Multidimensional stimulus encoding in the auditory nerve of the barn owl

Brian J. Fischer,¹ Jacob L. Wydick,¹ Christine Köppl,² and José L. Peña³

¹Department of Mathematics, Seattle University, Seattle, Washington 98122, USA

²Cluster of Excellence “Hearing4all” and Research Centre Neurosensory Science, Department of Neuroscience, School of Medicine and Health Science, Carl von Ossietzky University, Oldenburg, Germany

³Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, New York, New York 10461, USA

(Received 20 April 2018; revised 7 September 2018; accepted 10 September 2018; published online 12 October 2018)

Auditory perception depends on multi-dimensional information in acoustic signals that must be encoded by auditory nerve fibers (ANF). These dimensions are represented by filters with different frequency selectivities. Multiple models have been suggested; however, the identification of relevant filters and type of interactions has been elusive, limiting progress in modeling the cochlear output. Spike-triggered covariance analysis of barn owl ANF responses was used to determine the number of relevant stimulus filters and estimate the nonlinearity that produces responses from filter outputs. This confirmed that ANF responses depend on multiple filters. The first, most dominant filter was the spike-triggered average, which was excitatory for all neurons. The second and third filters could be either suppressive or excitatory with center frequencies above or below that of the first filter. The nonlinear function mapping the first two filter outputs to the spiking probability ranged from restricted to nearly circular-symmetric, reflecting different modes of interaction between stimulus dimensions across the sample. This shows that stimulus encoding in ANFs of the barn owl is multidimensional and exhibits diversity over the population, suggesting that models must allow for variable numbers of filters and types of interactions between filters to describe how sound is encoded in ANFs. © 2018 Acoustical Society of America.

<https://doi.org/10.1121/1.5056171>

[ICB]

Pages: 2116–2127

I. INTRODUCTION

A central question in neuroscience is how sensory signals are encoded in the spiking responses of primary sensory neurons. Here we focus on how sounds are encoded in the auditory nerve of the barn owl. Previous work described the responses of barn owl auditory-nerve fibers (ANF) to tones and noise stimuli in order to characterize frequency tuning, phase locking, and level sensitivity in these neurons (Köppl, 1997a,b; Köppl and Yates, 1999; Fontaine *et al.*, 2015). Together, these descriptions define individual parameters of a model for how the auditory nerve encodes sounds. However, a signal-processing model has not been explicitly tested in the barn owl.

Studies in mammalian systems have used a variety of modeling approaches to characterize the signal-processing occurring at the level of the auditory nerve. A fundamental component of many models of ANF responses in mammals is the use of multiple stimulus filters to describe frequency selectivity (Goldstein, 1990; Meddis *et al.*, 2001; Zhang *et al.*, 2001; Sumner *et al.*, 2002; Bruce *et al.*, 2003; Tan and Carney, 2003; Irino and Patterson, 2006; Zilany and Bruce, 2006). Indeed, having multiple stimulus filters whose outputs can interact nonlinearly has proven useful for modeling complex phenomena such as two-tone rate suppression and combination tones (Robles and Ruggero, 2001; Heil and Peterson, 2015).

Auditory-nerve models often include a specified number of stimulus filters with parameters that may be fit to

physiological data. On the other hand, approaches have been used to approximate the nonlinear input-output properties of auditory-nerve fibers using Wiener-Volterra expansion modeling, which allows relevant stimulus dimensions to be discovered from the data (Eggermont *et al.*, 1983a,b; Eggermont *et al.*, 1983c; Eggermont, 1993; van Dijk *et al.*, 1994; Yamada and Lewis, 1999). This type of approach is more flexible than using a parametric model, and is particularly useful for analyses where the relevant stimulus dimensions and their interactions are not yet well understood, as is the case for the barn owl. While Wiener-Volterra models can identify multiple stimulus dimensions that are important for the responses, they are typically limited to finding a single quadratic approximation to the spiking nonlinearity (Marmarelis and Marmarelis, 1978). This is a strong limitation, as the quadratic approximation fails to describe nonlinear responses such as half-wave rectification or divisive normalization and thus limits the ability to assess interactions between multiple filters. Therefore, a more flexible approach is required to identify nonlinearities that are expected from physiological responses.

We sought a model of barn owl ANF responses that identifies the number of relevant stimulus dimensions and the interactions along these dimensions. The relevant stimulus dimensions are each represented as a separate filter that is applied to the stimulus. The frequency-selectivity of the filters, including whether they are bandpass or broadband filters, is determined from the data. We applied spike triggered

covariance (STC) analysis to responses to Gaussian white noise sounds to identify the number of relevant stimulus dimensions (filters) and their properties (Fig. 1) (van Steveninck and Bialek, 1988; Bialek and van Steveninck, 2005; Schwartz *et al.*, 2006; Aljadeff *et al.*, 2016). We found that ANFs respond to multiple relevant stimulus dimensions and that stimulus components along these dimensions interact to drive spiking responses in varied nonlinear forms over the population.

II. METHODS

Experiments were performed on six adult barn owls (five *Tyto alba* and one *Tyto furcata*) (range from 3.5 months to 4 years of age and both sexes) in two laboratories, but using largely identical procedures. The procedures complied with guidelines set forth by the National Institutes of Health and were approved by the Albert Einstein College of Medicine's Institute of Animal Studies, and the Animal Ethics Authority (LAVES) of Lower Saxony, Germany, respectively.

A. Anesthesia and surgery

General anesthesia was induced and maintained by intramuscular injections of ketamine hydrochloride (5–20 mg/kg) and xylazine (2–4 mg/kg). Maintenance of body temperature was either assisted by a heating pad underneath the body of the owl (American Medical Systems) or held at 39°C by a feedback-controlled blanket system (Harvard Instruments). Animals breathed unaided, however, the youngest owl was intubated through a tracheotomy to avoid potential problems with salivation. A metal piece was fixed to the skull with dental cement, to hold the head firmly. The bone and meninges overlying the right cerebellum were removed, and the posterior part of the right cerebellum was aspirated to expose the surface of the auditory brainstem on that side.

FIG. 1. (Color online) Auditory nerve model structure. A sound stimulus at a fixed level is filtered with a bank of filters that is determined by STC analysis. The first filter is the STA. The remaining filters are significant STC filters that are different from the STA. The total number of filters is the number of stimulus dimensions encoded by the neuron. The spiking nonlinearity is estimated from the data by finding the proportion of times the filtered stimuli elicit a spike. This schematic shows only the two-dimensional nonlinearity for the stimulus projections along the STA and first STC filters. The color indicates the spiking probability. Spikes are produced as Bernoulli random variables with a spiking probability given by the spiking nonlinearity at each time.

Auditory nerve fibers were recorded at the point of entry to the brainstem. The electrode positioning was visually guided. In this portion of the owl's brainstem, the only other auditory neurons that are occasionally encountered belong to the cochlear nucleus angularis, known for their lack of phase locking (Sullivan and Konishi, 1984), or low-frequency regions of the nucleus magnocellularis that are distinguished by higher spontaneous discharge rates and longer click latencies (Köppl, 1997b). These criteria were routinely checked.

B. Acoustic stimulation

All recordings were performed in a double-walled sound-attenuating chamber (Industrial Acoustics Corporation). Closed sound systems were inserted into both ear canals and consisted of a miniature speaker (Knowles 1914, Aiwa HP-V14, or Etymotic ER-2) and a microphone (Knowles 1319 or Knowles FG 23329), each contained in a custom-made housing that fit the owl's ear canal. Previous calibration of the Knowles microphones with a Brüel and Kjaer (model 4190 or 4134) microphone made it possible to translate the voltage output of the Knowles microphone into sound level in dB sound pressure level (SPL). The Knowles microphones were used to calibrate the earphone assemblies at the beginning of each experiment. The calibration data contained the amplitudes and phase angles measured in frequency steps of 100 Hz. The stimulus generation software then used these calibration data to automatically correct irregularities in the amplitude and phase response of each earphone.

Tonal signals and Gaussian broadband (0.5–10 or 12 kHz) noise were synthesized by custom-written software and a signal processing board (AP2, Tucker-Davis Technologies) at a sampling rate of 48 kHz. Noise stimuli were generated by the computer via inverse fast Fourier transformation (FFT). Briefly, a random initial phase was generated for each frequency between lower and upper bounds of the frequency range. For all these frequencies, the same amplitude was chosen. Calibration was then applied by multiplying the amplitude and adding the phase for each frequency channel to compensate for magnitude and phase of the earphones, respectively. The signals were scaled to fit into the output range and avoid clipping and gated with 5 ms rise/fall times. Manual attenuators were used to adjust the overall sound levels in both ears; software-controlled attenuators (PA4, Tucker-Davis Technologies) varied the stimulus levels during data collection. The signals had a Gaussian-amplitude distribution that did not rely on the random number generator of the Tucker-Davis Tucker equipment.

C. Electrophysiology

Glass microelectrodes, filled with either 2 M Na-acetate or 3 M KCl and with impedances mostly between 30 and 60 MΩ, were positioned under visual control above the surface of the brainstem and then advanced under remote control. Neural signals were serially amplified by a Multiclamp 700B (Molecular Devices) or a WPI 767 (World Precision Instruments) and an AC amplifier (PC1, Tucker-Davis Technologies). A spike discriminator (SD1, Tucker-Davis Technologies) converted neural impulses into transistor-

transistor logic (TTL) pulses for an event timer (ET1, Tucker-Davis Technologies), which recorded the timing of the pulses. In parallel, the analog waveforms were stored in a personal computer via an analog-to-digital converter (DD1, Tucker Davis Technologies) with a sampling rate of 48 kHz and 16-bit resolution.

D. Data collection

Stimulation protocols were exclusively ipsilateral. After isolating the spikes of a single nerve fiber, several basic tests were run, using stimuli of 50 ms duration, 5 ms rise/fall times, presented at a rate of up to 5/s. Best frequency (BF) was estimated audiovisually by reducing the stimulus level until the response could only be discerned within approximately ± 200 Hz. Rate-level curves to broadband noise were recorded, in 5 dB steps, repeated 10 or 20 times. Rate-level curves were used to estimate sound-level threshold to noise and to determine the stimulus level to be used in subsequent noise presentations for model fitting and testing. The stimulus level was selected to be near saturation, based on the rate-level curve [mean = 58.6 dB SPL, standard deviation (s.d.) = 10.2 dB SPL].

Data for model fitting were obtained by presenting 200-ms *de novo*-synthesized Gaussian broadband noise signals (with power between 0.5 and 10 or 12 kHz). Stimuli were presented with an inter-stimulus interval of 500 ms, and we attempted to collect over 4000 spikes for each stimulus condition (6775 ± 2574 spikes, mean $\pm$ s.d.). We only included units where at least 2000 spikes were collected for model fitting. Data for model testing were obtained using a frozen noise protocol, in which a single 200-ms Gaussian broadband noise stimulus (with power between 0.5 and 10 or 12 kHz) was repeatedly presented. We attempted to collect frozen noise responses over 300 repetitions (296 ± 138 repetitions, mean $\pm$ s.d.).

E. Model fitting

We used spike-triggered analyses to model auditory-nerve fiber responses (de Boer and Kuyper, 1968; de Boer and de Jongh, 1978; van Steveninck and Bialek, 1988; Bialek and van Steveninck, 2005; Aljadeff *et al.*, 2016). The model consists of a linear filtering stage with a small number of filters, followed by a nonlinearity that specifies the probability of a spike being generated in each time bin. We used STC method to find the low-dimensional stimulus subspace that influences neural responses (Aljadeff *et al.*, 2016). Modeling fitting utilized computing resources provided by the Neuroscience Gateway (Sivagnanam *et al.*, 2013).

The stimulus at each time was defined as a segment of the sound at the current time and the preceding $N - 1$ time points:

$$s(t) = [x(t) x(t-1) \cdots x(t-(N-1))]^T, \quad (1)$$

where $x(t)$ is the sound at time t and superscript T indicates the transpose of a vector. The stimulus length N was determined by the length of the spike-triggered average (STA) stimulus. We first determined the time required to keep all

points where the STA was at least 5% of its maximum amplitude. The stimulus length N was then given by the number of time points required to extend 0.5 ms beyond the point where the STA was at least 5% of its maximum amplitude. The stimulus length was between 117 and 469 time points for all neurons. To ensure that segments of the inter-stimulus interval and the rise period of the stimulus were not included in the analysis (or, in other words, to guarantee that the analysis was done on a signal with stationary statistics), spikes that occurred in the 15 ms immediately following stimulus onset were excluded; this exclusion was also sufficient to guarantee that the onset transient was excluded and that the neuron had reached a stable firing rate, and we encountered no neurons that had onset-only responses.

The STC method analyzes the statistical properties of the set of stimuli that preceded a spike to determine the significant stimulus dimensions (van Steveninck and Bialek, 1988; Bialek and van Steveninck, 2005; Schwartz *et al.*, 2006; Aljadeff *et al.*, 2016). The STA is the mean of the stimuli that preceded a spike

$$\phi_{STA} = \frac{1}{n} \sum_{i=1}^n s(t_i), \quad (2)$$

where $t_1, t_2, \dots, t_n$ are the spike times. While the STA provides a single stimulus dimension that influences spiking, an analysis of the covariance matrix of the stimuli that preceded a spike can reveal additional stimulus dimensions that influence spiking. Additionally, STC analysis will identify relevant stimulus dimensions even in neurons with weak phase locking where the STA may not be well-defined. We looked for dimensions where the covariance matrix of the stimuli that preceded a spike differs from the covariance matrix of the entire set of stimulus segments $s(t)$ that were used for model fitting. The stimulus covariance matrix is

$$C_p = \frac{1}{K-1} \sum_{t=1}^K (s(t) - \bar{s})(s(t) - \bar{s})^T, \quad (3)$$

where there are K time points in the entire stimulus set and

$$\bar{s} = \frac{1}{K} \sum_{t=1}^K s(t) \quad (4)$$

is the mean stimulus. The mean stimulus will be zero here because the stimuli are drawn from a Gaussian distribution with zero mean. The STC matrix is

$$C_s = \frac{1}{n-1} \sum_{i=1}^n (s(t_i) - \phi_{STA})(s(t_i) - \phi_{STA})^T, \quad (5)$$

where, as above, $t_1, t_2, \dots, t_n$ are the spike times. We next compute eigenvalues and eigenvectors of the difference between the covariance matrices: $\Delta C = C_s - C_p$. The number of eigenvectors of ΔC is equal to the length of the stimulus N , but the response of the neuron likely depends on a much smaller set of vectors. The aim of this analysis is to find the small set of vectors that span the low-dimensional subspace

that is relevant for the response of the neuron. Those vectors that influence the response are the significant stimulus dimensions. To find significant stimulus dimensions, we look for eigenvalues that are outside the range of eigenvalues found in a null distribution. The null distribution is found by computing the eigenvalues of the difference between the stimulus covariance matrix and STC matrices at n random spike times. Random spike times are found by taking the original spike times and shifting all times by the same random amount (Aljadeff *et al.*, 2016). The null distribution of eigenvalues is computed over 1000 random STC matrices.

To determine the number of STC eigenvectors to keep in the analysis, we remove eigenvectors that are very similar to the STA (projection between unit STA and STC vectors >0.9) (Aljadeff *et al.*, 2016). We then transform each STC eigenvector to be orthogonal to the STA using Gram-Schmidt orthonormalization. The D stimulus dimensions in the model are the STA and the significant STC eigenvectors that differ from the STA: $\{\phi_{STA}, \phi_{STC,1}, \phi_{STC,2}, \dots, \phi_{STC,D-1}\}$, where the STC eigenvectors $\phi_{STC,1}, \phi_{STC,2}, \dots, \phi_{STC,D-1}$ are ordered according to the magnitude of the associated eigenvalues.

We used a nonparametric estimate of the nonlinearity that maps a low-dimensional stimulus representation to a spiking probability (Aljadeff *et al.*, 2016). The spiking nonlinearity was estimated for one- and two-dimensional stimulus spaces. The full two-dimensional nonlinearity

$$P(\text{spike}|s(t)) = g_2(\phi_{STA}^T s(t), \phi_{STC,1}^T s(t)) \quad (6)$$

was estimated using a histogram by first discretizing the stimulus projections into a finite number of bins and then computing the relative frequency with which spikes occurred for stimulus projections in each bin. The one-dimensional nonlinearity that depends only on the STA

$$P(\text{spike}|s(t)) = g_1(\phi_{STA}^T s(t)) \quad (7)$$

is the marginal spiking probability found by averaging the two-dimensional nonlinearity $g_2(\phi_{STA}^T s(t), \phi_{STC,1}^T s(t))$ over all values of the projections $\phi_{STC,1}^T s(t)$.

Even though many neurons had more than two significant filters, we limited the estimation of the nonlinearity to two-dimensional stimulus spaces for all neurons because estimation of high-dimensional nonlinearities requires a prohibitively large data set to estimate reliably using the histogram method (Schwartz *et al.*, 2006; Aljadeff *et al.*, 2016). The two-dimensional spiking nonlinearities were smoothed with a two-dimensional Gaussian with standard deviations equal to 7/8 of the histogram bin step size in each dimension for plotting.

F. Model evaluation

We assessed the accuracy of models with different numbers of filters using responses to repeated presentations of a 200 ms duration frozen Gaussian broadband noise stimulus that was not used in the fitting analysis. We used a frozen noise stimulus that shared the same statistics as the noise used to fit the model so that we can expect the model to

generalize to the test stimulus. We used frozen noise, rather than unfrozen noise, so that the peri-stimulus time histogram (PSTH) of the response to repeated presentations has a structure that reflects the neuron's selectivity for stimulus features. For each model, we computed the correlation coefficient between the predicted PSTH and the PSTH of the experimental data computed at the stimulus temporal resolution of 20.8 μ s, denoted CC_{model} . To compute the predicted PSTH, we used the model to produce 10 000 simulated spike trains. To determine how well we can expect the model to match the data, we computed the correlation between the measured PSTHs on 1000 randomly selected different subsets of trials (Hsu *et al.*, 2004; Heitman *et al.*, 2016; Schoppe *et al.*, 2016), denoted CC_{half} . CC_{half} was used to determine the maximum we can expect for the correlation between the model and the measured PSTH. This bound, denoted CC_{max} , is given by

$$CC_{max} = \sqrt{\frac{2CC_{half}}{1 + CC_{half}}}. \quad (8)$$

Note that we have rewritten the expression for CC_{max} given in Schoppe *et al.* (2016) to avoid dividing by small numbers when CC_{half} is near zero. The square of the normalized correlation coefficient

$$CC_{norm} = \frac{CC_{model}}{CC_{max}} \quad (9)$$

provides a measure of the percent of explainable variation in the response that is described by the model (Schoppe *et al.*, 2016).

G. Filter analysis

The BF of each filter was the weighted average of the frequencies where the amplitude spectrum of the filter exceeded 50% of the peak of the amplitude spectrum. The height of the amplitude spectrum, normalized so that the sum across frequency equals one, was used to weight the frequencies. We used this measure instead of the frequency at the peak because some filters were broadband and the peak can reflect random noise rather than a measure of the center of the frequency range where the filter has power.

The bandwidth of each filter was the frequency range at half height of the amplitude spectrum of the filter.

All filters were fit with gammachirp functions (Irino and Patterson, 1997). Gammachirp functions are commonly used to describe impulse responses of ANFs (Fontaine *et al.*, 2015). We determined whether all model filters were described by gammachirps in order to determine whether higher order dimensions of the response also reflect bandpass filters, or a different type of filter function. We fit the component of the filter where the envelope had $>10\%$ of its maximum value (Fischer *et al.*, 2011). The envelope of a signal $x(t)$ is defined as the magnitude of the analytic signal $x(t) + iy(t)$, where $y(t)$ is the Hilbert transform of $x(t)$. The gammachirp function is given by

$$f(t) = A(t-t_0)^3 \exp\left(-\frac{t-t_0}{\tau}\right) \times \cos\left(2\pi\left(f(t-t_0) + 0.5c(t-t_0)^2\right) + \phi\right) H(t-t_0), \quad (10)$$

where $H(t)$ is the unit step function. Parameters of the gamma-chirp filter were fit by minimizing the mean squared error.

H. Spiking nonlinearity analysis

We quantified the asymmetry of the spiking nonlinearity along the STA and STC 1 dimensions individually. The asymmetry index is defined as

$$\text{asymmetry index} = \frac{R - L}{R + L}, \quad (11)$$

where R and L are the sums of the spiking probabilities for stimulus projections that are greater than zero and less than zero, respectively (Atencio *et al.*, 2008). The asymmetry index ranges between -1 and 1 . It will be greater than zero when the spiking probability is higher for positive stimulus projections than for negative stimulus projections.

We used the marginal spiking nonlinearities along the STA and STC 1 dimensions to determine whether the two-dimensional spiking nonlinearity reflects a synergy of the inputs along the dimensions or is merely a product of the marginal spiking nonlinearities. The marginal spiking nonlinearities are the spiking nonlinearities along the STA and STC 1 dimensions that are found by averaging the two-dimensional spiking nonlinearity over the other variable. The definition of a marginal spiking nonlinearity is motivated by the definition of a marginal probability density in probability theory. To quantify the degree of dependence in the two-dimensional spiking nonlinearity, we used an inseparability index that is based on the singular value decomposition (SVD) of the two-dimensional spiking nonlinearity (Atencio *et al.*, 2008). The SVD represents a two-dimensional matrix as a weighted sum of rank one matrices $g_2 = \sum_{i=1}^n \sigma_i u_i v_i^T$, where the weights σ_i are called singular values. If the two-dimensional spiking nonlinearity is a product of the marginal nonlinearities, then only the first singular value will be nonzero. If the two-dimensional spiking nonlinearity reflects a dependence on pairs of inputs along the STA and STC 1 dimensions, then more than one singular value will be nonzero. The inseparability index quantifies this dependence by measuring the relative strength of the first singular value

$$\text{inseparability index} = 1 - \frac{\sigma_1^2}{\sum_{i=1}^n \sigma_i^2}. \quad (12)$$

The inseparability index ranges between 0 and 1 with a value of 0 indicating that the two-dimensional spiking nonlinearity is a product of the marginal nonlinearities.

We quantified the circular spread of two-dimensional spiking nonlinearities using a measure that is similar to the

vector strength commonly used to quantify phase locking strength (Goldberg and Brown, 1968; Köppl, 1997a). We first computed the strength of the nonlinearity in each direction by summing the nonlinearity over all points at that direction

$$g_\theta(\theta) = \sum_r g(r, \theta), \quad (13)$$

where $g(r, \theta)$ is the two-dimensional nonlinearity in polar coordinates. Since the spiking nonlinearity is the conditional probability of spiking given the stimulus, $g_\theta(\theta)$ corresponds to the marginal probability over direction. The nonlinearity vector strength (NVS) compares the length of the sum of unit vectors at each direction weighted by the spiking probability to the length of the vector that would result if the spiking probability was concentrated at one direction

$$\text{NVS} = \frac{\left| \sum_\theta u(\theta) g_\theta(\theta) \right|}{\sum_\theta g_\theta(\theta)}, \quad (14)$$

where $u(\theta)$ is a unit vector pointing in direction θ . This matches the calculation of the vector strength used with sample data (Goldberg and Brown, 1968). The NVS measure is one minus the circular variance; we use NVS rather than the variance because of its relationship to the sample vector strength that is commonly used to quantify the strength of phase locking.

III. RESULTS

The aim of this study was to characterize the input-output properties of ANF in the barn owl. We used a non-parametric modeling approach based on spike-triggered techniques, STA, and STC, to estimate the number of relevant stimulus dimensions and the nonlinear interactions along these dimensions that produce ANF responses.

A. Multiple stimulus dimensions

We first determined the number of stimulus dimensions that are important for producing ANF responses. We used STA and STC analysis of responses to Gaussian white noise stimuli at a given stimulus level to determine the number of significant stimulus dimensions. By construction in the analysis, the first stimulus dimension is the STA. As described previously, the relatively high level of phase locking in barn owl ANFs allows the STA to be well-defined for all neurons (Köppl, 1997a; Fontaine *et al.*, 2015). The remaining stimulus dimensions were the significant STC filters that were different from the STA. We found that all neurons had multiple significant stimulus dimensions [Fig. 2(A); median = 4, interquartile range = 1, range 2–7; $n = 123$]. There was a weak negative correlation between the number of dimensions and the BF of the STA [Fig. 2(B); $r = -0.3$, $p < 0.001$; $n = 123$]. Thus, the STC analysis shows that each ANF has multiple significant stimulus dimensions, whose properties we describe below.

FIG. 2. Number of stimulus dimensions. (A) The total number of filters (dark gray) and the number of excitatory (light gray) and suppressive (black) filters. The total number of filters is the number of stimulus dimensions for the neuron. (B) The number of filters vs the BF of the STA.

B. Frequency selectivity of model filters

Many parametric models of mammalian ANF responses that contain multiple filters assume a fixed relationship between the frequency selectivity of the model filters (Goldstein, 1990; Meddis *et al.*, 2001; Zhang *et al.*, 2001; Sumner *et al.*, 2002; Bruce *et al.*, 2003; Tan and Carney, 2003; Irino and Patterson, 2006; Zilany and Bruce, 2006). We analyzed the frequency selectivity of the STA and STC filters to determine whether there is a consistent relationship between the first and the secondary filters for the barn owl ANFs. We computed the BF for each filter as the center of mass of the frequency range where the amplitude spectrum of the filter was at least 50% of the maximum value. The BFs of the first filters (STA) in the sample covered most of the audible spectrum for the barn owl (Fig. 3). The BFs of the first and second (STC 1) stimulus dimensions were highly correlated [$r^2=0.94$, $\text{BF}_{\text{STC},1}=0.33 + 0.95\text{BF}_{\text{STA}}$; $n=123$; Fig. 3(A)]. The BFs of the first and third (STC 2) stimulus dimensions showed lower correlation [$r^2=0.57$, $\text{BF}_{\text{STC},2}=1.83 + 0.63\text{BF}_{\text{STA}}$; $n=117$; Fig. 3(B)]. The difference between the BFs of the first and third stimulus dimensions could be up to 4.2 kHz. Many of the third filters were broadband filters, resulting in BFs concentrated near the center of the frequency range of the stimulus. The correlation between the BF of the first and the fourth and fifth dimensions was smaller still (STA and STC 3: $r^2=0.5$, $n=102$; STA and STC 4: $r^2=0.47$, $n=53$; filters 1 and 6: $r^2=0.14$, $n=13$). This shows that there is similar frequency selectivity of the first and second stimulus dimensions; however, there is not a precise relationship between the

FIG. 3. Filters' best frequencies. (A) Correlation between the BFs of the first filter (STA) with the second filter (STC1). (B) Correlation between the BFs of the first filter (STA) and the third filter (STC2). The solid line is the identity line.

frequency selectivity of the first stimulus dimension and dimensions beyond the second.

The STAs are bandpass filters that are well-described by gammachirp filters, as described in detail previously (Fontaine *et al.*, 2015). The significant STC filters had diverse properties [Figs. 4(A) and 4(B)]. The first STC filters were mostly bandpass filters that are well-described by gammachirp filters, but the accuracy of the gammachirp fit decreased for higher modes [Fig. 4(C)]. The reduced accuracy of the gammachirp fit for the STC filters resulted from multiple factors. Many filters had a central portion that was well-described by a gammachirp filter, but with additional noise around it [e.g., Fig. 4(A) STC 1 and STC 2]. This noise is evident in the increased baseline of the amplitude spectrum of the STC modes relative to the STA [Fig. 4(A), STC 1 and STC 2]. Other filters were broadband, with power over most of the spectrum of the sound stimulus and were therefore poorly fit by the gammachirp [Fig. 4(A), STC 3]. Corresponding to the accurate gammachirp fits, the majority of the first mode STC filters had narrow bandwidths, as found for the STAs [Fig. 4(D)]. In contrast, the majority of higher order STC modes had larger bandwidths, consistent with the reduced accuracy of the fit by the gammachirp filter [Fig. 4(D)]. Since the bandwidth was measured at half the maximum of the amplitude spectrum, the increased bandwidth of higher order STC modes did not reflect the increased baseline of the amplitude spectrum. In summary, for most neurons, the first two stimulus dimensions were given by gammachirp filters at nearly the same frequency, while higher order dimensions were more often broadband.

C. Excitatory and suppressive dimensions

We next examined whether the significant stimulus dimensions had an excitatory or suppressive effect on the responses. Stimulus dimensions were judged excitatory or suppressive when stimulus projections along those dimensions increased or decreased spiking probability, respectively. A stimulus dimension from the STC analysis was considered excitatory or suppressive when its associated eigenvalue was positive or negative, respectively (Schwartz *et al.*, 2006; Aljadeff *et al.*, 2016). Note that these designations alone do not inform whether a neuron will have a high or low maximum firing rate. The first stimulus dimension, corresponding to the STA, was excitatory for all neurons. The second dimension could be either excitatory or suppressive, and the relative numbers of each type varied with BF. The median BF for neurons with an excitatory second dimension was significantly higher than the median BF for neurons with a suppressive second dimension [$p < 10^{-4}$, Mann-Whitney U test; $n_{\text{excitatory}}=97$, $n_{\text{suppressive}}=26$; Fig. 5(A)]. Similarly, the median BF for neurons with an excitatory third dimension was significantly higher than the median BF for neurons with a suppressive third dimension [$p < 10^{-4}$, Mann-Whitney U test; $n_{\text{excitatory}}=64$, $n_{\text{suppressive}}=54$; Fig. 5(B)]. This also demonstrates variation in model properties over the population and suggests that a fixed model structure will not capture the responses of all neurons.

FIG. 4. (Color online) Filter shapes. (A) Examples of filters with gammachirp fits and the amplitude spectrum of the filter for one neuron. All amplitude spectra were normalized by the maximum of the amplitude spectrum of the STA. (B) Eigenvalues used to determine the significant filters shown in (A). Significant eigenvalues are shown as red dots. (C) Root-mean-square (rms) error in gammachirp fits to model filters. The fits are poor for some filters that are broadband, rather than bandpass [as shown in (D)]. (D) Filter bandwidths (width at half height of the amplitude spectrum) are low, corresponding to bandpass filters, for most of the first and second filters. Higher order filters are often broadband as seen in the example neuron.

D. Spiking nonlinearity

The spiking nonlinearity specifies how the responses along the stimulus dimensions interact to determine the probability of spiking. We estimated the spiking nonlinearity using a non-parametric approach by calculating the spiking probability for a grid of response values along the first two dimensions. We limited the analysis to two-dimensional nonlinearities because the estimation of three-dimensional nonlinearities required prohibitively large data sets. There was diversity in the shape of the nonlinearity in the first two stimulus dimensions, ranging from highly localized [Fig. 6(A), left] to nearly circularly-symmetric [Fig. 6(A), right].

We first analyzed the marginal spiking probabilities along the STA and STC 1 mode separately to determine how responses along these dimensions influence spiking. To gain insight into the function of the nonlinearity, we quantified the asymmetry of the one-dimensional nonlinearities using an asymmetry index that varies from -1 to 1 (Atencio *et al.*, 2008). A nonlinearity such as half-wave rectification will be asymmetric, while a squaring nonlinearity that measures energy will be symmetric. The asymmetry index is positive when the spiking probability is higher for positive projections of the stimulus along the dimension. The asymmetry

index is negative when the spiking probability is higher for negative projections. The asymmetry index will be near zero when the nonlinearity is symmetric. The STA nonlinearity was typically asymmetric, where spiking probability increased as the projection of the stimulus along the STA dimension increased in the positive direction (Fig. 6). This type of nonlinearity corresponds to half-wave rectification of the projection of the stimulus along the STA dimension. The asymmetry index for the STA nonlinearity was positive for most neurons, consistent with a half-wave rectification [Fig. 6(B)]. The STC 1 nonlinearity was often symmetric, with an asymmetry index that was distributed around zero [Fig. 6(B)].

We next determined whether the two-dimensional nonlinearity reflects a dependence between responses along the STA and the STC 1 mode, and thus that the neuron is selective for particular combinations of these modes (Atencio *et al.*, 2008). To test this, we examined how well the two-dimensional nonlinearity could be described by the product of the marginal nonlinearities (Atencio *et al.*, 2008). The accuracy of this approximation is quantified by an inseparability index that ranges from zero to one for independent to cooperative inputs, respectively (Atencio *et al.*, 2008). The distribution of inseparability indices was skewed to the right, with values ranging from 0.01 to 0.5 [Fig. 6(C)]. Although the inseparability indices were small for some neurons, the presence of nonzero inseparability indices shows that the two-dimensional nonlinearity reflects some degree of cooperative coding between the STA and STC 1 dimensions in most neurons.

For those neurons where the first two filters are matched in frequency selectivity, a restricted nonlinearity reflects strong phase locking to tones, while a circularly-symmetric nonlinearity reflects a phase invariant response [Fig. 7(A)]. When the first two filters are not matched in frequency, selectivity localized nonlinearities reflect response suppression by frequencies outside the STA. Conversely, broad

FIG. 5. Excitatory and suppressive filters. (A) BF for neurons where the first STC mode (STC 1) is excitatory or suppressive. (B) BF for neurons where the second STC mode (STC 2) is excitatory or suppressive.

FIG. 6. (Color online) Spiking nonlinearity. (A) Example two-dimensional spiking nonlinearities ranging from restricted in the plane (left) to nearly circularly symmetric (right) for dimensions 1 and 2. The curves on the top and side show the marginal STA and STC 1 nonlinearities, respectively. (B) Asymmetry indices for the STA nonlinearity and STC 1 nonlinearity. The dashed line is the identity line. (C) Nonlinearity inseparability indices vs. STA BF. The values for the nonlinearities in (A) are shown by the red triangle (left), green square (center), and blue diamond (right).

nonlinearities reflect facilitation by frequencies outside the STA. We quantified the spread of the nonlinearity around the circle using a vector strength measure that is analogous to the vector strength that is commonly used to quantify the strength of phase locking (Goldberg and Brown, 1968). The NVS is equal to one when the spiking probability is concentrated at a single angle, and is equal to zero when the spiking probability is circularly symmetric. The NVS varied with BF in a manner that is similar to the experimentally measured dependence of spiking vector strength on BF [Fig. 7(B)] (Köpl, 1997a). Most neurons with low BFs had a high NVS, indicating that the spiking nonlinearity was restricted in the plane spanning the first two filters. This is the type of nonlinearity expected for a neuron with an excitatory first stimulus dimension and suppressive second stimulus dimension (Rust *et al.*, 2005; Fairhall *et al.*, 2006). For neurons with BFs above 3 kHz, there was weak negative correlation between BF and NVS for the nonlinearity in dimensions 1 and 2 ($r = -0.35$, $p = 0.0002$). The large spread of the NVS at high frequencies suggests that the shape of the nonlinearity reflects not only the strength of phase locking strength but also additional nonlinear interactions across stimulus dimensions that influence the neuron's responses.

FIG. 7. (Color online) Spiking nonlinearity shape and phase locking. (A) An example neuron where the first two filters (the STA and STC 1) have similar BFs and therefore approximate a quadrature pair of filters. In this case, the nonlinearity reflects the strength of phase locking for the neuron. (B) We used a vector strength measure to quantify the shape of the two-dimensional spiking nonlinearity for dimensions 1 and 2. The nonlinearity vector strength is weakly negatively correlated with BF. Black points are for neurons where the difference between the BFs of the filters is less than 200 Hz.

E. Prediction accuracy

We tested the one- and two-dimensional models by predicting the response to a repeated frozen Gaussian white noise stimulus that was not used to fit the models. The models showed a range of prediction accuracy over the population when assessed at the sampling rate of the stimulus [bin size $20.8 \mu\text{s}$; Figs. 8(A) and 8(B)]. Neurons with low BFs tended to have clearly defined portions of the stimulus that elicited spikes, resulting in stripes in the spike raster across trials [Fig. 8(A)]. For these neurons, the predicted spiking probability matched the general shape of the measured PSTH, producing r^2 values greater than 0.4 [Fig. 8(C)]. Neurons with high BFs did not exhibit clearly visible structure in the raster plot at a fine time scale. Here, the models poorly predicted the measured PSTH, producing r^2 values less than 0.1 [Fig. 8(B)]. Over the population, the prediction accuracy was negatively correlated with BF [Fig. 8(C)]. The prediction accuracy of models with one or two stimulus dimensions was not significantly different over the population when measuring responses at the sampling rate of the stimulus ($p > 0.28$, Kruskal Wallis). While there was no difference at the population level, the prediction accuracy for the test stimuli was greater for the two-dimensional model than for the one-dimensional model for 70 of 123 individual neurons.

We next sought to explain the prediction accuracy of the models in terms of response properties of ANFs. The performance of any model is limited by the variability of the neural responses (Sahani and Linden, 2003; Hsu *et al.*, 2004; Schoppe *et al.*, 2016). To assess the variability of neural responses, we computed the correlation between the PSTHs of the data on different subsets of trials (Hsu *et al.*, 2004; Heitman *et al.*, 2016; Schoppe *et al.*, 2016). We used this correlation between the PSTHs of the data on different subsets of trials to approximate the maximum value we can expect for the correlation between the model and the data, called CC_{max} (Schoppe *et al.*, 2016). The value of CC_{max} decreased as BF increased [Fig. 8(D)]. We then computed the normalized correlation coefficient as a measure of the percent of explainable variation in the measured PSTH that

FIG. 8. (Color online) Model prediction accuracy. (A) An example neuron with high prediction accuracy. The plot on top shows the experimentally measured raster plot of the neuron's response to repeated presentations of a broadband noise stimulus. The bottom plot shows the measured PSTH (black) along with the model predictions of the spiking probability for models with one (red) and two (blue) filters at high temporal resolution, matching the sampling rate of the stimulus. (B) An example neuron with low prediction accuracy. Plots are as in (A). (C) R^2 between the model prediction and the measured PSTH models with one and two filters vs the BF of the STA. (D) CC_{max}^2 , computed from the correlation between measured PSTH different subsets of trials vs BF. CC_{max} is the maximum possible correlation between the model and the measured PSTH. (E) CC_{norm}^2 is the percent of explainable variation that is described by the model. (F) Difference in CC_{norm}^2 for one-dimensional (1-D) and two-dimensional (2-D) models (2-D-1-D).

is described by the model (Schoppe *et al.*, 2016). The normalized correlation coefficient shows that the model is performing nearly as well as can be expected for a subset of neurons with BFs below 5 kHz, but is missing aspects of the response for high BF neurons when predicting responses at the sampling rate of the stimulus [Fig. 8(E)]. The normalized correlation coefficient was greater for the two-dimensional model than for the one-dimensional model for 70 of 123 individual neurons [Fig. 8(F)].

We next determined how the prediction accuracy changed when the PSTHs were measured at lower temporal resolution [Figs. 9(A) and 9(B)]. We computed the normalized correlation coefficient for PSTH bin sizes ranging from 20.8 μ s to 4.2 ms, and found the bin size that maximized the normalized correlation coefficient. The majority of neurons (greater than 60%) maximized the normalized correlation coefficient for PSTH bin sizes that were less than 1 ms [Fig. 9(C)]. At this neuron-specific temporal resolution, the model prediction accuracy was higher for all neurons, but particularly so for the high-BF neurons [Figs. 9(D) and 9(E)]. While the models had a higher normalized correlation coefficient at the lower temporal resolution, they were still unable to describe the magnitude of response fluctuations with the stimulus envelope [Fig. 9(B)].

This suggests that there are limitations on prediction accuracy placed by the simple model structure.

IV. DISCUSSION

We applied STA and STC analyses to barn owl ANF responses to Gaussian white noise. We found that multiple significant stimulus dimensions interact in a variety of ways to determine ANF responses. The results here show that a complete auditory nerve encoding model for the barn owl should include multiple filters whose outputs can interact in diverse nonlinear ways.

Several aspects of the analysis presented here show that ANF responses in the barn owl depend on multiple stimulus dimensions. First, the STC analysis detected multiple relevant stimulus dimensions for each neuron in the sample. Second, the two-dimensional model outperformed the one-dimensional model when predicting responses to frozen stimuli in a cross validation test for the majority of the neurons. As seen in Fig. 8(A), both the one- and two-dimensional models followed the basic periodicity of the response, but the two-dimensional model was better able to capture rapid fluctuations in the amplitude of the PSTH.

FIG. 9. (Color online) Model prediction accuracy at lower temporal resolution. (A) An example neuron with high prediction accuracy. The raster plot on top shows the experimentally measured raster plot of the neuron's response to repeated presentations of a broadband noise stimulus. The bottom plot shows the PSTH and model predictions with a PSTH bin size of 0.5 ms, which maximized the CC_{norm}^2 for this neuron. (B) An example neuron with low prediction accuracy. Plots are as in (A), but with a PSTH bin size of 1.05 ms, which maximized the CC_{norm}^2 for this neuron. (C) PSTH bin size that maximizes CC_{norm}^2 . (D) CC_{norm}^2 at the PSTH bin size that maximizes CC_{norm}^2 . (E) Difference between the maximal CC_{norm}^2 and the value at the sampling rate of the stimulus.

However, there were neurons where the two-dimensional model performed worse than the one-dimensional model in the cross validation test. This may reflect a low-dimensional response space in this subset of cells, or a failure of the multidimensional model to generalize to the test stimulus (Schwartz *et al.*, 2006; Aljadeff *et al.*, 2016).

A. Added value of the present model

Spike-triggered analyses have a long history in auditory neuroscience, with the STA being a common measure of the primary linear filter describing neural responses (de Boer and Kuiper, 1968; Eggermont *et al.*, 1983c). Auditory nerve responses have demonstrated nonlinear interactions between excitation and suppression across frequencies whose description requires more than a single linear temporal filter (Robles and Ruggero, 2001; Heil and Peterson, 2015). The spectrotemporal receptive field (STRF) provides a description of how excitation and inhibition in different frequency bands produce ANF responses at low temporal resolution (Eggermont *et al.*, 1983c; Lewis and van Dijk, 2004). However, the STRF alone does not describe nonlinear interactions across frequency and does not describe responses on a time scale that allows for phase locking. Wiener-Volterra expansion models use time-dependent kernels to make a polynomial approximation to the neural response, which are typically limited to quadratic approximations due to data limitations (Eggermont *et al.*, 1983a,b; Eggermont *et al.*, 1983c; Eggermont, 1993; van Dijk *et al.*, 1994; Yamada and Lewis, 1999). Wiener-Volterra expansion models share some similar properties to the model we present here. The first Wiener kernel is the STA and the second order Wiener kernel is related to the STC filters (Schwartz *et al.*, 2006; Sandler and Marmarelis, 2015). However, the restriction of the nonlinearity to a quadratic form in the Wiener-Volterra expansion model will fail to describe the many forms of nonlinearity we observed in the data. The use of the STA and STC analyses, together with the histogram estimation of the nonlinearity, overcomes this limitation of the Wiener-Volterra expansion models.

B. Informing future modeling efforts

The non-parametric analysis performed here is complementary to those used in parametric modeling by informing the structure of the model, which can then be fit to data. The performance of the models we tested can be improved by incorporating elements of parametric mammalian auditory nerve models. For example, a barn owl model should minimally include a hair cell-auditory nerve synapse model (Meddis, 1986; Westerman and Smith, 1988) and a spike history term to produce adaptation and a refractory period (e.g., Zhang *et al.*, 2001). We expect that a parametric approach will be required to produce a model that predicts spiking on a timescale of tens of microseconds, as required for the barn owl. Our analysis shows that a population model of the barn owl auditory nerve should include multiple filters where the frequency selectivity of each filter can be fit to data because there was no fixed relationship between the frequency selectivity of the STA and STC filters. Additionally, the form of the spiking nonlinearity should change with the

best frequency of the neuron to allow for frequency-dependence in whether secondary stimulus dimensions are excitatory or suppressive. A more general model will also allow the stimulus level to change dynamically (Goldstein, 1990; Meddis *et al.*, 2001; Zhang *et al.*, 2001). These elements could be included in a purely statistical model or a model that attempts to model aspects of cochlear mechanics.

C. Possible physiological correlates

The STC analysis used here makes no specific predictions about where the filters are physically implemented in cochlear physiology. STC analysis uncovers a low-dimensional subspace of relevant stimulus dimensions, but does not uniquely identify particular filters (Schwartz *et al.*, 2006; Aljadeff *et al.*, 2016). However, by focusing on the STA, and the STC modes that are additional filters to the STA, we frame the model in terms of a primary filter whose frequency selectivity has been shown to match the frequency selectivity of the neuron (Fontaine *et al.*, 2015). This probably reflects a combination of mechanisms that contribute to cochlear tuning. In birds, the most likely candidates are hair-bundle micromechanics and membrane channel kinetics (Gleich and Manley, 2000), less likely longitudinal traveling waves (Gummer *et al.*, 1987; Köppl, 2011; Xia *et al.*, 2016). The presence of suppressive dimensions that are frequency selective with BFs shifted from the BF of the STA may reflect mechanisms that produce two-tone rate suppression (Schwartz and Simoncelli, 2001) and are assumed to arise micromechanically in the cochlea (Chen *et al.*, 1996; Robles and Ruggero, 2001). Broadband filters (Fig. 4) perform a gain control operation that may reflect multiple mechanisms, such as middle-ear transmission (Manley, 2017). While elements of the models are consistent with a mechanistic interpretation, we take caution in identifying particular STC modes as individual filters that have a physiological basis.

D. Multidimensional coding in other species

Wiener-Volterra models suggest that multiple stimulus dimensions are relevant for mammalian and amphibian ANF responses (Lewis *et al.*, 2002a; Recio-Spinoso *et al.*, 2005). The second Wiener kernel is closely related to the STC matrix (Sandler and Marmarelis, 2015), and therefore a neuron with multiple significant singular vectors of the second Wiener kernel will likely have multiple significant STC modes. Previous studies in the chinchilla and bullfrog have shown that one or two singular vectors account for most of the structure in the second Wiener kernel (Lewis *et al.*, 2002a; Recio-Spinoso *et al.*, 2005). In contrast, in the gerbil, there are multiple singular vectors of the second Wiener kernel with similar weights (Lewis *et al.*, 2002b). Although the dimension of the sound space encoded by ANFs has not been quantified in other species using STC analysis, our results are consistent with evidence of multiple relevant stimulus dimensions in auditory nerve coding from other species.

E. Conclusion

This study provides an analysis of the stimulus space encoded in barn owl auditory nerve responses. The barn owl

is a model system for studies of sound localization, and many studies have indicated that spectrotemporal processing by neurons at various stages of the sound localization pathway is important for computing sound location and identity (Keller and Takahashi, 2005; Fischer *et al.*, 2011; Steinberg *et al.*, 2013). Our results show that complex and varied spectrotemporal processing occurs in the auditory nerve. This establishes a basis for extracting various useful features of sound at early stages of the auditory system.

ACKNOWLEDGMENTS

This work was supported by Grant No. DC007690 to J.L.P. and the Deutsche Forschungsgemeinschaft (Collaborative Research Center ‘Active Hearing,’ project A14) to C.K. We thank E. Hunt for help with data management, M. Beckert for help with figures and comments on the manuscript, and R. Ferger for comments on the manuscript. We enthusiastically thank the Neuroscience Gateway for computing resources.

- Aljadeff, J., Lansdell, B. J., Fairhall, A. L., and Kleinfeld, D. (2016). “Analysis of neuronal spike trains,” *Deconstructed. Neuron* **91**, 221–259.
- Atencio, C. A., and Sharpee, T. O. (2008). “Schreiner CE cooperative nonlinearities in auditory cortical neurons,” *Neuron* **58**, 956–966.
- Bialek, W., and van Steveninck, R. R. (2005). “Features and dimensions: Motion estimation in fly vision,” ArXivq-Bio0505003.
- Bruce, I. C., Sachs, M. B., and Young, E. D. (2003). “An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses,” *J. Acoust. Soc. Am.* **113**, 369–388.
- Chen, L., Salvi, R. J., Trautwein, P. G., and Powers, N. (1996). “Two-tone rate suppression boundaries of cochlear ganglion neurons in normal chickens,” *J. Acoust. Soc. Am.* **100**, 442–450.
- de Boer, E., and de Jongh, H. R. (1978). “On cochlear encoding: Potentialities and limitations of the reverse-correlation technique,” *J. Acoust. Soc. Am.* **63**, 115–135.
- de Boer, E., and Kuyper, P. (1968). “Triggered correlation,” *IEEE Trans. Biomed. Eng.* **15**, 169–179.
- Eggermont, J. J. (1993). “Wiener and Volterra analyses applied to the auditory system,” *Hear. Res.* **66**, 177–201.
- Eggermont, J. J., Aertsen, A. M., and Johannesma, P. I. (1983a). “Prediction of the responses of auditory neurons in the midbrain of the grass frog based on the spectro-temporal receptive field,” *Hear. Res.* **10**, 191–202.
- Eggermont, J. J., Aertsen, A. M., and Johannesma, P. I. (1983b). “Quantitative characterisation procedure for auditory neurons based on the spectro-temporal receptive field,” *Hear. Res.* **10**, 167–190.
- Eggermont, J. J., Johannesma, P. M., and Aertsen, A. M. (1983c). “Reverse-correlation methods in auditory research,” *Q. Rev. Biophys.* **16**, 341–414.
- Fairhall, A. L., Burlingame, C. A., Narasimhan, R., Harris, R. A., Puchalla, J. L., and Berry, M. J. (2006). “Selectivity for multiple stimulus features in retinal ganglion cells,” *J. Neurophysiol.* **96**, 2724–2738.
- Fischer, B. J., Steinberg, L. J., Fontaine, B., Brette, R., and Peña, J. L. (2011). “Effect of instantaneous frequency glides on interaural time difference processing by auditory coincidence detectors,” *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18138–18143.
- Fontaine, B., Köppl, C., and Peña, J. L. (2015). “Reverse correlation analysis of auditory-nerve fiber responses to broadband noise in a bird, the barn owl,” *J. Assoc. Res. Otolaryngol.* **16**, 101–119.
- Gleich, O., and Manley, G. A. (2000). “The hearing organ of birds and crocodilian,” in *Comparative Hearing: Birds and Reptiles. Springer Handbook of Auditory Research* (Springer, New York), pp. 70–138.
- Goldberg, J. M., and Brown, P. B. (1968). “Functional organization of the dog superior olivary complex: An anatomical and electrophysiological study,” *J. Neurophysiol.* **31**, 639–656.
- Goldstein, J. L. (1990). “Modeling rapid waveform compression on the basilar membrane as multiple-bandpass-nonlinearity filtering,” *Hear. Res.* **49**, 39–60.
- Gummer, A. W., Smolders, J. W., and Klinke, R. (1987). “Basilar membrane motion in the pigeon measured with the Mössbauer technique,” *Hear. Res.* **29**, 63–92.
- Heil, P., and Peterson, A. J. (2015). “Basic response properties of auditory nerve fibers: A review,” *Cell Tissue Res.* **361**, 129–158.
- Heitman, A., Brackbill, N., Greschner, M., Sher, A., Litke, A. M., and Chichilnisky, E. J. (2016). “Testing pseudo-linear models of responses to natural scenes in primate retina,” bioRxiv:045336.
- Hsu, A., Borst, A., and Theunissen, F. E. (2004). “Quantifying variability in neural responses and its application for the validation of model predictions,” *Netw. Bristol Engl.* **15**, 91–109.
- Irino, T., and Patterson, R. D. (1997). “A time-domain, level-dependent auditory filter: The gammachirp,” *J. Acoust. Soc. Am.* **101**, 412–419.
- Irino, T., and Patterson, R. D. (2006). “A dynamic compressive gammachirp auditory filterbank,” *IEEE Trans. Audio Speech Lang. Process.* **14**, 2222–2232.
- Keller, C. H., and Takahashi, T. T. (2005). “Localization and identification of concurrent sounds in the owl’s auditory space map,” *J. Neurosci.* **25**, 10446–10461.
- Köppl, C. (1997a). “Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, *Tyto alba*,” *J. Neurosci.* **17**, 3312–3321.
- Köppl, C. (1997b). “Frequency tuning and spontaneous activity in the auditory nerve and cochlear nucleus magnocellularis of the barn owl *Tyto alba*,” *J. Neurophysiol.* **77**, 364–377.
- Köppl, C. (2011). “Birds—Same thing, but different? Convergent evolution in the avian and mammalian auditory systems provides informative comparative models,” *Hear. Res.* **273**, 65–71.
- Köppl, C., and Yates, G. (1999). “Coding of sound pressure level in the barn owl’s auditory nerve,” *J. Neurosci.* **19**, 9674–9686.
- Lewis, E. R., Henry, K. R., and Yamada, W. M. (2002a). “Tuning and timing of excitation and inhibition in primary auditory nerve fibers,” *Hear. Res.* **171**, 13–31.
- Lewis, E. R., Henry, K. R., and Yamada, W. M. (2002b). “Tuning and timing in the gerbil ear: Wiener-kernel analysis,” *Hear. Res.* **174**, 206–221.
- Lewis, E. R., and van Dijk, P. (2004). “New variation on the derivation of spectro-temporal receptive fields for primary auditory afferent axons,” *Hear. Res.* **189**, 120–136.
- Manley, G. A. (2017). “Comparative auditory neuroscience: Understanding the evolution and function of ears,” *J. Assoc. Res. Otolaryngol.* **18**, 1–24.
- Marmarelis, P. Z., and Marmarelis, V. Z. (1978). *Analysis of Physiological Systems: The White-Noise Approach* (Plenum Press, New York).
- Meddis, R. (1986). “Simulation of mechanical to neural transduction in the auditory receptor,” *J. Acoust. Soc. Am.* **79**, 702–711.
- Meddis, R., O’Mard, L. P., and Lopez-Poveda, E. A. (2001). “A computational algorithm for computing nonlinear auditory frequency selectivity,” *J. Acoust. Soc. Am.* **109**, 2852–2861.
- Recio-Spinoso, A., Temchin, A. N., van Dijk, P., Fan, Y.-H., and Ruggero, M. A. (2005). “Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers,” *J. Neurophysiol.* **93**, 3615–3634.
- Robles, L., and Ruggero, M. A. (2001). “Mechanics of the mammalian cochlea,” *Physiol. Rev.* **81**, 1305–1352.
- Rust, N. C., Schwartz, O., and Movshon, J. A. (2005). “Simoncelli EP spatiotemporal elements of macaque v1 receptive fields,” *Neuron* **46**, 945–956.
- Sahani, M., and Linden, J. F. (2003). “How linear are auditory cortical responses?,” in *Proceedings of Advances in Neural Information Processing Systems*, December 8–13, Vancouver, Canada.
- Sandler, R. A., and Marmarelis, V. Z. (2015). “Understanding spike-triggered covariance using Wiener theory for receptive field identification,” *J. Vis.* **15**, 16.
- Schoppe, O., Harper, N. S., Willmore, B. D. B., King, A. J., and Schnupp, J. W. H. (2016). “Measuring the performance of neural models,” *Front. Comput. Neurosci.* **10**.
- Schwartz, O., Pillow, J. W., Rust, N. C., and Simoncelli, E. P. (2006). “Spike-triggered neural characterization,” *J. Vis.* **6**, 484–507.
- Schwartz, O., and Simoncelli, E. P. (2001). “Natural signal statistics and sensory gain control,” *Nat. Neurosci.* **4**, 819–825.

- Sivagnanam, S., Majumdar, A., Yoshimoto, K., Astakhov, V. B. A., Martone, M., and Carnevale, N. T. (2013). "Introducing the neuroscience gateway," in *Proceedings of the CEUR Workshop*, May 22, Dresden, Germany.
- Steinberg, L. J., Fischer, B. J., and Peña, J. L. (2013). "Binaural gain modulation of spectrotemporal tuning in the interaural level difference-coding pathway," *J. Neurosci.* **33**, 11089–11099.
- Sullivan, W. E., and Konishi, M. (1984). "Segregation of stimulus phase and intensity coding in the cochlear nucleus of the barn owl," *J. Neurosci.* **4**, 1787–1799.
- Sumner, C. J., Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R. (2002). "A revised model of the inner-hair cell and auditory-nerve complex," *J. Acoust. Soc. Am.* **111**, 2178–2188.
- Tan, Q., and Carney, L. H. (2003). "A phenomenological model for the responses of auditory-nerve fibers. II. Nonlinear tuning with a frequency glide," *J. Acoust. Soc. Am.* **114**, 2007–2020.
- van Dijk, P., Wit, H. P., Segenhout, J. M., and Tubis, A. (1994). "Wiener kernel analysis of inner ear function in the American bullfrog," *J. Acoust. Soc. Am.* **95**, 904–919.
- van Steveninck, R. R., and Bialek, W. (1988). "Real-time performance of a movement-sensitive neuron in the blowfly visual system: Coding and information transfer in short spike sequences," *Proc. R. Soc. London B* **234**, 379–414.
- Westerman, L. A., and Smith, R. L. (1988). "A diffusion model of the transient response of the cochlear inner hair cell synapse," *J. Acoust. Soc. Am.* **83**, 2266–2276.
- Xia, A., Liu, X., Raphael, P. D., Applegate, B. E., and Oghalai, J. S. (2016). "Hair cell force generation does not amplify or tune vibrations within the chicken basilar papilla," *Nat. Commun.* **7**, 13133.
- Yamada, W. M., and Lewis, E. R. (1999). "Predicting the temporal responses of non-phase-locking bullfrog auditory units to complex acoustic waveforms," *Hear. Res.* **130**, 155–170.
- Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (2001). "A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression," *J. Acoust. Soc. Am.* **109**, 648–670.
- Zilany, M. S. A. (2006). "Bruce IC Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," *J. Acoust. Soc. Am.* **120**, 1446–1466.