A point process model for auditory neurons considering both their intrinsic dynamics and the spectro-temporal properties of an extrinsic signal

Eric Plourde[Memberm, IEEE],
Neuroscience Statistics Research Laboratory, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114 USA and also with the Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139

Bertrand Delgutte, and
Department of Otology and Laryngology, Harvard Medical School, Boston, MA 02114 USA and also with the Harvard–MIT Division of Health Science and Technology and the Research Laboratory of Electronics, MIT, Cambridge, MA 02139 USA

Emery N. Brown[Fellow, IEEE]
Neuroscience Statistics Research Laboratory, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114 USA and also with the Harvard–MIT Division of Health Science and Technology and the Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139 USA

Eric Plourde: eplourde@mit.edu; Bertrand Delgutte: Bertrand_Delgutte@meei.harvard.edu; Emery N. Brown: enb@neurostat.mit.edu

Abstract

We propose a point process model of spiking activity from auditory neurons. The model takes account of the neuron's intrinsic dynamics as well as the spectro-temporal properties of an input stimulus. A discrete Volterra expansion is used to derive the form of the conditional intensity function. The Volterra expansion models the neuron's baseline spike rate, its intrinsic dynamics - spiking history - and the stimulus effect which in this case is the analog of the spectro-temporal receptive field (STRF). We performed the model fitting efficiently in a generalized linear model framework using ridge regression to address properly this ill-posed maximum likelihood estimation problem. The model provides an excellent fit to spiking activity from 55 auditory nerve neurons. The STRF-like representation estimated jointly with the neuron's intrinsic dynamics may offer more accurate characterizations of neural activity in the auditory system than current ones based solely on the STRF.

Index Terms

Spike train model; point process; spectro-temporal receptive field; generalized linear model; auditory system

---

I. Introduction

Understanding the factors that are responsible for inducing neurons to spike is an important, active topic of investigation in neuroscience. One approach is to fit statistical models

---

Correspondence to: Eric Plourde, eplourde@mit.edu.

containing the most salient factors to neural spiking activity and use the fitted model to evaluate the relative importance of the factors. Two key factors or covariates to consider in standard neurophysiology experiments are the intrinsic dynamics of the neuron such as the absolute and relative refractory periods, bursting and network dynamics whereas the primary extrinsic factor is the external stimulus applied in the experiment. In an auditory experiment the stimulus is typically a specific sound pattern.

The intrinsic dynamics modeled typically in terms of the neuron's past spiking history has been established as an important descriptor of spiking propensity in a number of neural systems [1]–[5]. Analyses of auditory neurons have focused on constructing a spectro-temporal receptive field (STRF) by estimating a linear relation between a spectro-temporal representation of the auditory stimulus and the rate function of the neuron [6]. The coefficients of the linear model comprise the STRF. To date no statistical model has characterized the spiking propensity of auditory neurons by representing simultaneously the intrinsic dynamics and the complete spectro-temporal representation of the auditory stimulus. Given the point process nature of neural spiking activity, a principled approach to constructing the model would be to relate the covariates to the spiking propensity of the neuron in terms of the conditional intensity (rate) function (CIF) because a point process is completely described by its CIF.

We present a point process model for auditory neural spiking activity that considers both the neuron's intrinsic dynamics and the spectro-temporal properties of the auditory stimulus. We formulate the log of the conditional intensity function in terms of a discrete-time Volterra series expansion of the neuron's spiking history and the spectro-temporal decomposition of the auditory stimulus. The Volterra expansion contains a parameter representing the baseline spike rate, a set describing the intrinsic dynamics and a second set characterizing the stimulus effect, the analog of the STRF. Using the generalized linear model (GLM) in a ridge regression framework to address properly the ill-posed inverse nature of this maximum likelihood estimation problem we illustrate our approach by fitting the model to the spiking activity of 55 auditory nerve neurons in an anesthetized cat in response to an auditory stimulus.

This paper comprises the following sections: the proposed statistical model is derived in Section II; the model estimation is presented in Section III; results, including goodness of fit of the model and model parameter analyses, are presented in Section IV and Section V concludes the work.

II. The statistical model

Given an observation interval $(0, T]$ and spike times $0 < u_1, u_2, \ldots, < u_{N(T)} < T$. The CIF of the spike train is defined by [1]

$$\lim_{\Delta \to 0} \frac{\Pr\{[N(t+\Delta) - N(t)]=1|H_t\}}{\Delta} = \lambda(t|H_t) \tag{1}$$

where $N(t)$ is the number of spikes in the interval $(0, t]$ for $t \in (0, T]$ and H_t is the relevant history of the covariates at t. It follows that for Δ small

$$\Pr(\text{spike in}(t, t+\Delta]|H_t) \approx \lambda(t|H_t)\Delta. \tag{2}$$

The CIF is for a point process a history-dependent generalization of the rate function of a Poisson process. To obtain a discrete formulation of the CIF we choose K sufficiently large so that each subinterval $\Delta = K^{-1}T$ contains at most one spike. We index the subintervals $k = 1, \ldots, K$ and define n_k to be 1 if there is a spike in the subinterval $((k-1)\Delta, k\Delta]$ and 0 if there is no spike. For our analysis we choose K so that Δ is 1 millisecond, consistent with the absolute refractory period of a neuron.

Let $s_{k,j}$ be the value of a spectro-temporal representation of the sound stimulus with frequency band j at time $k\Delta$ for $j = 1, \ldots, J$. Define the relevant history of the sound stimulus for predicting the current spiking propensity as $H_{k,j} = \{s_{k,j}, \ldots, s_{k-L,j}\}$, assuming a dependence that goes back L time periods. Similarly, define the relevant spiking history for predicting the current spiking propensity as $H_{k,J+1} = \{n_{k-1}, \ldots, n_{k-P}\}$, assuming a dependence that goes back P time periods. Let $H_k = \{H_{k,1}, \ldots, H_{k,J+1}\}$. If we assume that there is a functional F which describes the relation between H_k and the CIF $\lambda(k\Delta|H_k)$ then we can expand the log of the CIF in a discrete Volterra series as [7]

$$
\begin{aligned}
\log(\lambda(k\Delta|H_k,\beta)) &= F(H_k,\beta) \\
&= \beta_0 + \sum_{j=1}^{J}\sum_{l=0}^{L-1}\beta_{l,j}s_{k-l,j} + \sum_{p=1}^{P}\beta_{p,J+1}n_{k-p} + \text{higher terms}
\end{aligned}
\tag{3}
$$

where $\beta = \{\beta_0, \beta_{0,1}, \ldots, \beta_{L-1,J}, \beta_{1,J+1}, \ldots, \beta_{P,J+1}\}$ is the $(JL + P + 1) \times 1$ vector of Volterra kernels. We interpret the Volterra series expansion as the sum of the outputs of $J + 1$ linear filters having Volterra kernels as the impulse responses. The kernels $\beta_{l,j}$ are the analogs of the STRFs used to characterize auditory neurons. The kernel $\beta_{p,J+1}$ models the effect of the spiking history and β_0 governs the mean spiking rate. Exponentiating both sides of (3) yields the CIF

$$
\lambda(k\Delta|H_k,\beta) = \exp\left\{\beta_0 + \sum_{j=1}^{J}\sum_{l=0}^{L-1}\beta_{l,j}s_{k-l,j} + \sum_{p=1}^{P}\beta_{p,J+1}n_{k-p}\right\}
\tag{4}
$$

where we have neglected the higher order terms.

Constructing the Volterra series expansion in terms of the log of the discrete CIF ensures that the CIF is non-negative and that the relation between the spiking activity and the sound stimulus and the spiking history can be modeled using a generalized linear model (GLM) with either a binomial or a Poisson link function [3]. Models with a similar structure as the one derived in (4) have been considered previously for other neural systems [1]–[5]. None of these analyses was motivated by a Volterra series expansion nor has this model been used in an analysis of auditory spiking activity in which the intrinsic dynamics and the STRF were simultaneously estimated.

III. Model estimation

We can rewrite (3) in a more compact form as

$$
\log(\lambda) = \mathbf{X}\beta
\tag{5}
$$

where $\mathbf{X}$ is the $RK \times (JL + P + 1)$ matrix of covariates, R has been added to take account of the number of trial and the logarithm function is applied element-wise. It follows that the log likelihood function for estimating $\boldsymbol{\beta}$ is [1]–[4]

$$\mathbb{L}(\beta)=\sum_{k=1}^{K}\sum_{r=1}^{R}n_{k,r}\log(\lambda_k|\beta) - \sum_{k=1}^{K}\sum_{r=1}^{R}(\lambda_k|\beta)\Delta$$

(6)

where the inner sum is over trials. An advantage of (5) is that it shows that (6) is equivalent to a GLM with a Poisson log likelihood function. We can therefore use the GLM framework to estimate $\boldsymbol{\beta}$.

It is well known that the Fisher scoring algorithm for GLM parameter estimation with canonical link functions can be solved by iteratively reweighted least-squares (IRLS) [8]. In the IRLS algorithm, the maximum likelihood estimate of $\boldsymbol{\beta}$ is computed iteratively by solving successive weighted least squares (WLS) subproblems. The WLS subproblems are solved using a conjugate gradient algorithm that greatly reduces the computational time. Our problem has an additional feature that must be considered. The design matrix $\mathbf{X}$ includes 1 ms delayed version of the input spectro-temporal representations. As a consequence, many of the columns of $\mathbf{X}$ are highly correlated, especially when the auditory stimulus is presented across a large number of trials. This suggests that the estimation of $\boldsymbol{\beta}$ is an ill-posed inverse problem which can be solved by regularization. We can estimate $\boldsymbol{\beta}$ using the truncated regularized iteratively reweighted least squares (TR-IRLS) algorithm for GLM parameter estimation [8]. The TR-IRLS is a variant of the IRLS algorithm in which a ridge parameter can be included in each WLS subproblem to provide a quadratic regularization.

Introducing the regularization term τ the estimate of $\boldsymbol{\beta}$ at iteration $i + 1$ of the TR-IRLS is computed as

$$\widehat{\beta}_{i+1}=(\mathbf{X}^T\mathbf{W}_i\mathbf{X}+\tau\mathbf{I})^{-1}\mathbf{X}^T\mathbf{W}_i\mathbf{z}_i$$

(7)

where $\mathbf{W}_i$ is an $RK \times RK$ diagonal weight matrix whose diagonal elements are $\boldsymbol{\lambda}_i$, the vector of CIF estimates at iteration i, $\mathbf{z}_i=\mathbf{X}\widehat{\beta}_i+\mathbf{W}_i^{-1}(\mathbf{n} - \boldsymbol{\lambda}_i)$ is the so-called adjusted dependent covariate [8] and $\mathbf{n}$ is the column vector of all of the $n_{k,r}$ spiking activity across all trials and across all times in the experiment.

IV. Results

We applied the proposed statistical model to neural spiking activity recorded in the auditory nerves of anesthetized cats following the presentation of the input sentence "Wood is best for making toys and blocks" spoken by a male voice and sampled at 10 kHz [9]. The dataset was composed of the spike train responses of 55 distinct neurons each recorded across $R = 20$ trials. The spectro-temporal representation of the input speech was obtained through a modified version of an auditory spectrogram [10] which applied a gammatone filterbank to the input speech signal. The bandwidths of the filters were modified according to [11] to represent adequately the processing performed by the cat's cochlea. To obtain comparable frequency bins, each was normalized by its Euclidean norm over the entire time domain. We used $P = 40$, $L = 10$ and $J = 25$, where the center frequencies ranged from 20 Hz to 4.4 kHz, and $\tau = 0.1$ was the regularization parameter in the TR-IRLS algorithm.

A. Model Goodness of fit

To evaluate the model goodness-of-fit, we used the time rescaling theorem with rescaled times computed from the estimated CIF [3]. If the latter is a good approximation to the true CIF of the point process, then the rescaled times will be independent and uniformly distributed on the interval [0, 1). We used Kolmogorov-Smirnov (KS) plots to assess the uniformity of the rescaled times and the autocorrelation function (ACF) of transformed rescaled times to assess their independence [3], [12].

Fig. 1 shows the KS plots of the best and worst fits in which the upper and lower 95% confidence bounds are indicated by dashed lines. If the model is accurate, the KS plot should show an empirical distribution $\hat{F}(x)$ versus the fitted cumulative distribution $F(x)$ that lies along the 45° line. As shown in Fig. 1a, the curve for the best fit is indistinguishable from the 45° line indicating a close model fit. Even for the worst case (Fig. 1b), the fitting is also quite good, with the KS plot being always within or extremely close to the confidence bounds.

To evaluate the fitting in the entire dataset, we used the normalized KS statistic which is defined as:

$$\widehat{D} = \sup \left| \frac{\widehat{F}(x) - F(x)}{\widehat{B}} \right|$$

(8)

where $\hat{B}$ is the width of 95% confidence bound. A value of $\hat{D} < 1$ thus indicates that the entire KS plot is within the 95% confidence bound. Table I shows the number of occurrences of normalized KS statistics values determined using 4-fold cross-validation. Forty-four of the 55 (80%) $\hat{D}$ values were less than 1 which supports the excellent fit of the models. Most of the $\hat{D}$ values that were greater than 1 were less than 1.05, suggesting that these models as well were in close agreement with the data.

The ACF of the transformed rescaled times were also evaluated to assess their independence. It was observed (results not shown here) that the rescaled times were highly independent with 92.5% of the ACF values in the entire auditory nerve dataset ($n = 55$) being inside their respective confidence bounds for lags up to 100 ms. The findings from the KS and ACF analyses suggest that the estimated CIFs provide excellent approximations to the true CIFs.

B. Model parameter analyses

We present the estimated values of the model parameters, i.e. the Volterra kernels β_0, $\beta_{p,J+1}$ and $\beta_{l,j}$ obtained using the 20 trials for each neuron recording.

Figure 2a shows a scatter plot of the estimated baseline parameter $\exp(\beta_0)$ versus the baseline firing rate measured in [9]. As can be observed, there is a strong correlation between $\exp(\beta_0)$ and the measured baseline firing rate (Pearson correlation coefficient $\rho = 0.92$).

The mean values of the exponentiated history parameters $\beta_{p,J+1}$ with the error bars of the 95% confidence interval is plotted in Fig. 2b. This plot shows that these parameters accurately capture the refractory behavior of the neurons because the mean value of $\exp(\beta_{p,J+1})$ is appreciably less than 1.

The values of the exponentiated stimulus parameters $\beta_{l,j}$ for a neuron with a characteristic frequency of 1024 Hz are shown in Fig. 2c. This representation resembles closely that of an

STRF. It has a well defined preferred spectro-temporal region that is fairly restricted for auditory nerve recordings. Moreover, the center frequency of the filter with the highest parameter value (1140 Hz) corresponds to the measured characteristic frequency (1024 Hz). We plot in Fig. 2d the center frequency l of the filter corresponding to the spectro-temporal parameter $\beta_{l,j}$ with the highest value as a function of the measured characteristic frequency for the entire dataset. There is a high correlation between the two ($\rho = 0.82$) indicating a good description of the CF by the model.[1]

V. Conclusion

We presented a point process model of neural spiking activity in the auditory system that takes account of both the neuron's intrinsic dynamics and the spectro-temporal properties of an input sound stimulus. We derived the associated CIF using a discrete Volterra expansion formulated in terms of the spiking history and the spectro-temporal components. This CIF model makes it possible to assess the relative importance of the neuron's intrinsic dynamic (spiking history) and the STRF (spectro-temporal components). We fit the model to actual auditory nerve spiking activity by regularized maximum likelihood estimation. The models gave accurate descriptions of neural spiking activity in terms of KS goodness-of-fit analyses. This model, which considers both the neuron's intrinsic dynamics and the STRF, may offer a way of obtaining more accurate characterizations of activity in other parts of the auditory system.

Acknowledgments

References

1. Brown, EN.; Barbieri, R.; Eden, UT.; Frank, LM. Likelihood methods for neural spike train data analysis. In: Feng, J., editor. Computational neuroscience: a comprehensive approach. London: CRC Press; 2003. p. 253-286.

2. Paninski L. Maximum likelihood estimation of cascade point-process neural encoding models. Network: Comput Neural Syst. 2004; 15:243–262.

3. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. J Neurophysiol. 2005; 93:1074–1089. [PubMed: 15356183]

4. Sarma SV, Eden UT, Cheng ML, Williams ZM, Hu R, Eskandar E, Brown EN. Using point process models to compare neural spiking activity in the subthalamic nucleus of parkinson's patients and a healthy primate. IEEE Trans Biomed Eng. 2010; 57(6):1297–1305. [PubMed: 20172804]

5. Truccolo W, Hochberg LR, Donoghue JP. Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. Nature Neurosci. 2010; 13(1):105–111. [PubMed: 19966837]

6. Theunissen FE, Sen K, Doupe AJ. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J Neurosci. 2000; 20(6):2315–2331. [PubMed: 10704507]

7. Marmarelis, VZ. Nonlinear dynamic modeling of physiological systems. Piscataway, N.J: IEEE Press; 2005.

[1]Additional examples of exponentiated $\beta_{l,j}$ values as well as estimated CIF for different neurons from the dataset can be found at http://www.neurostat.mit.edu/publications.

8. Komarek, P. PhD thesis. Carnegie Mellon University; 2004. Logistic Regression for Data Mining and High-Dimensional Classification.

9. Delgutte, B.; Hammond, BM.; Cariani, PA. Neural coding of the temporal envelope of speech: Relation to modulation transfer functions. In: Palmer, AR.; Reese, A.; Summerfield, AQ.; Meddis, R., editors. Psychophysical and physiological advances in hearing. London: Whurr; 1998. p. 595-603.

10. Ellis, D. Gammatone-like spectrograms," 2009, web ressource. http://www.ee.columbia.edu/_dpwe/resources/matlab/gammatonegram/

11. Carney LH, Yin TCT. Temporal coding of resonances by low-frequency auditory nerve fibers: single fiber responses and a population model. J Neurophysiol. 1988; 60:1653–1677. [PubMed: 3199176]

12. Haslinger R, Pipa G, Brown EN. Discrete time rescaling theorem: determining goodness of fit for discrete time statistical models of neural spiking. Neural Computation. Oct; 2010 22(10):2477–2506. [PubMed: 20608868]
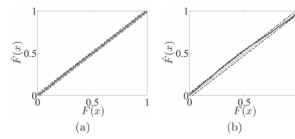
Fig. 1.
KS plots (black lines) of (a) the best and (b) worst fits from a set of 55 auditory nerve recordings. Dashed lines indicate the 95% confidence bound.
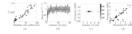
Fig. 2.
CIF parameter values. (a) Scatter plot of $\exp(\beta_o)$ vs. the measured baseline firing rate with Pearson correlation coefficient $\rho = 0.92$. (b) Mean of $e^{\beta_{p,J+1}}$, with 95% confidence interval, versus a latency p up to $P = 40$ ms (c) Example of exponentiated $\beta_{l,j}$ values displayed according to the center frequency of the corresponding gammatone filter j and latency l (CF = 1024 Hz) with $J = 25$ and $L = 10$ ms. (d) Scatter plot of the center frequency of the filter corresponding to the $\beta_{l,j}$ with the highest value versus the measured characteristic frequency of the corresponding neuron; $\rho = 0.82$.

TABLE I

Occurences of normalized KS statistic values $\hat{D}$ for the complete 55 auditory nerve dataset obtained using 4-fold cross validation.

$\hat{D}$ values	$0 - 0.5$	$0.5 - 1$	$1 - 1.5$	$1.5 - 2$	> 2
number of occurences	1	43	10	1	0