Published in final edited form as:

J Comput Neurosci. 2010 August ; 29(0): 193–201. doi:10.1007/s10827-009-0146-6.

A dynamical point process model of auditory nerve spiking in response to complex sounds

Andrea Trevino,

Department of Electrical & Computer Engineering Neuroscience Program, University of Illinois, Urbana, USA

Todd P. Coleman, and

Department of Electrical & Computer Engineering Neuroscience Program, University of Illinois, Urbana, USA

Jont Allen

Department of Electrical & Computer Engineering Neuroscience Program, University of Illinois, Urbana, USA

Andrea Trevino: atrevin2@illinois.edu; Todd P. Coleman: colemant@illinois.edu; Jont Allen: jontalle@illinois.edu

Abstract

In this paper, we develop a dynamical point process model for how complex sounds are represented by neural spiking in auditory nerve fibers. Although many models have been proposed, our point process model is the first to capture elements of spontaneous rate, refractory effects, frequency selectivity, phase locking at low frequencies, and short-term adaptation, all within a compact parametric approach. Using a generalized linear model for the point process conditional intensity, driven by extrinsic covariates, previous spiking, and an input-dependent charging/discharging capacitor model, our approach robustly captures the aforementioned features on datasets taken at the auditory nerve of chinchilla in response to speech inputs. We confirm the goodness of fit of our approach using the Time-Rescaling Theorem for point processes.

Keywords

Cochlea; Auditory nerve; Spiking model; Statistical model; Point process; Conditional intensity; Time rescaling theorem

1 Introduction

The cochlea is the organ of the auditory pathway that translates sound pressure waves into neural spike trains. A large body of detailed work over the last century has helped shed a significant amount of light on the specifics of cochlear processing. One of the earliest landmark discoveries was that a sound waveform is split into its individual frequency components as the input traveling wave moves along the basilar membrane of the cochlea.

This leads to a tonotopic decomposition (Von Békésy 1944) that is retained in some sense through the entire auditory pathway up to the auditory cortex. Besides this frequency decomposition, the cells of the auditory nerve exhibit varying spiking thresholds and spontaneous rates, as well as a series of nonlinear responses that include tone-on-tone masking (Wegel and Lane 1924), forward masking, and a dynamic-range amplification related to the workings of the outer hair cells (Dallos and Harris 1978).

Finding the correct model of this functionality would not only be beneficial to hearing aid design like cochlear implants, but would also help give a sense of what elements of a signal carry the information necessary for speech and music perception. Recent work by Shannon et al. (1995) has shown that only a few bands of speech modulated noise are enough for normal hearing listeners to hear intelligible speech. His work implies that temporal cues alone may carry a significant amount of the information necessary for speech perception. This observation immediately lends itself to the idea of using a model that is sensitive to precise temporal cues. A point process model is exactly this kind of model, with the basic concept being that at each point in time there is a time-varying, history and input-dependent probability of a spike.

Point process models have recently been applied to neural spiking patterns with a large amount of success. One such model is the low-complexity signal processing and inference algorithm that has been developed for characterizing place receptive fields of neurons in the hippocampus (Brown et al. 1998). We propose a point process framework that uses a parametric, generalized linear model approach (Truccolo et al. 2005) to model the conditional intensity function of spiking at the auditory nerve. We determine the parameter weights in the linear model by using maximum likelihood estimation to extract parameter values from recordings of chinchilla auditory nerve fibers. Such an approach gives us a compact model. We show that our model provides an accurate estimate of spiking distribution over time for neurons with high to medium spontaneous firing rates. For these neurons, the majority of the results fall within or at the 95 percent confidence interval of a statistical goodness-of-fit measure.

Other models of auditory nerve spiking do exist (Zhang et al. 2001) and statistics have already been applied to measure first-spike latency (Heil and Neubauer 2003). But neither of these models are able to focus on temporal characteristics for an extended complex stimulus in the way that our point process approach allows. Our model is computationally quite simple and only uses parameters that reflect the biological processes of the cochlea.

Each component of our model has a tie to well-known cochlear phenomena with a focus on being faithful to the biology and micromechanics of the system. These components include spontaneous firing rate, short-term refractory effects, frequency selectivity, phase locking and short-term adaptation in the inner hair cell as a result of its inherent capacitance. Our model parameters are determined through analysis of auditory nerve spiking data provided by Prof R. E. Wickesberg at the University of Illinois. This data is taken from the auditory nerve of anesthetized chinchillas. The full experimental procedure can be found in Stevens and Wickesberg (1999). We analyzed the spiking patterns of 23 neurons with

properties that span a large range of characteristic frequencies (CF), thresholds and spontaneous rates.

1.1 Background

Neural spiking at the auditory nerve is induced through displacement of the inner hair cells along the cochlea. The most basic feature of auditory nerve spiking that we take into account is the spontaneous firing rate. Measured rates from our data range from less than 1 to 142 spikes/second.

Due to its varying width and stiffness, each point along the basilar membrane resonates at a different frequency, decomposing the input signal (Von Békésy 1944) before the waveform even reaches the neural signaling pathways. The relationship between the spiking response and supra-threshold sound pressure level that is necessary to induce spiking is often represented as a tuning curve. For each place along the basilar membrane of the cochlea, the tuning curve for each neuron will be lowest (require the lowest input level to spike) and sharpest at the characteristic or best frequency. Each neuron has a unique tuning curve due to varying threshold levels, but since several neurons innervate each frequency-specific section of the cochlea, a robust representation of the input waveform is able to be encoded at the auditory nerve. This was first explained by Wever's volley theory in the 1930's (Wever and Bray 1937). This theory suggested that neurons with similar characteristic frequencies would fire out of phase in order to capture a more robust representation of the input signal. This would allow neural activity to track frequencies that exceed a neuron's maximum firing rate. In humans, the cochlea maps the frequencies from 20 Hz to 20 kHz in a continuum. Similarly, the hearing range of a chinchilla at 60 dB is approximately 50 Hz to 33 kHz (Heffner and Heffner 1991).

In the late 1970's, phase locking at the auditory nerve as a result of cochlear micromechanics was discovered (Johnson 1980). The structure of the organ of corti and the inner hair cells essentially half-wave rectifies an input waveform, meaning that neural spiking only occurs at the times when the waveform is positive. The term phase locking refers to the phenomena where the neural spiking activity synchronizes with the waveform such that it occurs at relatively the same point in the phase of the waveform for a particular frequency and level. The phase dependence on frequency and level is explored in detail by Anderson et al. (1971). In the chinchilla, the loss of phase locking begins around 1 kHz and is complete above 4–5 kHz. Hence, we assume in our model that all neurons with characteristic frequencies below 1 kHz exhibit phase locking and that the rest are only able to follow the general envelope of the input waveform.

The spiking response of auditory nerve fibers, in response to a single tone input, is characterized by a short burst of high frequency spiking at the onset followed by a significantly lower, slowly decreasing rate for the duration of the tone. We seek to capture this short-term adaptation by modeling the effects of the capacitance of the inner hair cell membrane on the current flow through the stereocilia.

2 Statistical modeling approach

2.1 Point process approach

Sensory neurons encode information about exogenous stimuli in their spiking patterns. Given that the randomness and variability in the width and height of each spike is on a much smaller order than the variability of the spike timings, many computational neuroscientists develop models where the information is encoded in the temporal sequence of spikes. As this is inherently stochastic, a natural way to model this is with the nerve spiking being a *point process*—a random process pertaining to the random times of events occurring.

Our point process model assumes that an exogenous input signal X , pertaining to a sound waveform, drives a system that generates a point process at the output. The system's propensity to spike is allowed to depend not only on the exogenous input X , but also its previous spiking. We first define

$$a^i \triangleq (a_1, \dots, a_i)$$

and consider discretizing time into $\Delta \simeq 0.001$ seconds.

The essence of the statistical structure of Y is the conditional intensity function $\lambda(i|H_i)$ (Daley and Vere-Jones 2003; Brown et al. 2003):

$$\begin{aligned} H_i &\triangleq \{y_1, \dots, y_i, x_1, \dots, x_i\} \\ \lambda(i|H_i) &\triangleq \lim_{\Delta \rightarrow 0} \frac{\mathbb{P}(y_{i+1}=1|H_i)}{\Delta} \end{aligned}$$

where $y_i = 1$ if a spike occurs in time bin i and equals 0 otherwise. The joint distribution of a realization $y^n \triangleq y$ of a point process, given an exogenous input $x^n \triangleq x$, is given by Daley and Vere-Jones (2003), Brown et al. (2003):

$$\mathbb{P}(y|x) \approx \exp \left\{ \sum_{i=1}^n \log(\lambda(i|H_i)) y_i - \lambda(i|H_i) \Delta \right\} \quad (1)$$

2.2 The dynamical point process model

The main focus of this work is finding a formulation for the conditional intensity $\lambda(i|H_i)$ that captures the statistical relationships between previous spiking, other system states (e.g. membrane voltage), and exogenous inputs.

We model $\log \lambda(i|H_i; \theta, \beta, \gamma)$ according to a parametric generalized linear model (GLM) that depends upon previous spiking, the sound pressure wave input, and other causal 'history' states that are predictive of future spiking.

$$\begin{aligned}
\log \lambda(i|H_i; \theta, \beta, \gamma) &= \theta_0 + \sum_{j=1}^J \theta_j y_{i-j} + \beta E_x(i) + \gamma C_x(i) \\
E_x(i) &= \text{bandpass filtered sound signal} \\
&\quad \text{envelope at time } i \\
C_x(i) &= \text{inner hair cell current at time } i
\end{aligned} \tag{2}$$

θ_0 contains the spontaneous firing rate; ($\theta_j : j = \{1, 2, \dots, J\}$) pertain to spiking history that capture refractory effects; β corresponds to the band-pass filtered signal, with the filter centered at the characteristic frequency (CF)—this quantifies how much the energy at this frequency contributes to the probability of spiking; finally, γ corresponds to the charging model which captures the short-term adaptation effects that are a result of the inherent capacitances and conductances of each hair cell.

$E_x(i)$ is calculated by band-pass filtering the input sound wave at CF. We use low-order Butterworth filters for this operation; the bandwidth of each neuronspecific filter is chosen by using the bandwidth measurements of Fletcher (1952) and were validated through likelihood techniques. Butterworth filters were chosen due to their linear phase characteristics. For frequencies below 1 kHz, the input signal is not half-wave rectified and is instead directly presented, to account for low probability of spiking during the troughs of the waveform—in other words, phase locking. If the center frequency of the filter is greater than 1 kHz, we further process the waveform by half wave rectifying and extracting the envelope to model loss of synchrony/phase locking. Specifically, by considering the continuous-time waveform $x(t)$ and the band-pass filter with impulse response $h_{CF}(t)$, we then have:

$$\begin{aligned}
[a]^+ &\triangleq \begin{cases} a & a \geq 0 \\ 0 & \text{otherwise} \end{cases} \\
e_x(t) &\triangleq \begin{cases} h_{CF} * x(t) & CF < 1 \text{ kHz} \\ [h_{CF} * x]^+ * h_{LP}(t) & CF \geq 1 \text{ kHz} \end{cases} \tag{3}
\end{aligned}$$

$$E_x(i) \triangleq e_x(i\Delta) \tag{4}$$

where $[a]^+$ is a half-wave rectification, h^{LP} performs envelope extraction and $*$ represents the convolution operation. This formulation models the loss of phase locking above 1 kHz by only preserving the waveform envelope beyond this point.

$C_x(i)$ is calculated by using the Weiss model as its basis (Allen 1983). This parameter is meant to capture short-term adaptation effects due to the capacitance of the inner hair cell membrane. We treat the potassium channels at the stereocilia as a variable resistor with a conductance of zero at the negative points of the waveform and a conductance that is linearly proportional to waveform amplitude at the positive points, as shown in Fig. 1 (R). This linear relationship is a reasonable estimation based on resistances observed in physiological experiments (Weiss et al. 1974). The resting potential of the hair cell V_H is initially set to -70 mV. The endocochlear potential V_E is set to 90 mV (Sewell 1984). In

series with the resistor at the stereocilia, the inner hair cell itself is modeled as a resistor R_M and capacitor C_M in parallel. R_M and C_M are set at 140 M Ω and 10 pF, respectively (Raybould et al. 2001). At the positive points of the input waveform, the hair cell is depolarized. This is modeled as an RC circuit composed of the input-dependent resistance at the stereocilia and C_M . During the negative points of the input waveform, the stereocilia act as an open circuit, effectively creating an RC circuit from the two elements in parallel with a time constant of 1.4 ms. Thus, C_M charges during the positive points of the input waveform and discharges at the negative points. The equations that define C_x follow from standard circuit theory:

$$x_{CF}(t) \triangleq h_{CF} * x(t) \quad (5)$$

$$last(t) \triangleq \sup\{\tau \leq t: \{x_{CF}(\tau_-) > 0, x_{CF}(\tau_+) = 0\} \cup \{x_{CF}(\tau_-) = 0, x_{CF}(\tau_+) > 0\}\}$$

$$v_C(t) = \begin{cases} v_{RC} \left(1 - e^{-(t-last(t))x_{CF}(t)/\tau_1}\right) & x_{CF}(t) > 0 \\ v_o \left(e^{-(t-last(t))/\tau_2}\right) & x_{CF}(t) \leq 0 \end{cases} \quad (6)$$

$$c_x(t) = C_m \frac{dv_C(t)}{dt} \quad (7)$$

$$C_x(i) \triangleq c_x(i\Delta) \quad (8)$$

where CF in Eq. (5) again corresponds to the characteristic frequency of the particular neuron; the upper part of Eq. (6) reflects charging of the inner hair cell at a rate proportional to the waveform input; and the lower part of Eq. (6) reflects discharging. v_{RC} represents the voltage across the stereocilia and hair cell and v_o is the initial voltage across C_M . The variable τ_1 is set so that the total time constant for the charging equation remains below .7 ms to match measurements of inner hair cell voltage (Allen 1983); τ_2 was 1.4 ms.

We use maximum likelihood estimation to find the optimal estimates of (θ, β, γ) :

$$(\hat{\theta}, \hat{\beta}, \hat{\gamma}) = \underset{\theta, \beta, \gamma}{\operatorname{argmin}} -\log \mathbb{P}(y|x; \theta, \beta, \gamma) \quad (9)$$

where $\mathbb{P}(y|x; \theta, \beta, \gamma)$ is given by Eq. (1) and $\lambda(i|H_i; \theta, \beta, \gamma)$ is given by Eq. (3). Because of the structure of the point process likelihood and the GLM structure of the conditional intensity, the maximum-likelihood problem Eq. (9) is convex with a unique optimal solution and can be implemented with iteratively weighted least squares. The variable J is chosen using the AIC model selection criterion (Akaike 1974).

3 Validation on experimental data

We develop a parametric generalized linear model for 23 neurons with the measured characteristic frequencies ranging from 100 Hz to 5453 Hz. The threshold levels for these neurons fall in the range of 3 dB to 41 dB SPL. The spontaneous firing rates range from 0 sp/s to 142 sp/s. Auditory nerve fiber responses were recorded while presenting the speech

inputs at varying decibel levels with periods of silence in between each speech input. We use the results for an input at 70 dB SPL for the figures in this section. A plot of the waveform input used for all of these results is shown in Fig. 2.

In this section, we first show how the model coefficients θ , β , γ give us a sense of the physical properties of each neuron. We then show that our point process model, summarized by Eqs. (1) and (3), gives an accurate (within a statistical 95 percent confidence interval) representation of the spiking pattern at the auditory nerve in response to a complex speech sound.

3.1 History coefficients

The history coefficients ($\theta_1 \dots \theta_j$) of the model (3) represent the effects of previous spiking (within the same neuron) on the current probability of spiking. Due to the refractory period, we often see strongly negative coefficients for the parameters that represent a spike in the most recent past (1–3 ms). Beyond the recent past, the magnitudes of the coefficient values drop significantly and may even become positive. An example of these coefficient values is shown in Fig. 3. Using AIC model selection, we found that, on average, only about 5 ms of previous spiking history was statistically relevant, with a maximum of 9 ms and a minimum of 2 ms.

3.2 Spontaneous rate

Our model gives a measure of the spontaneous firing rate in the parameter θ_0 .

Since θ_0 is the only variable that affects the probability of spiking given no waveform input or past spiking, we infer that this value can give us a sense of the neuron's actual spontaneous firing rate. We compare our spontaneous firing representation in the conditional intensity model θ_0 to the experimental spontaneous firing rate in Fig. 4 (Left). We see that the model has a small bias towards overestimation of the spontaneous rate, but is nonetheless accurate enough to capture the general statistics of the system, as shown in the following goodness of fit section.

3.3 Phase locking at low frequencies

The point process model is able to capture phase locking at low frequencies by combining a variety of elements, including previous spiking activity and the interplay between frequency envelopes, in determining the probability of a spike at each time point. The ability of the model to follow phase locking activity is demonstrated in Fig. 5. This figure compares the bandpassed waveform envelope at CF to an overlay of the conditional intensity and a scaled sum of the ensemble spikes at a single neuron for 50 repetitions of the stimulus. The scale factor on the spikes is included for the sake of visualization. Clearly, the neuron is most likely to spike at the most positive points of the envelope. We see that this is mirrored in the conditional intensity function, where the highest conditional intensities also correspond to the most positive points of the waveform envelope, showing a strong model representation of the phase locking effect.

3.4 Goodness of fit

We use a point process-specific statistical measure of goodness of fit in order to assess the model. More specifically, we implement a Komolgorov-Smirnoff (KS) test on the time-rescaled inter-spike intervals with respect to our estimated conditional intensity function. By the Time Rescaling Theorem (Brown et al. 2002; Daley and Vere-Jones 2003), if a perfect model were found, the re-scaled inter-spike intervals, given by

$$Z_i = \int_{T_{i-1}}^{T_i} \lambda(t|H_t) dt$$

of a point process with occurrence times $T_1, T_2, \dots$ are independent unit-rate exponentials, thus forming the epoch times of a unit-rate Poisson process. To assess how well our estimated conditional intensity has fit the data, we construct $\lambda(t|H_t)$ from our estimated coefficients (θ, β, γ) , and construct each $\hat{Z}_i$ accordingly. We subsequently transform each $\hat{Z}_i$ to uniform $[0, 1]$ random variables and perform a Komolgorov-Smirnoff (KS) plot. If our estimation procedure has been accurate and the point process model truly captured all elements of cochlear spiking at the auditory nerve, then the empirical CDF of the resulting $\hat{Z}_i$'s should fit the true CDF of a uniform distribution, that of a 45 degree line. KS plots for a variety of neurons, along with the 95% confidence intervals, are given in Figs. 6 and 7.

Figure 6 compares the spiking response pattern for one stimulus repetition to the model estimate of the conditional intensity function. We see that the periods over which the conditional intensity stays high matches those when the measured spiking rate is greatest. A closer look shows that small bursts in the conditional intensity often match small bursts in spiking, if slightly offset. These overlay plots show how the statistics of the complex phenomena that control spiking at the auditory nerve can be well approximated by a relatively simple point process model with appropriate parameters. The corresponding KS plots further show that the model estimate for these two neurons is accurate in terms of the statistics of the process, falling well within the 95% confidence interval.

4 Discussion

We have shown that our point process model is able to accurately predict the stochastic spiking activity that occurs at the auditory nerve in response to a complex speech sample. These results hold for neurons of high to medium spontaneous rates across a large range of characteristic frequencies (100–5453 Hz) and threshold levels (3–41 dB SPL). We tested the model at different set input levels (60, 70, and 90 dB SPL) and found that the model, when calculated for each level, was able to capture the statistical properties of the system within the designated 95% confidence interval for almost all cases. Thus, the statistical nature of the point process approach allows us to capture a wide range of neuron-specific characteristics, while still maintaining a simple, compact model.

The point process approach allows the statistics of the experimental data to determine the weights/ coefficients of the various parameters in the GLM model (3). This generates a neuron-specific representation that accounts for differences in frequency selectivity, threshold and spontaneous rate. The history coefficients give a direct sense the refractory

period and, more abstractly, any specific bursting pattern that might be characteristic of the specific neuron. The envelope at CF provides a sense of the observed input signal without the effects of nonlinear cochlear processing. The short term adaptation component of the model tracks a very well documented phenomena, the heightened response pattern at the onset of a stimulus.

Of note are the similarities between our model and those developed by Johnson (1974). The representation of spontaneous rate in the conditional intensity function θ_0 is analogous to a_0 in Johnson's rate function. In both cases, the estimated values vary from the experimentally measured values, suggesting that other elements besides spontaneous rate may be being encoded. As shown in Fig. 4, θ_0 consistently overestimates the log of the spontaneous rate by a small amount. Johnson instead found that a_0 deviated significantly from the log of the spontaneous rate and in fact was somewhat correlated to the stimulus frequency (Johnson 1974, Fig. 4.18). We do not find this correlation; these differences are most likely due to the statistical point process framework of our model.

In the future, we plan to add a representation of the outer hair cells to the model in order to better capture long-term adaptation and level-specific effects. As our model is statistically accurate while staying computationally simple, we see potential to apply this work to cochlear implant research.

Acknowledgments

The authors would like to first thank Bob Wickesberg and Hanna Stevens, for the rich dataset that made this work possible; Emery Brown for his advice and mentorship; Anne Dreyer for her input on modeling auditory spikes; and Bryce Lobbell, for his input on using critical bandwidth filters.

A. Trevino would like to acknowledge the financial support from the UIUC SURGE fellowship, the NIH-UIUC Sensory Neuroscience Training Grant, and the NSF Graduate Research Fellowship. T. P. Coleman would like to acknowledge the financial support from the AFOSR Complex Networks Program via award no FA9550-08-1-0079.

References

- Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 1974; 19(6):716–723.
- Allen, JB. A hair cell model of neural response. In: de Boer, E., et al., editors. *proceeding of the IUTAM; ICA symposium; Martinus, The Hague, and Delft UP, Delft, The Netherlands*. 1983. p. 193-202.
- Anderson DJ, Rose JE, Hind JE, Brugge JF. Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: Frequency and intensity effects. *The Journal of the Acoustical Society of America*. 1971; 49:1131. [PubMed: 4994692]
- 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: Chapman & Hall; 2003.
- Brown EN, Barbieri R, Ventura V, Kass RE, Frank LM. The time-rescaling theorem and its application to neural spike train data analysis. *Neural Computation*. 2002; 14:325–346. [PubMed: 11802915]
- Brown EN, Frank LM, Tang D, Quirk MC, Wilson MA. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *Journal of Neuroscience*. 1998; 18(18):7411–7425. [PubMed: 9736661]
- Daley, DJ.; Vere-Jones, D. *An introduction to the theory of point processes*. New York: Springer; 2003.

- Dallos P, Harris D. Properties of auditory nerve responses in absence of outer hair cells. *Journal of Neurophysiology*. 1978; 41(2):365–383. [PubMed: 650272]
- Fletcher, H. *Speech and hearing in communication*. New York: Van Nostrand; 1952.
- Heffner RS, Heffner HE. Behavioral hearing range of the chinchilla. *Hearing Research*. 1991; 52(1): 13–6. [PubMed: 2061202]
- Heil P, Neubauer H. A unifying basis of auditory thresholds based on temporal summation. *Proceedings of the National Academy of Sciences*. 2003; 100(10):6151–6156.
- Johnson, DH. PhD thesis. Massachusetts Institute of Technology; 1974. The response of single auditory-nerve fibers in the cat to single tones; synchrony and average discharge rate.
- Johnson DH. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *The Journal of the Acoustical Society of America*. 1980; 68:1115. [PubMed: 7419827]
- Raybould NP, Jagger DJ, Housley GD. Positional analysis of guinea pig inner hair cell membrane conductances: Implications for regulation of the membrane filter. *JARO-Journal of the Association for Research in Otolaryngology*. 2001; 2(4):362–376.
- Sewell WF. The relation between the endocochlear potential and spontaneous activity in auditory nerve fibres of the cat. *The Journal of Physiology*. 1984; 347(1):685–696. [PubMed: 6707972]
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. *Science*. 1995; 270(5234):303. [PubMed: 7569981]
- Stevens HE, Wickesberg RE. Ensemble responses of the auditory nerve to normal and whispered stop consonants. *Hearing Research*. 1999; 131(1–2):47–62. [PubMed: 10355604]
- 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. *Journal of Neurophysiology*. 2005; 93:1074–1089. [PubMed: 15356183]
- Von Békésy G. Über die mechanische Frequenzanalyse in der Schnecke verschiedener Tiere. *Akustik Zurnal*. 1944; 9:3–11.
- Wegel RL, Lane CE. The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Physical Review*. 1924; 23(2):266–285.
- Weiss TF, Mulroy MJ, Altmann DW. Intracellular responses to acoustic clicks in the inner ear of the alligator lizard. *The Journal of the Acoustical Society of America*. 1974; 55:606. [PubMed: 4819861]
- Wever EG, Bray CW. The perception of low tones and the resonance-volley theory. *Journal of Psychology*. 1937; 3:101–114.
- Zhang X, Heinz MG, Bruce IC, Carney LH. A phenomenological model for the responses of auditorynerve fibers: I. Nonlinear tuning with compression and suppression. *The Journal of the Acoustical Society of America*. 2001; 109:648. [PubMed: 11248971]

Fig. 1.

Inner Hair Cell Model (*left*): Circuit model of an inner hair cell $V_M = -70$ mV $V_E = 90$ mV, figure is an adaptation from (Weiss et al. 1974) (*right*): Stereocilia conductance $g(x)$ as a function of position x , showing that the conductance increases linearly during the positive points of the input waveform and goes to zero during the negative points of the waveform

Fig. 2.
Input Waveform—spoken ‘dodn’, followed by an equivalently long period of silence

Fig. 3.

Example of the estimated history coefficients for a neuron where J , the maximum amount of delay, equals 5

Fig. 4.
Comparison of experimental and model estimate spontaneous rates

Fig. 5.

(*Top*): Band-passed input waveform envelope at CF [180 Hz] (*Bottom*): Overlay plot of the model conditional intensity and a scaled ensemble sum of spikes for 50 stimulus repetitions

Fig. 6. (Left): CF 824 Hz, spontaneous rate 41 spks/s, threshold 11 dB SPL (right): CF 1921 Hz, spontaneous rate 39 spks/s, threshold 12 dB SPL (Top): Ensemble sum of experimentally recorded spikes over 50 repetitions (Middle): The model's estimated conditional intensity function (Bottom): Corresponding goodness of fit plots (input waveform shown in Fig. 2)

Fig. 7. Goodness of fit plots that correspond to neurons with characteristic frequencies: (*top*) 504 Hz (*middle*) 1111 Hz (*bottom*) 4685 Hz