent variance given σ and the period p was $4\mu f_p(t, \sigma)$. The time-dependent modulation function for the square-wave stimulus was

$$f_p(t, \sigma) = 1 + (\sigma - 1) \left\lfloor \frac{1}{2} \sin\left(\frac{2t\pi}{p}\right) + 1 \right\rfloor \quad (15)$$

where $\lfloor \cdot \rfloor$ denotes the floor operator, and the function for the sine-wave stimulus was similarly defined as

$$f_p(t, \sigma) = 1 + (\sigma - 1) \left(\frac{1}{2} \sin\left(\frac{2t\pi}{p}\right) + \frac{1}{2} \right). \quad (16)$$

The parameter μ was calibrated so that with no variance modulation (i.e., $\sigma = 1$), the simulated cells produced $\sim 10spk/s$.

2.3. Generalized Linear Models

The GLM defines the spiking activity as an autoregressive Poisson process with **(Figure 3A)**. The spike rate at time t is given as a linear–nonlinear function of the stimulus and the spike history

$$\lambda_t = f\left(\mathbf{k}_{stim}^\top \mathbf{x}_t + \mathbf{h}_{spk}^\top \mathbf{y}_{hist,t} + b\right) \quad (17)$$

where \mathbf{x}_t is the stimulus vector preceding time t (the values of I_{stim}), and \mathbf{y}_{hist} is the spike history vector. The parameters of the GLM are the stimulus filter (\mathbf{k}_{stim}), the spike history filter (\mathbf{h}_{spk}), and baseline rate (b). For the inverse-link function, f , we used the canonical exponential function except where otherwise noted.

The log-likelihood of a binned spike train, \mathbf{y} , given the model parameters is then

$$\log p(\mathbf{y}|\mathbf{k}_{stim}, \mathbf{h}_{spk}, b) = \sum_{t=1}^T -\lambda_t \Delta_t + y_t \log(\lambda_t) + const. \quad (18)$$

For all model fits and simulations, we set $\Delta_t = 1$ ms. We numerically maximized the log-likelihood using conjugate-gradient methods.

To reduce the number of model parameters, we parameterized the filters using smooth basis functions **(Figure 3B)**. The stimulus filter was parameterized using 15 raised cosine basis functions:

$$\mathbf{k}_{stim}(t) = \sum_{j=1}^{15} z_j \mathbf{g}_j(t),$$

$$\mathbf{g}_j(t) = \begin{cases} \frac{1}{2} \cos\left(\frac{\log[t+c]-\phi_j}{a}\right) + \frac{1}{2} & \text{for } \frac{\log[t+c]-\phi_j}{a} \in [-\pi, \pi] \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

where t is in seconds. To fit \mathbf{k}_{stim} , we optimized the weights z_j . We set $c = 0.02$ and $a = 2(\phi_2 - \phi_1)/\pi$. The ϕ_j were evenly spaced from $\phi_1 = \log(T_0/1000 + c)$, $\phi_{15} = \log(T_{end}/1000 + c)$ where the peaks of the filters are in the range $T_0 = 0$ and $T_{end} = 100$ ms.

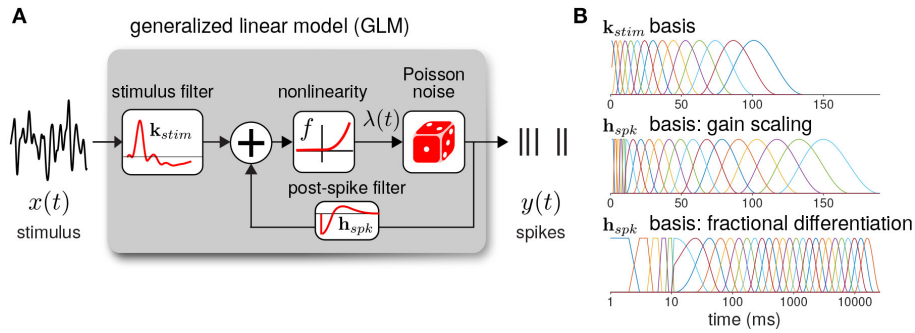


FIGURE 3 | (A) Diagram of the neural GLM that describes spiking as an autoregressive Poisson process. **(B)** The basis functions used to parameterize the GLM filters. (top) The stimulus basis used for all GLMs. (middle) The spike history basis used for the gain scaling simulations. (bottom) The spike history basis used for the fractional differentiation simulations. Due to the length of the spike history filters needed to capture fractional differentiation, the time axis is shown in log scale.

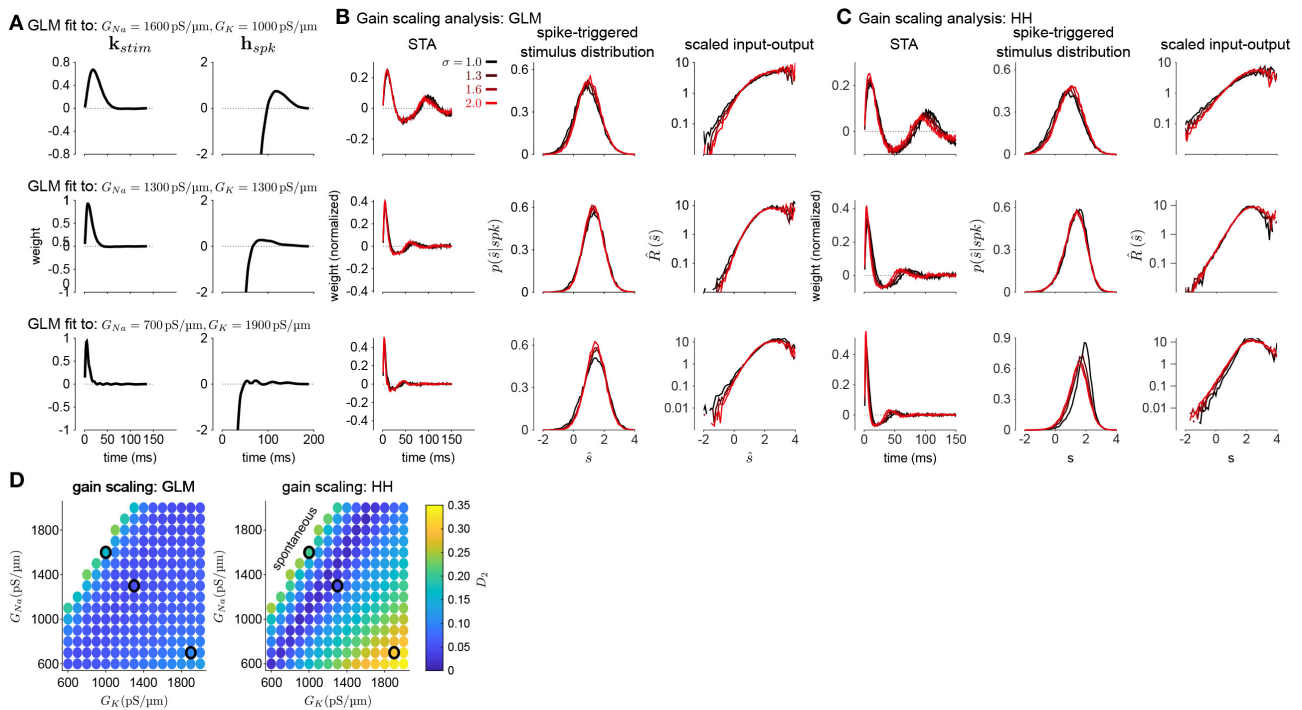
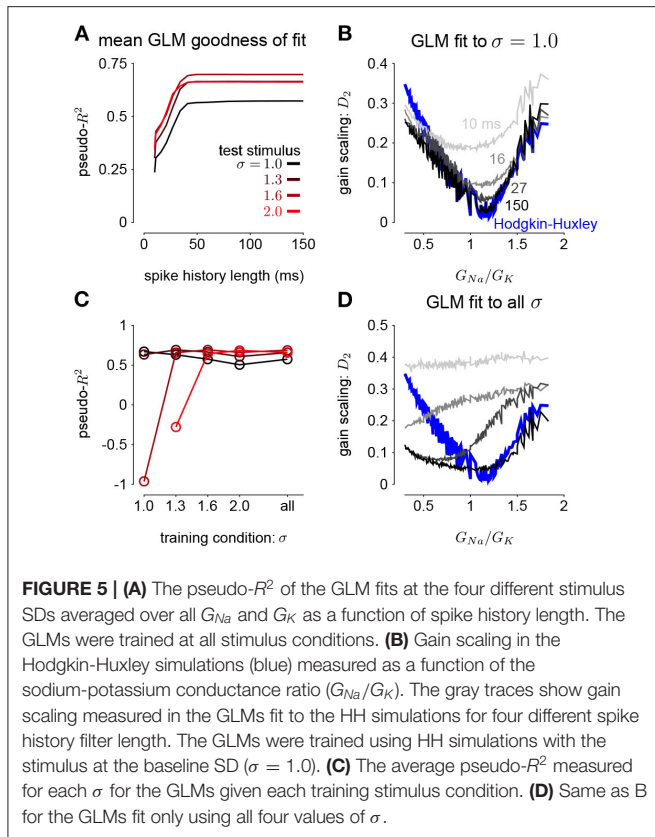


FIGURE 4 | (A) Example filters from GLM fits to HH simulations with three different spiking conductance levels (rows). Large negative values driving the refractory period in the spike history filter (right) have been truncated. **(B)** The spike-triggered averages (left), scaled spike-triggered stimulus distributions (right), and scaled input-output functions (right) for the GLM fits in A for all four stimulus SDs. **(C)** Same as B for the HH simulations. **(D)** Gain scaling (D_2 from Equation (4), Wasserstein distance between the spike-triggered distributions at $\sigma = 1$ and $\sigma = 2$) at all the spiking conductance levels explored for the GLM simulations (top) and the Hodgkin-Huxley simulations (bottom). Lower values of D_2 correspond to stronger gain scaling. The three black circles indicate the conductance levels for the GLM examples in A and B. Gain scaling was not computed for values of G_{Na} and G_K that resulted in spontaneous spiking in the Hodgkin-Huxley simulations.

The spike history filter bases were constructed in two parts. To account for the absolute refractory period, we used 5 box car filters of width 2 ms for the first 10 ms of the spike history. The remaining spike history filter was parameterized using raised cosine basis functions with the parameter $c = 0.05$. For the gain scaling simulations, $N = 15$ cosine basis functions were used with spacing $T_0 = 10$ and $T_{end} = 150$ ms. For the fractional differentiation simulations, $N = 25$ cosine basis functions were

used with spacing $T_0 = 10$ and $T_{end} = 16,000$ ms. To explore how the timescale of spike history affected adaptation in the GLM, for each model we fit the GLM using only the first i cosine basis functions for each $i = 0$ (using only the refractory box-car functions) to $i = N$. Thus, we obtained $N + 1$ nested model fits across a range of spike history lengths. When stated, the length of the spike history filter, T_{hist} , denotes the time of the peak of the i th basis function.



2.3.1. Evaluating Model Performance

We evaluated the GLM performance by assessing the ability of the GLM to predict the HH model response to a 32 s novel stimulus. For the gain scaling simulations, we tested the response to the test stimulus at each stimulus SD (σ). For the fractional differentiation simulations, the stimulus SD was modulated by a sine or square wave with a 4 s period and a modulation height of $\sigma = 2.0$. Predictive performance was evaluated using the pseudo- R^2 score (Cameron and Windmeijer, 1997). We selected this measure because it can be applied to Poisson process observations instead of trial-averaged firing rates as is required by the standard R^2 measure of explained variance (Benjamin et al., 2018). Thus, it is especially appropriate for comparing the stochastic GLM to a spike train simulated by the deterministic HH model. The pseudo- R^2 is written as the ratio of deviances:

$$\begin{aligned} \text{pseudo} - R^2 &= 1 - \frac{D(\mathbf{y}^*, \text{GLM})}{D(\mathbf{y}^*, \text{null})} \\ &= 1 - \frac{\log(p_{\text{GLM}}(\mathbf{y}^* | \mathbf{k}_{\text{stim}}, \mathbf{h}_{\text{spk}}, b)) - \log(p_{\text{sat.}}(\mathbf{y}^*))}{\log(p_{\text{null}}(\mathbf{y}^* | \bar{\mathbf{y}}^*)) - \log(p_{\text{sat.}}(\mathbf{y}^*))} \end{aligned} \quad (20)$$

where \mathbf{y}^* is the test spike train. The GLM likelihood is $p_{\text{GLM}}(\mathbf{y}^* | \mathbf{k}_{\text{stim}}, \mathbf{h}_{\text{spk}}, b)$ and the likelihood of the null model ($p_{\text{null}}(\mathbf{y}^* | \bar{\mathbf{y}}^*)$) is the probability of the spike train given only the mean firing rate, $\bar{\mathbf{y}}^*$. The saturated model likelihood ($p_{\text{sat.}}(\mathbf{y}^*)$) is the probability of observing \mathbf{y}^* given one parameter per bin:

that is, the Poisson probability of observing \mathbf{y}^* given a model with rate $\lambda = 1$ in each bin in which the HH model spiked and rate $\lambda = 0$ in each bin that the HH did not spike. Thus, the pseudo- R^2 measures the fraction of explainable log-likelihood captured by the GLM.

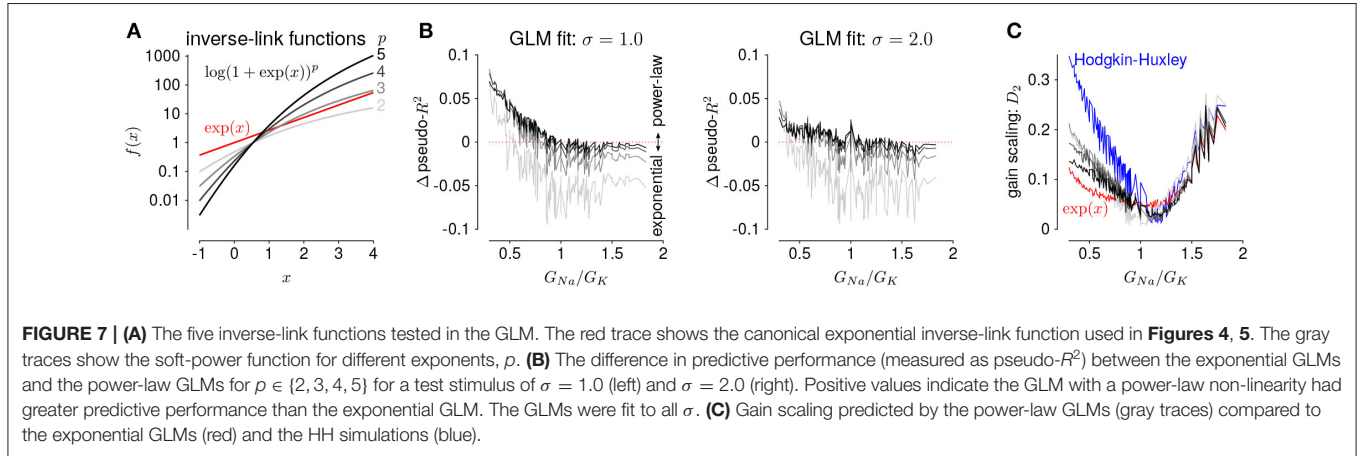
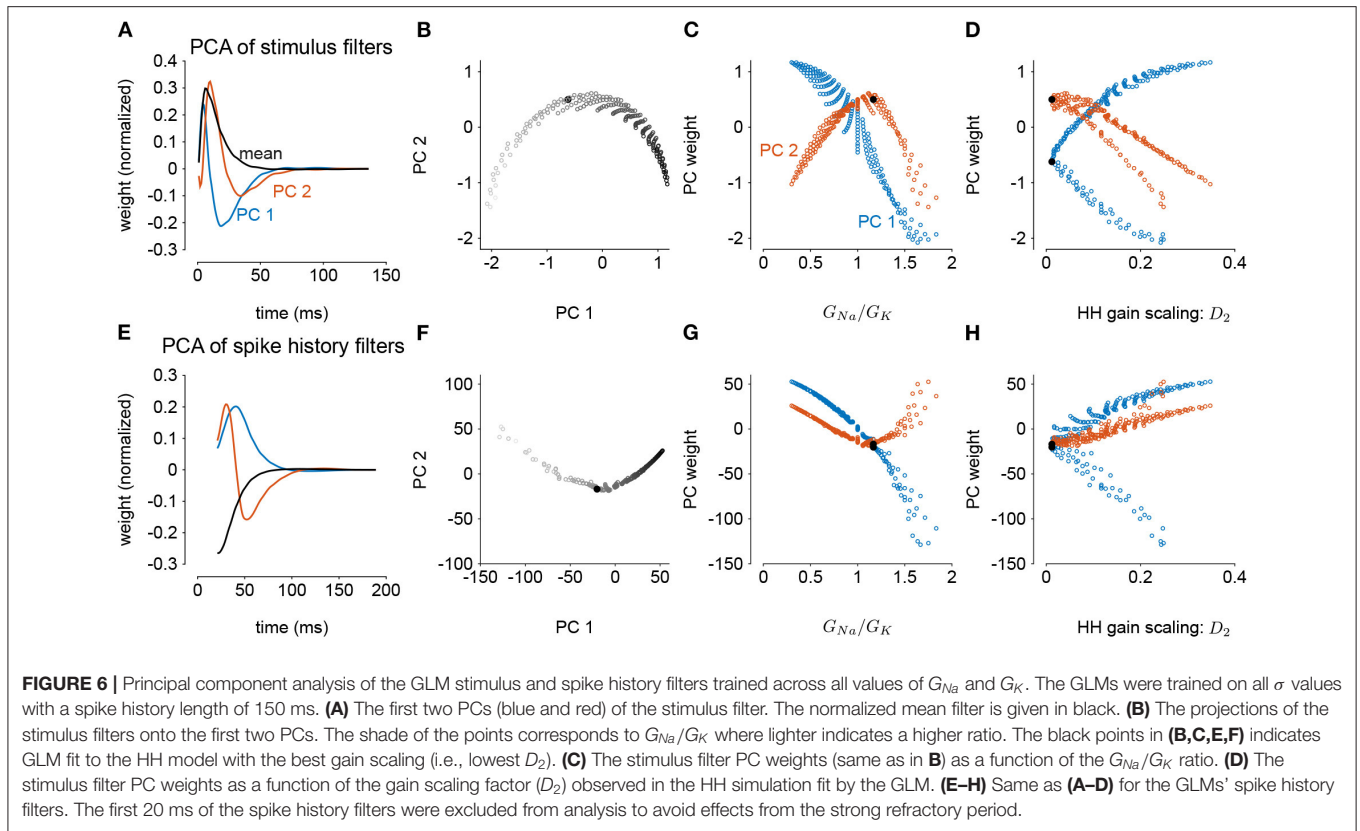
3. RESULTS

3.1. GLMs Capture Gain Scaling Behavior

To investigate how GLMs can capture biophysically realistic gain scaling, we fit the Hodgkin-Huxley simulations with GLMs (Figure 4A). We fit a unique GLM for each value of G_{Na} and G_K in the HH model, and the GLMs were fit using the entire range of stimulus SDs ($\sigma = 1.0, 1.3, 1.6, \text{ and } 2.0$). Applying the STA analysis at the four stimulus SDs, we quantified gain scaling in GLM fits and compared the gain scaling in the GLM simulations to the HH neurons (Figures 4B,C). Across the range of spiking conductance values, we found that the GLM fits consistently showed gain scaling (Figure 4D). The HH neurons showed the greatest degree of gain scaling when the G_{Na}/G_K ratio was close to one, with the lowest D_2 score occurring at a ratio of 1.17 (Mease et al., 2013). We observed the same pattern in the GLM simulations, but the GLM fits generally exhibited stronger gain scaling when $G_{Na}/G_K < 1$ than the HH neurons. We note that in general the optimal G_{Na}/G_K ratio depends on the leak conductance; however, here we assumed a fixed leak for simplicity.

The GLM's characterization of the HH neurons depended on the spike history filter. This is revealed by comparing the stimulus filters (Figure 4A) to the stimulus features extracted by spike-triggered averaging (Figure 4B): While the STA showed multiphasic responses, the GLM stimulus filter was consistent with a simple, monophasic integration. This demonstrates that the STA reflects the combination of stimulus and spike history effects (Agüera y Arcas and Fairhall, 2003; Stevenson, 2018). A spike history filter of sufficient length was necessary to achieve accurate model fits across all stimulus SDs (Figures 5A,B). A possible interpretation of this finding is that the spike history is acting as a measure of the "context" that serves to normalize the stimulus response, and that a sufficiently long sampling of the spiking output is needed in order to perform this normalization appropriately.

We also explored how the stimulus conditions used to fit the GLM determined the model's ability to capture gain scaling. Remarkably, we found that the GLM fit only to the baseline stimulus SD ($\sigma = 1.0$) captured the gain scaling pattern seen in the HH neuron (Figure 5B). The gain scaling observed in the GLMs required a sufficiently long spike history filter, on the order of at least 50 ms. With shorter spike history, the GLM did not obtain the same level of gain scaling at the optimal G_{Na}/G_K ratio. However, these GLM fits failed to generalize across stimulus SDs. The GLM trained only at $\sigma = 1.0$ explained less variance in the spiking responses to a stimulus at $\sigma = 2.0$ than a model capturing only the mean firing rate for all values of G_{Na} and G_K (predictive pseudo- $R^2 < 0$; Figure 5C). Therefore, the GLM trained at $\sigma = 1.0$ does not accurately characterize the HH responses despite accurately predicting gain scaling in



those cells. In contrast, GLMs trained at all four σ values failed to capture the lack of gain scaling at low G_{Na}/G_K values despite showing improved model fit across all σ (**Figure 5D**; a detailed example is provided in **Supplementary Figure 2A**). Because the GLM trained on all σ showed both consistent generalization performance and strong gain scaling behavior, the remaining analyses considered only that training condition.

We next considered how the GLM parameters related to the gain scaling computation and the space of G_{Na} and G_K in the HH models. To visualize the geometry of the model parameters, we performed PCA on the stimulus and

spike history filters (**Figures 6A,E**). The filters produced across the two HH parameters spanned a two-dimensional subspace (variance explained: stimulus 98.8%, spike history 97.3%). The PCA reconstructions for example stimulus filters are given in **Supplementary Figure 3**. This decomposition of the stimulus filter shows that the first mode is primarily an integrator, while the second acts as a derivative, and will serve to adjust the timescale of the filter's integration window. The inflection point for the second mode suggests that there may be an optimal time constant of stimulus and spike history integration needed to support the gain scaling property (**Figures 6B,F**). The first

component for both filters correlated with the G_{Na}/G_K ratio (Figures 6C,G; stimulus PC1 $r = -0.97, p < 10^{-4}$; spike history PC2 $r = 0.97, p < 10^{-4}$). The second PC correlates with the gain scaling value observed in the corresponding HH model (Figures 6D,H; stimulus PC2 $r = -0.89, p < 10^{-4}$; spike history PC2 $r = 0.90, p < 10^{-4}$).

3.1.1. Power-Law Firing Rate Nonlinearities

The GLMs we considered used the canonical inverse-link function, the exponential nonlinearity (McCullagh and Nelder, 1989), to transform the filtered stimulus plus spike history into a firing rate. However, it is known that firing rate nonlinearities that instead have a power-law relationship of the input produce gain scaling (Miller and Troyer, 2002; Murphy and Miller, 2003). We therefore considered a range of soft-power nonlinearities over a range of exponents for the GLM firing rate (Figure 7A; Equation 17):

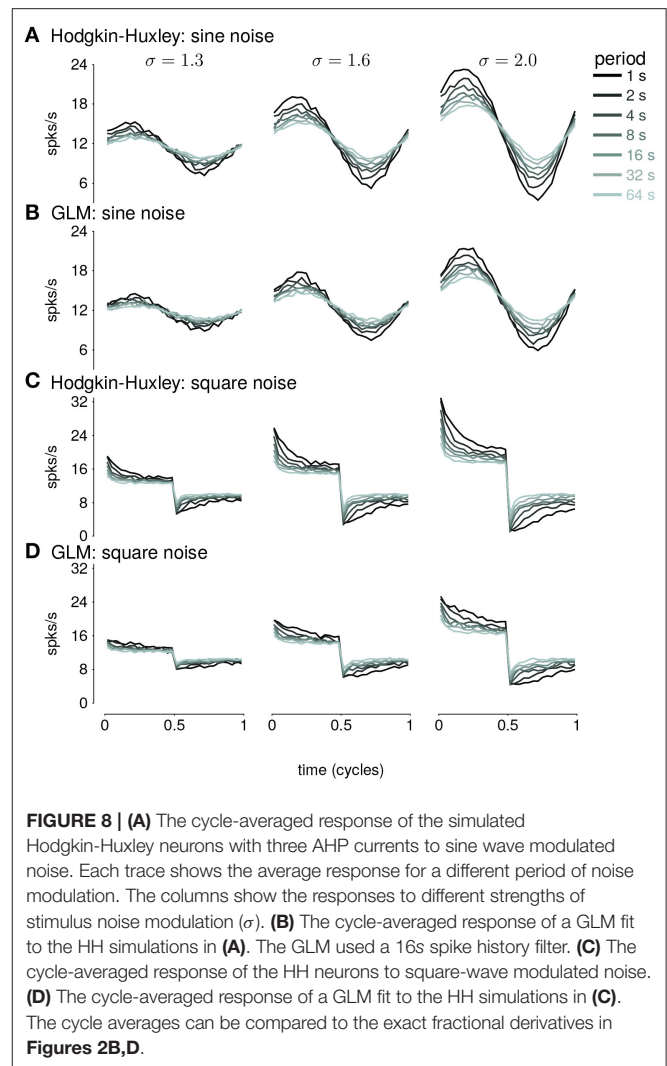
$$f(x) = \log(1 + \exp(x))^p \quad (21)$$

for $p \in \{2, 3, 4, 5\}$ (for $p = 1$, the model performed poorly for all HH simulations and the results are not shown). We found that the power-law nonlinearity produced better predictive fit than the exponential for HH simulations with low G_{Na}/G_K ratios (Figure 7B). For those ratios, the exponential GLM in fact predicted *greater* gain scaling than the HH simulation actually showed (Figure 5A and Supplementary Figure 2A). We found the power-law nonlinearities showed *less* gain scaling in the low G_{Na}/G_K regime, which was more consistent with the HH simulations (Figure 7C). This perhaps counter-intuitive result is likely due to the temporal processing of the GLM: the spike history filter shapes the effective stimulus-response function over longer timescales. Thus, the instantaneous spike rate function need not be a power law to produce gain scaling and an instantaneous power-law function may not result in strong gain scaling in the presence of spike history dependencies.

3.2. GLMs Capture Fractional Differentiation With Long Timescales of Adaptation

In this section, we address adaptive computations occurring over multiple timescales spanning tens of seconds, instead of instantaneous gain. We consider adaptation to changes in stimulus variance in the responses of HH simulations with three AHP currents (Lundstrom et al., 2008). The neurons were injected with noise stimuli with a periodically modulated SD. The stimulus SD followed either a sine or square wave. We focused our analyses on the cycle-averaged firing rate to see how the neural responses reflect fractional differentiation of the stimulus SD envelope in the cycle-averages.

We fit GLMs to HH simulations in response to either sine- or square-wave SD modulation. The training data included simulations with noise modulation periods of 1–64s. We considered GLMs with different lengths of spike history filters. Cycle-averaged responses of HH and GLM simulations appear qualitatively similar (Figure 8), and thus we aimed to characterize how well the GLM fits captured the fractional



differentiation properties of the HH neuron. Although the AHP conductances act to provide a simple linear filtering of the spike train similar to the GLM's spike history filter, the GLM effectively assumes that the spike history is instead a current such that the total conductance of the cell remains constant (Latimer et al., 2019b). Therefore, it is not given that the GLM can replicate the computational effects of the AHP conductances.

The sinusoidal noise simulations show two properties of fractional differentiation. First, we estimated response gain (i.e., the strength of the sinusoidal modulation in the cycle-averaged response as a function of stimulus period; Figure 9A). In an ideal fractional differentiator, the log gain is proportional to the log of the stimulus period. The HH neuron shows a near linear response ($r^2 = 0.99, p < 10^{-4}$). Although the GLM with short history shows an almost flat relationship, increasing the spike history length shows similar slope to the HH neuron. The second property was the phase lead of the cycle-averaged response relative to the stimulus (Figure 9B). The phase lead should be constant under perfect fractional differentiation. The

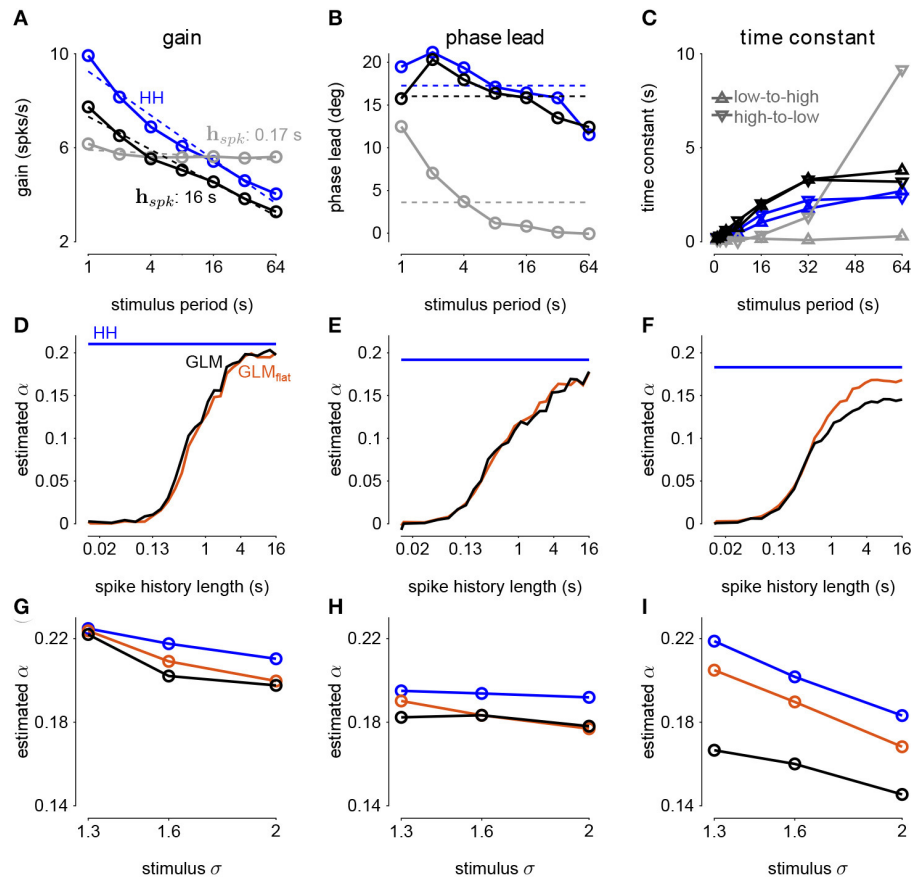


FIGURE 9 | (A) The gain of the average responses to sine-wave modulated noise as a function of stimulus period for the HH and GLMs (Figures 8A,B right). The GLM fit with a 0.17s spike history filter (gray) is compared to the GLM with the full 16s spike history (black). The HH simulation is given in blue. The noise stimulus was modulated with $\sigma = 2$. **(B)** The phase lead of the average responses to sine-wave modulated noise as a function of stimulus period for the HH and GLMs. The fraction of variance explained in the HH phase lead curve by the GLM with 16s spike history was $R^2 = 0.61$. **(C)** The time constant of an exponential function fit to the cycle-averaged response to square wave noise for each stimulus period (Figures 8C,D right). The markers denote time constants estimated for steps from low to high variance or step from high to low variance. The fraction of variance explained of the log time constants of the HH simulation by the GLM with 16s spike history was $R^2 = 0.80$. **(D)** The fractional differentiation order (α) of the GLM estimated by the slope of gain as a function of the log stimulus period in **(B)**. The value is estimated for each spike history lengths (black) and compared to α estimated from the HH simulation (blue). The red trace shows α estimated from the GLM fit only to unmodulated noise. **(E)** α estimated by the average phase lead across stimulus periods. **(F)** α estimated by fitting a the square-wave responses with a fractional differentiating filter. **(G–I)** α estimated at different noise modulation strengths for the 16s spike history GLM and HH simulation.

phase lead declines with longer period, but the HH simulation still shows strong phase lead in a 64s period. Short spike history filter GLMs exhibit a phase lead that tends to zero with long SD periods. However, the GLM fit with a long spike history filter closely tracks the HH neuron’s phase lead.

The final signature of fractional differentiation was the exponential decay of the cycle-averaged response under square-wave noise simulation (Figure 9C). We estimate the time constant of the decay on the square noise cycle average for both steps up and steps down in stimulus SD. The time constant increases approximately linearly with the SD period, and GLMs with long spike history showed time constants closely approximated the HH neuron.

From each signature, we estimated the order of the fractional differentiation (α) in both the HH neurons and the GLM fits. We estimated the order using the slope of log-period compared

to log-gain and mean phase lead across all stimulus periods for the sine-wave SD simulations (Figures 9D,E). A least-squares fit of FD filter of order α was applied to the square noise stimuli (Figure 9F). We considered α for the GLM fits as a function of the spike history length. The order estimates for the HH neuron, although slightly different for each signature, were approximately $\alpha = 0.2$. The GLM’s FD order tends toward that of the HH neuron as the spike history length increases from below. Surprisingly, when we considered a GLM trained only to a flat noise stimulus (no sine or square modulation; stimulus SD $\sigma = 1.0$) showed similar estimates of α (Figures 9D–F, red traces). Thus, the response properties giving rise to fractional differentiation of the noise envelope could be detected by the GLM even without driving with long timescale noise modulation.

We then considered how the estimated fractional differentiation order depended on the strength of the SD

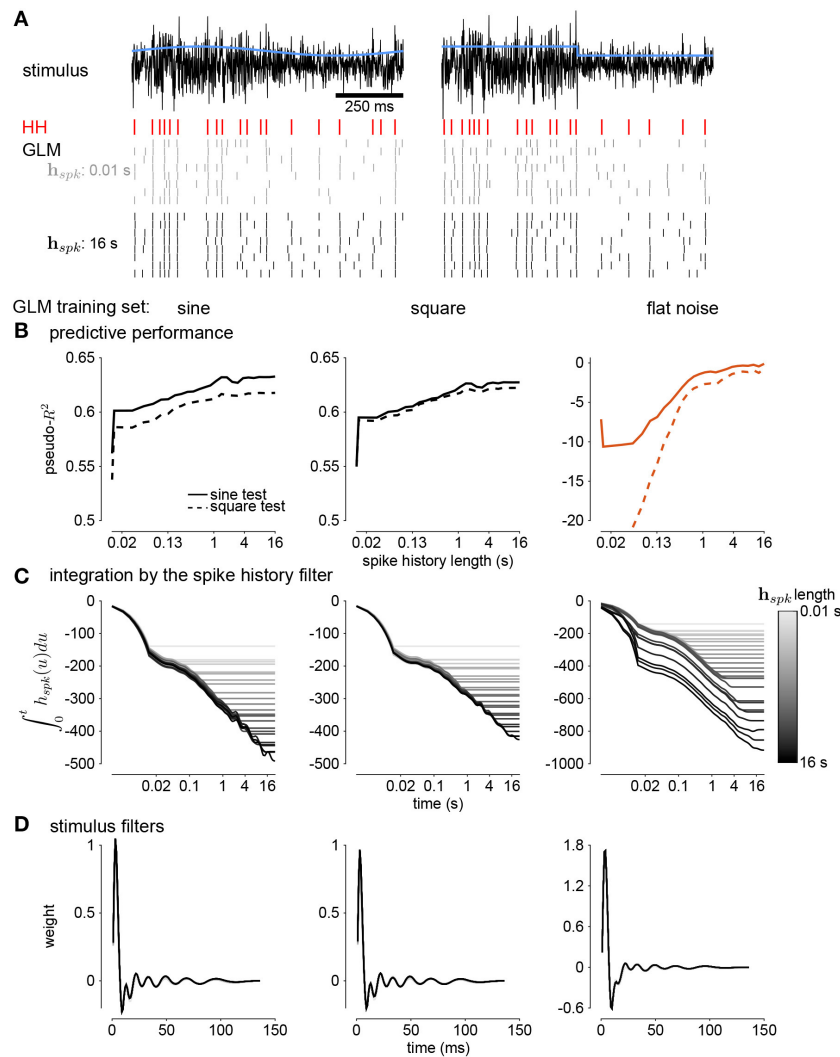


FIGURE 10 | (A) Comparing simulations of the HH model to sine wave modulated (left) and square wave modulated (right) noise. The black traces show the stimulus and the blue traces show the standard deviation envelope. The red raster denotes the HH spike response. Several repeated simulations of the GLM are shown for a GLM trained with only a 10ms spike history (gray raster) and a GLM trained with the full 16s spike history (black raster). **(B)** Assessing the model fitness for GLMs fit to sine wave modulated noise (left), square wave noise (middle), and unmodulated noise (right). The pseudo- R^2 measured on a withheld training set simulated from the HH model as a function of spike history length. The stimulus was sine (solid lines) or square (dashed lines) modulated noise with a 4s period and a modulation strength of $\sigma = 2$. **(C)** The integral over time (i.e., cumulative sums) of each spike history filter. **(D)** The stimulus filters for all GLM fits.

modulation. We found a slightly higher α for lower stimulus SDs (note that $\sigma = 2.0$ was used to fit the GLMs) for the gain and timescale estimates (Figures 9G–I). However, the phase lead estimate was fairly stable across SDs.

Next, we quantified how well the GLM predicted the HH responses to new stimuli (Figure 10A). Spike history filters with timescales of several seconds improved the GLM’s ability to predict spike trains, and the improvement continued for spike histories of several seconds (Figure 10B). However, training only on unmodulated noise did not result in a good GLM fit despite predicting α (Figure 10B).

We examined the parameter estimates in the GLM as a function of spike history length. We plotted the integral of

the spike history filter to show how the filter integrates spikes over time. The integrals show long timescales for the GLM fit to either sine- or square- wave noise (Figure 10C). The GLM fit to either type of noise predicted over 60% of the variance in the HH responses to both sine- and square-wave noise. The flat noise GLM also showed long timescales, but the integral changed substantially with changes in the length of spike history. This indicates that the combination of spike-history dependent timescales is not well-constrained in the flat noise condition despite predicting α , perhaps due to biases present in the data without modulations (Stevenson, 2018). The stimulus filters are short timescale and showed little dependence on spike history length (Figure 10D). Thus, the GLM captured fractional

differentiation in the HH neuron by linearizing the long timescale AHP currents.

4. DISCUSSION

Individual neurons can adapt their responses to changes in input statistics. Here, we studied two adaptive computations to changes in the stimulus variance that are captured by biophysically realistic neurons. First, we examined gain scaling of the inputs so that the spike-triggered stimulus distribution was independent of the stimulus variance. The ability of the neuron to gain scale depended on the ratio of the spike-generating potassium and sodium conductances. Second, we considered spiking responses that approximate a fractional derivative of the stimulus standard deviation, which can be produced by a set of AHP currents with different timescales. Although HH neurons can produce these adaptive effects, it is difficult to fit the HH to data.

Our results demonstrate that the GLM provides a tractable statistical framework for modeling adaptation to stimulus variance in single-neurons. The GLM provides an alternative representation of the spiking responses as two linear filters (stimulus and spike history filters) with a fixed spiking non-linearity instead of a multidimensional (and potentially stochastic) dynamical system (Meng et al., 2011, 2014). Importantly, a single GLM could accurately approximate the responses of HH neurons across multiple levels of input variance or across multiple timescales of variance modulation. The GLM accomplished this by linearizing the effect of recent spiking into a non-linear and stochastic spiking mechanism to adjust for the current stimulus statistics, which can act as a measure of the current stimulus context. To reproduce gain scaling, around 150ms of spike history is required, in line with the rapid expression of the gain scaling property with changes in stimulus statistics (Fairhall et al., 2001a; Mease et al., 2013). In the fractional derivative case, the GLM summarized the multiple AHP currents of the HH models as a single linear autoregressive function with multiple timescale effects.

This approach shows that at least part of the spectrum of adaptive behaviors to stimuli with time-varying characteristics can be captured through a single linear spike history feedback. Effective alternative approaches have captured time-varying context with a multiplicity of filters acting on different timescales of the stimulus alone (Kass and Ventura, 2001; McFarland et al., 2013; Qian et al., 2018; Latimer et al., 2019a).

REFERENCES

- Agüera y Arcas, B., and Fairhall, A. L. (2003). What causes a neuron to spike? *Neural Comput.* 15, 1789–1807. doi: 10.1162/08997660360675044
- Agüera y Arcas, B., Fairhall, A. L., and Bialek, W. (2003). Computation in a single neuron: Hodgkin and huxley revisited. *Neural Comput.* 15, 1715–1749. doi: 10.1162/08997660360675017
- Anastasio, T. J. (1998). Nonuniformity in the linear network model of the oculomotor integrator produces approximately fractional-order dynamics and more realistic neuron behavior. *Biol. Cybernet.* 79, 377–391. doi: 10.1007/s004220050487
- Benjamin, A. S., Fernandes, H. L., Tomlinson, T., Ramkumar, P., VerSteeg, C., Chowdhury, R. H., et al. (2018). Modern machine learning as a

The simulations explored here assumed the input to a cell was an injected current generated from a Gaussian distribution. However, neurons receive input as excitatory and inhibitory conductances, which can be integrated across complex dendritic processes (Ozuysal et al., 2018; Latimer et al., 2019b). Additionally, realistic input statistics may not follow a Gaussian distribution (Heitman et al., 2016; Maheswaranathan et al., 2018). Further work toward understanding the adaptive computations performed by single neurons should consider the inputs the neuron receives within a broader network and should consider non-linear stimulus processing (McFarland et al., 2013; Benjamin et al., 2018).

Neural coding and computations that occur across a wide range of input levels depend heavily on adaptation to the stimulus variance (Wark et al., 2007). The GLM, despite being a simple approximation, can provide a good representation of adaptive computations in biophysically realistic neurons.

DATA AVAILABILITY STATEMENT

All simulations are publicly available at <https://github.com/latimerk/GainScalingGLM>.

AUTHOR CONTRIBUTIONS

KL and AF designed the study and wrote the manuscript. KL performed the simulations and statistical analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Human Frontiers in Science Program to AF. KL was supported by a Chicago Fellowship.

ACKNOWLEDGMENTS

We thank Jonathan Pillow and Misha Proskurin for their help in developing this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnsys.2020.00060/full#supplementary-material>

benchmark for fitting neural responses. *Front. Comput. Neurosci.* 12:56. doi: 10.3389/fncom.2018.00056

Buhry, L., Grassia, F., Giremus, A., Grivel, E., Renaud, S., and Saïghi, S. (2011). Automated parameter estimation of the Hodgkin-Huxley model using the differential evolution algorithm: application to neuromimetic analog integrated circuits. *Neural Comput.* 23, 2599–2625. doi: 10.1162/NECO_a_00170

Cameron, A. C., and Windmeijer, F. A. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *J. Econometr.* 77, 329–342. doi: 10.1016/S0304-4076(96)01818-0

Cserssik, D., Hangos, K. M., and Szederkényi, G. (2012). Identifiability analysis and parameter estimation of a single Hodgkin-Huxley type voltage dependent ion channel under voltage step measurement conditions. *Neurocomputing* 77, 178–188. doi: 10.1016/j.neucom.2011.09.006

- Fairhall, A. (2014). "Adaptation and natural stimulus statistics," in *The Cognitive Neurosciences, Chapter 26*, eds S. M. Gazzaniga and R. G. Mangun (Cambridge, MA: MIT Press), 283–293.
- Fairhall, A. L., Lewen, G. D., Bialek, W., and de Ruyter van Steveninck, R. R. (2001a). Efficiency and ambiguity in an adaptive neural code. *Nature* 412:787. doi: 10.1038/35090500
- Fairhall, A. L., Lewen, G. D., Bialek, W., and de Ruyter van Steveninck, R. R. (2001b). "Multiple timescales of adaptation in a neural code," in *Advances in Neural Information Processing Systems*, eds T. K. Leen, T. G. Dietterich, and V. Tresp (Cambridge MA: MIT Press), 124–130.
- Heitman, A., Brackbill, N., Greschner, M., Sher, A., Litke, A. M., and Chichilnisky, E. (2016). Testing pseudo-linear models of responses to natural scenes in primate retina. *bioRxiv* 045336. doi: 10.1101/045336
- Kass, R. E., and Ventura, V. (2001). A spike-train probability model. *Neural Comput.* 13, 1713–1720. doi: 10.1162/08997660152469314
- Lankarany, M., Zhu, W.-P., and Swamy, M. (2014). Joint estimation of states and parameters of Hodgkin-Huxley neuronal model using Kalman filtering. *Neurocomputing* 136, 289–299. doi: 10.1016/j.neucom.2014.01.003
- Latimer, K. W., Barbera, D., Sokoletsky, M., Awwad, B., Katz, Y., Nelken, I., et al. (2019a). Multiple timescales account for adaptive responses across sensory cortices. *J. Neurosci.* 39, 10019–10033. doi: 10.1523/JNEUROSCI.1642-19.2019
- Latimer, K. W., Rieke, F., and Pillow, J. W. (2019b). Inferring synaptic inputs from spikes with a conductance-based neural encoding model. *eLife* 8:e47012. doi: 10.7554/eLife.47012
- Lundstrom, B. N., Higgs, M. H., Spain, W. J., and Fairhall, A. L. (2008). Fractional differentiation by neocortical pyramidal neurons. *Nat. Neurosci.* 11:1335. doi: 10.1038/nn.2212
- Maheswaranathan, N., Kastner, D. B., Baccus, S. A., and Ganguli, S. (2018). Inferring hidden structure in multilayered neural circuits. *PLoS Comput. Biol.* 14:e1006291. doi: 10.1371/journal.pcbi.1006291
- Mainen, Z. F., Joerges, J., Huguenard, J. R., and Sejnowski, T. J. (1995). A model of spike initiation in neocortical pyramidal neurons. *Neuron* 15, 1427–1439. doi: 10.1016/0896-6273(95)90020-9
- McCullagh, P., and Nelder, J. A. (1989). *Generalized Linear Models, Volume 37 of Monographs on Statistics and Applied Probability, 2nd Edn*. London: Chapman & Hall; CRC.
- McFarland, J. M., Cui, Y., and Butts, D. A. (2013). Inferring nonlinear neuronal computation based on physiologically plausible inputs. *PLoS Comput. Biol.* 9:e1003143. doi: 10.1371/journal.pcbi.1003143
- Mease, R. A., Famulare, M., Gjorgjieva, J., Moody, W. J., and Fairhall, A. L. (2013). Emergence of adaptive computation by single neurons in the developing cortex. *J. Neurosci.* 33, 12154–12170. doi: 10.1523/JNEUROSCI.3263-12.2013
- Mease, R. A., Lee, S., Moritz, A. T., Powers, R. K., Binder, M. D., and Fairhall, A. L. (2014). Context-dependent coding in single neurons. *J. Comput. Neurosci.* 37, 459–480. doi: 10.1007/s10827-014-0513-9
- Meng, L., Kramer, M. A., and Eden, U. T. (2011). A sequential monte carlo approach to estimate biophysical neural models from spikes. *J. Neural Eng.* 8:065006. doi: 10.1088/1741-2560/8/6/065006
- Meng, L., Kramer, M. A., Middleton, S. J., Whittington, M. A., and Eden, U. T. (2014). A unified approach to linking experimental, statistical and computational analysis of spike train data. *PLoS ONE* 9:e85269. doi: 10.1371/journal.pone.0085269
- Miller, K. D., and Troyer, T. W. (2002). Neural noise can explain expansive, power-law nonlinearities in neural response functions. *J. Neurophysiol.* 87, 653–659. doi: 10.1152/jn.00425.2001
- Murphy, B. K., and Miller, K. D. (2003). Multiplicative gain changes are induced by excitation or inhibition alone. *J. Neurosci.* 23, 10040–10051. doi: 10.1523/JNEUROSCI.23-31-10040.2003
- Oldham, K., and Spanier, J. (1974). *The Fractional Calculus Theory and Applications of Differentiation and Integration to Arbitrary Order*, Vol. 111. New York, NY: Elsevier.
- Ozuyal, Y., Kastner, D. B., and Baccus, S. A. (2018). Adaptive feature detection from differential processing in parallel retinal pathways. *PLoS Comput. Biol.* 14:e1006560. doi: 10.1371/journal.pcbi.1006560
- Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E., et al. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454:995. doi: 10.1038/nature07140
- Pozzorini, C., Naud, R., Mensi, S., and Gerstner, W. (2013). Temporal whitening by power-law adaptation in neocortical neurons. *Nat. Neurosci.* 16, 942–948. doi: 10.1038/nn.3431
- Qian, C., Sun, X., Zhang, S., Xing, D., Li, H., Zheng, X., et al. (2018). Nonlinear modeling of neural interaction for spike prediction using the staged point-process model. *Neural Comput.* 30, 3189–3226. doi: 10.1162/neco_a_01137
- Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W. (1999). *Spikes: Exploring the Neural Code*. Cambridge, MA: MIT Press.
- Rubner, Y., Tomasi, C., and Guibas, L. J. (1998). "A metric for distributions with applications to image databases," in *Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271)* (Mumbai: IEEE), 59–66. doi: 10.1109/ICCV.1998.710701
- Stevenson, I. H. (2018). Omitted variable bias in GLMs of neural spiking activity. *Neural Comput.* 30, 3227–3258. doi: 10.1162/neco_a_01138
- Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., and Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J. Neurophysiol.* 93, 1074–1089. doi: 10.1152/jn.00697.2004
- Vavoulis, D. V., Straub, V. A., Aston, J. A., and Feng, J. (2012). A self-organizing state-space-model approach for parameter estimation in Hodgkin-Huxley-type models of single neurons. *PLoS Comput. Biol.* 8:e1002401. doi: 10.1371/journal.pcbi.1002401
- Wark, B., Lundstrom, B. N., and Fairhall, A. (2007). Sensory adaptation. *Curr. Opin. Neurobiol.* 17, 423–429. doi: 10.1016/j.conb.2007.07.001
- Weber, A. I., and Fairhall, A. L. (2019). The role of adaptation in neural coding. *Curr. Opin. Neurobiol.* 58, 135–140. doi: 10.1016/j.conb.2019.09.013
- Weber, A. I., and Pillow, J. W. (2017). Capturing the dynamical repertoire of single neurons with generalized linear models. *Neural Comput.* 29, 3260–3289. doi: 10.1162/neco_a_01021

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Latimer and Fairhall. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.