Available online at www.sciencedirect.com
SCIENCE DIRECT®
Hearing Research 186 (2003) 30–46
HEARING RESEARCH
www.elsevier.com/locate/heares

New variations on the derivation of spectro-temporal receptive fields for primary auditory afferent axons

Edwin R. Lewis [a,*], Pim van Dijk [b]

[a] *Department of EECS, University of California, Berkeley, CA 94720-1770, USA*
[b] *University Hospital Maastricht, Department of Otorhinolaryngology and Head and Neck Surgery, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands*

Received 5 June 2003; accepted 22 July 2003

Abstract

The spectro-temporal receptive field [Hear. Res 5 (1981) 147; IEEE Trans BME 15 (1993) 177] provides an explicit image of the spectral and temporal aspects of the responsiveness of a primary auditory afferent axon. It exhibits the net effects of the competition between excitatory and inhibitory (or suppressive) phenomena. In this paper, we introduce a method for derivation of the spectro-temporal receptive field directly from a second-order Wiener kernel (produced by second-order reverse correlation between spike responses and broad-band white-noise stimulus); and we expand the concept of the spectro-temporal receptive field by applying the new method not only to the second-order kernel itself, but also to its excitatory and inhibitory subkernels. This produces separate spectro-temporal images of the excitatory and inhibitory phenomena putatively underlying the competition. Applied, in simulations, to models with known underlying excitatory and suppressive tuning and timing properties, the method successfully extracted a faithful image of those properties for excitation and one for inhibition. Applied to three auditory axons from the frog, it produced images consistent with previously published physiology.

Key words: Wiener kernel; Peripheral tuning; Suppression; Frog; Amphibian papilla; Basilar papilla

1. Introduction

The second-order Wiener kernel was introduced to the hearing research community by Eggermont et al. (1983a,b,c) and Van Dijk et al. (1994, 1997a,b). It is a two-dimensional extension of the *first-order* reverse-correlation (REVCOR) function introduced to the community by de Boer and his colleagues (de Boer and Kuyper, 1968; de Boer and de Jongh, 1978) and used extensively by others (e.g., Møller, 1986; Carney and Yin, 1988; Evans, 1989). The existence of the second-order Wiener kernel for an individual auditory afferent axon depends on the presence of an even-order nonlinearity somewhere in the signal path between the auditory input and the axon's spike trigger. It is computed by *second-order* reverse correlation between a continuous, broad-band noise stimulus and the occurrence of spikes in the axon. It is a time-domain representation of the particular features in the noise stimulus that, on average, led to spikes in that particular axon. An alternative representation of those features, in frequency and time, is given by the spectro-temporal receptive field (Hermes et al., 1981), which is equivalent to a single Fourier transform of the second-order Wiener kernel (Eggermont, 1993; see also Victor and Shapley, 1980). Within the spectro-temporal receptive field, one can see regions of time and frequency in which the noise power was above average prior to spikes, and regions where it was below average. The former would be associated with excitation or activation, the latter with inhibition or suppression.

* Corresponding author.
E-mail address: lewis@eecs.berkeley.edu (E.R. Lewis).

Abbreviations: AP, amphibian papilla; BP, basilar papilla; BEF, best excitatory frequency; DFT, discrete Fourier transform; SV*n*, *n*th ranking singular vector

0378-5955/03/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0378-5955(03)00257-0

The second-order Wiener kernel is computed as a symmetric matrix with real values. It can be decomposed into a set of singular vectors (which for real symmetric matrices are the same as the eigenvectors), and from these one can construct an excitatory subkernel and an inhibitory or suppressive subkernel (Yamada, 1997; Lewis et al., 2002). These provide separate time-domain representations of the particular features in the noise stimulus that, on average, led to excitation or activation, and those that, on average, led to suppression or inhibition. In some cases, the tuning properties of excitation and suppression are embodied clearly in the highest-ranking singular vectors of the two submatrices (Yamada, 1997). In other cases, owing to the noisiness of the data, tuning is not represented well by any singular vector. The analytical method presented in this paper was designed to circumvent that problem- to generate spectral information about excitation and inhibition directly from the second-order Wiener kernel itself and directly from its two subkernels. It is a variation on the spectro-temporal receptive field- one that involves averaging along the main diagonal of the kernel. We test and calibrate the method by applying it to a parametric model with well-defined tuning. Then, for illustrative purposes, we present results of its application to three afferent axons from the auditory papillae of a ranid frog species- *Rana catesbeiana*.

2. The second-order Wiener kernel and its translation to spectro-temporal images

The data used to compute the second-order Wiener kernel comprise sampled versions of continuous, non-repeating broad-band white noise stimuli that have been applied to the ear and sampled versions of the corresponding intracellular potentials recorded from primary afferent axons from the frog's amphibian papilla (AP). For each data set, sampling is carried out at a fixed rate, m samples per s. Thus, in each data set, time (θ) is depicted as a sequence of successive real integers, each representing a finite sampling interval (duration $1/m$). The sampled intracellular potentials are translated into estimates of either the intervals during which the spike peaks occurred or intervals during which the spike potentials first exceeded a criterion threshold value. The second-order Wiener kernel is an $n \times n$ array of real numbers, $h_2(\tau_1,\tau_2)$, derived by second-order cross correlation between the occurrences of spike peaks or threshold crossings and the noise stimulus (Lee and Schetzen, 1965; Marmarelis and Marmarelis, 1978; Schetzen, 1980). The general expression for each element of the array is

$$h_2(\tau_1,\tau_2) = K[\mathrm{Ex}\{s(\theta_{sp}-\tau_1+1)s(\theta_{sp}-\tau_2+1)\} - \mathrm{Ex}\{s(\theta-\tau_1+1)s(\theta-\tau_2+1)\}];\ 1\le\tau_1\le n;\ 1\le\tau_2\le n \quad (1)$$

where $s(\theta)$ is the sampled value of the noise during interval θ; θ_{sp} is an (estimated) interval during which a spike peak or threshold-crossing occurred; Ex{} is the expected value of the quantity enclosed in the braces. The first term on the right side of Eq. 1 is taken over all n-interval stimulus segments that immediately preceded spikes. The second term is taken over all n-interval stimulus segments, and thus can be recognized as the discrete autocorrelation, R_s, of the noise stimulus at large (e.g., see Papoulis, 1977; Van Stokkum et al., 1986):

$$R_s(\tau_1,\tau_2) = \mathrm{Ex}\{s(\tau_1)s(\tau_2)\} \quad (2)$$

The autocorrelation in this case is short-term, being taken over noise segments of duration n (sampling intervals). The first term on the right side of Eq. 1 is the short-term autocorrelation R_{ss} computed over all of the n-element noise segments that immediately preceded spikes. Thus Eq. 1 can be interpreted as the difference between two short-term autocorrelations.

$$h_2(\tau_1,\tau_2) = K[R_{ss}(\tau_1,\tau_2)-R_s(\tau_1,\tau_2)] \quad (3)$$

The array h_2 is symmetric about its main diagonal. Its diagonals can be described as follows:

$$d(k,\delta) = h_2(k,k+|\delta|) = K[R_{ss}(k,k+\delta)-R_s(k,k+\delta)];\ 1\le k\le n-|\delta|;\ -n+1\le\delta\le n-1 \quad (4)$$

where $d(k,0)$ is the kth element of the main diagonal, $d(k,1)$ is the kth element of the first superdiagonal, $d(k,-1)$ is the kth element of the first subdiagonal, and so forth. If the noise-generating process were stationary, which we presume it to be, then $R_s(\tau_1,\tau_2)$ would depend only on the difference, $\tau_1-\tau_2$:

$$R_s(\delta) = \mathrm{Ex}\{s(\tau)s(\tau+\delta)\} \quad (5)$$

Thus, under the presumption of stationarity, the second term on the right side of Eq. 4 can be replaced by $R_s(\delta)$.

$$d(k,\delta) = K[R_{ss}(k,k+\delta)-R_s(\delta)] \quad (6)$$

Because R_{ss} is conditioned on the occurrence of a spike at $\tau_1=\tau_2=1$, one expects it not to equal $R_s(\delta)$, but to reflect instead something special about the noise stimulus producing that spike. In this case, it would be something special about the stimulus segments centered approximately k-1 sampling intervals prior to each spike.

For the patterns of parallel diagonal lines that appear to be ubiquitous in our second-order Wiener kernels from amphibian–papillar axons, $d(k,\delta)$ clearly varies slowly with k, rapidly with δ. In the analysis that follows, we take advantage of this attribute by assuming that the special stimulus features that led to each spike are invariant over small ranges of time prior to the spike. For stimulus segments centered T-1 intervals prior to the spike, for example, those features would be embodied in an idealized function $d(T,\delta)$ that is approximated by the observed functions $d(k,\delta)$ for all values of k in the range

$$T-M \le k \le T+M \tag{7}$$

where M is a positive integer defining the region over which $d(k,\delta)$ is taken to be independent of k. Under that assumption, we can rewrite Eq. 6 as follows:

$$R_{ss}(T,\delta)-R_s(\delta)=\frac{d(T,\delta)}{K};\quad -2M\le\delta\le 2M \tag{8}$$

Fixing T and taking the one-dimensional discrete Fourier transform (DFT) of each of the two autocorrelations translates it into a power spectrum (e.g., see Papoulis, 1977; Bracewell, 1986):

$$S_{ss}(T,\omega)=\mathrm{DFT}\{R_{ss}(T,\delta)\};\quad S_s(\omega)=\mathrm{DFT}\{R_s(\delta)]\} \tag{9}$$

where DFT{} is the DFT of the function contained in the braces. Taking advantage of the additivity property of the Fourier transform, we can combine Eqs. 8 and 9 as follows:

$$S_{ss}(T,\omega)-S_s(\omega)=\Delta S(T,\omega)=$$

$$\mathrm{DFT}\left\{\frac{d(T,\delta)}{K}\right\};\quad -2M\le\delta\le 2M \tag{10}$$

where $\Delta S(T,\omega)$ can be interpreted as the difference between the short-term power spectrum of the noise segments (of length $2M$+1) centered T-1 sampling intervals prior to spikes and the power spectrum of the noise stimulus in general. This power spectral difference tells us what is special about the noise T-1 intervals prior to spike occurrences see (Eggermont et al., 1983a,b,c; Eggermont, 1993). Over a given range of values of T and ω, the power spectral difference $\Delta S(T,\omega,)$ might be positive or negative. In other words, it might show that it is an increase in the power of certain spectral components of the noise segments that is correlated with the occurrences of the spikes, while it is a decrease in the power of other components that is so correlated. The sum of the power spectral difference over all spectral components is given by $d(T,0)/K$.

In order to apply Eq. 10 to our data, we must construct specific algorithms for translating the elements of the computed second-order Wiener kernels into estimates of the idealized functions $d(T,\delta)$. The region of h_2 implied by inequality 7 is a $2M$+1 by $2M$+1 submatrix whose central element is $h_2(T,T)$. One approach would be to equate each element in $d(T,\delta)$ to a single element of the appropriate diagonal of that submatrix (e.g., for $d(T,0)$ select an element of the main diagonal, for $d(T,1)$ select an element of the first superdiagonal, and so forth). One way to do this would be to select the elements at the intersections of the various diagonals with two adjacent counterdiagonals, one of which includes the central element, $h_2(T,T)$, and the other of which includes $h_2(T,T+1)$. The function $d(T,\delta)$ formed in this manner would have the following elements:

$$h_2(T+M,T-M),\ldots,h_2(T+1,T-1),h_2(T+1,T),$$

$$h_2(T,T),h_2(T,T+1),h_2(T-1,T+1),\ldots,$$

$$h_2(T-M,T+M)=d(T,-2M),\ldots,d(T,-2),$$

$$d(T,0),d(T,1),d(T,2),\ldots,d(T,2M) \tag{11}$$

Any element (e.g., $h_2(T,T+\delta)$) computed from the data always will be a noisy estimate of the corresponding ideal element (e.g., $d(T,\delta)$). To improve the estimate of $d(T,\delta)$, we could take the average values of the corresponding elements of several pairs of counterdiagonal. For example,

$$d(T,0)=\frac{1}{3}[h_2(T-1,T-1)+h_2(T,T)+h_2(T+1,T+1)];$$

$$d(T,1)=\frac{1}{3}[h_2(T-1,T)+h_2(T,T+1)+$$

$$h_2(T+1,T+2)]\ \text{etc} \tag{12}$$

For each value of δ this averaging could be extended to all of the elements in the δth diagonal of the $2M$+1 by $2M$+1 submatrix:

$$d(T,\delta)=\frac{1}{2M+1-|\delta|}\sum_{k=-M}^{M-|\delta|}h_2(T+k,T+k+|\delta|) \tag{13}$$

3. Materials and methods

The physiological data used in this paper were taken from afferent auditory axons of the American bullfrog (*R. catesbeiana*). The experimental methods used to obtain the data appear in a previous paper (Yamada and Lewis, 1999)[1]. The $n\times n$ second-order Wiener kernels were computed from the data in the manner described by Eq. 1. In the analysis used for this report, the func-

[1] All animal experiments were performed in accordance with protocols approved by the UC Berkeley Animal Care and Use Committee (protocol # R081-1097).

tion $d(T,\delta)$ was computed in the manner of Eq. 13. This was done for each value of T in the range

$$1 \le T \le n - M \tag{14}$$

For $T < M$, Eq. 13 was replaced by

$$d(T,\delta) = \frac{1}{2T-1-|\delta|} \sum_{k=-T+1}^{T-1-|\delta|} h_2(T+k, T+k+|\delta|) \tag{15}$$

For each value of T, the function $d(T,\delta)$ was zero-padded and translated to a circular coordinate prior to discrete Fourier transformation:

$$d(T,0), d(T,1), d(T,2), ..., d(T,2M), 0, 0,$$

$$..., 0, 0, d(T,-2M), ..., d(T,-2), d(T,-1) =$$

$$g(1), g(2), g(3), ..., g(1024);\ g(k+1024) = g(k) \tag{16}$$

For each value of T, discrete Fourier transformation was carried out on the function in Eq. 16. Because this function is even, the elements of the transform, $K\Delta S(T,\omega)$, all were positive or negative real numbers. These were displayed in three-dimensional plots against ω (frequency) and T-1 (time prior to each spike occurrence).

4. Tests with parametric models

To demonstrate the method and calibrate the resulting power-spectral difference functions, $K\Delta S(T,\omega)$, against known filter functions, we carried out simulated second-order reverse correlation on the model of Fig. 1, with various band-pass filter functions, f_1 and f_2, various low-pass filter functions, f_3, and various spike trigger algorithms (θ). Details of the results depended on the relative shapes and spectral and temporal overlaps of the filter functions and on the nature of the trigger function; but certain quantitative features were observed consistently. We illustrate these features with three examples- one with excitation only, one with inhibition against background noise, and one with both excitation and inhibition plus the same background

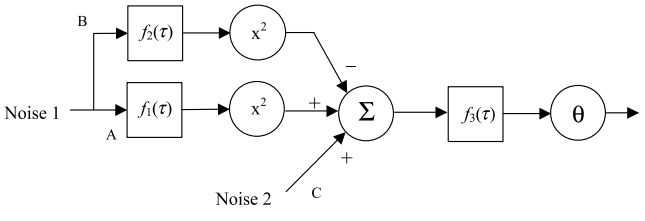

Fig. 1. The parametric model. This is a variation on the LNL (linear/nonlinear/linear) or sandwich model (sometimes called the *general* model) employed widely in theoretical studies of nonlinear system identification (e.g., Billings and Fakhouri, 1978; Hunter and Korenberg, 1986).

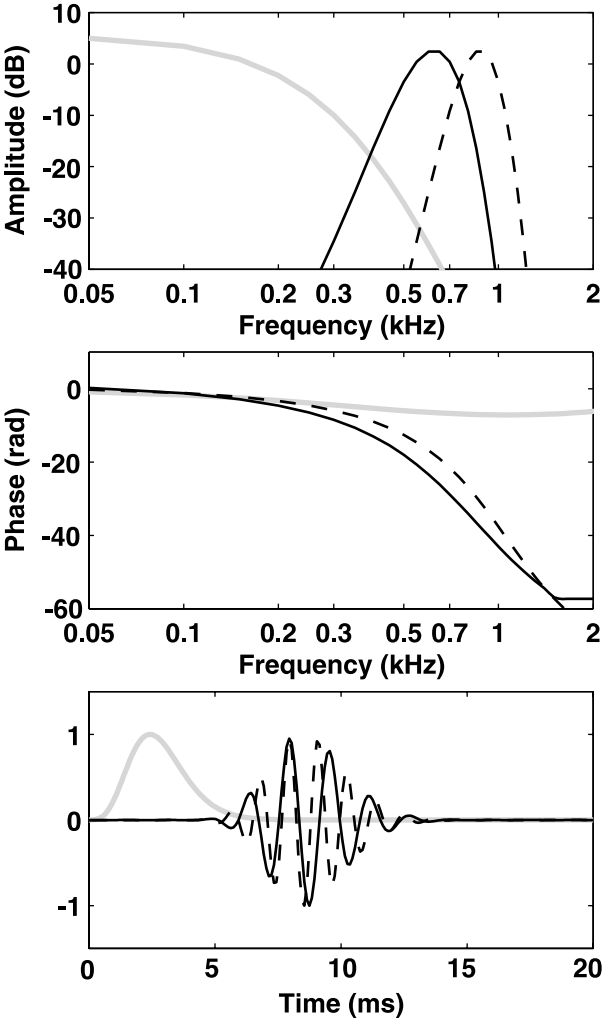

Fig. 2. The filter functions used in models I III and III. The impulse–response waveforms are displayed in the bottom panel; their DFTs are shown in the upper two panels. The excitatory band-pass filter, f_1, is represented by the solid black line; it was constructed as a 625-Hz gammatone. The suppressive filter, f_2, is represented by the dashed black line; it was constructed as a 875-Hz gammatone. The low-pass filter, f_3, is represented by the gray line.

noise. The model itself can be taken to be a primitive representation of the auditory periphery (e.g., see fig. 8 in Eggermont et al., 1983c; fig. 1 in Van Dijk et al., 1997a). Here, however, we use it merely as well-defined parametric model with which to calibrate a nonparametric model (a second-order Wiener kernel) derived from it by white-noise analysis.

Computations on the model are carried out in discrete time. The continuous-time filter functions are replaced by finite-impulse–response approximations. They are the same for all three examples, as are the simulated noise-stimulus waveforms, Noise1 and Noise2 (discrete-time approximations to broadband white-noise waveforms with Gaussian amplitude distribution). All three examples employ the same trigger algorithm to trans-

late the output of the low-pass filter, f_3, into a train of spikes. When the value of the output falls below 0.12, the trigger is armed. During the first subsequent discrete-time interval that the output exceeds 0.15, a spike occurs and the trigger is disarmed. Thus a spike is triggered nearly every time the output passes through the threshold (0.15) with a positive slope. The first example, Model I, contains only path A. The waveform Noise1 is applied only to the input of the excitatory band-pass filter, f_1, the output of which is squared and then applied to the input of the low-pass filter, f_3. The second example, Model II, contains paths B and C. The squared output of the suppressive filter, f_2, is subtracted from the simulated internal noise waveform (Noise2) and the difference is applied to the low-pass filter, f_3. The waveforms Noise1 and Noise2 are statistically independent of one another. The third example, Model III, contains all three paths; and the internal noise waveform (Noise2) is the same as that used in Model II. Because the same Noise1 waveform is used in all three examples, models I and III exhibit identical waveforms at the output of the excitatory filter, f_1, and models II and III exhibit identical waveforms at the output of the suppressive filter, f_2. For each model, the results presented here represent a simulated period of 10 min, with a sampling rate of 10 000 Hz. Prior to being summed, the squared outputs of f_1 and f_2 and the waveform Noise 2 all are normalized to yield maximum absolute values of 1.0. For Noise2, a finite segment of simulated Gaussian noise, this makes the root-mean-square amplitude very small. Before being applied to the threshold process, the output of the low-pass filter is normalized in the same way. The filter functions f_1, f_2 and f_3 are shown along with their DFTs in Fig. 2.

4.1. Excitation only

For Model I, Fig. 3 shows the second-order Wiener kernel (top panel) along with its excitatory and inhibitory subkernels. This 200×200-element Wiener kernel was derived from reverse correlation taken over 26 769 spikes (average simulated spike rate equal to approximately 44.6 sp/s over a simulated period of 600 s). The subkernels were derived by singular-value decomposition of the complete kernel (see Yamada, 1997; Lewis et al., 2002). Excitatory tuning is represented by the bold patch of parallel diagonal line segments centered about 10 ms in the upper panel. The central diagonal line segment comprising only positive values identifies this patch as being, at least in part, excitatory (see Yamada, 1997). The reproduction of the patch entirely within the excitatory subkernel identifies it as being entirely excitatory. The fainter patch, centered about 14 ms in the upper panel, exhibits a central diagonal line segment comprising only negative values and therefore is fully or partly inhibitory. Its reproduction entirely within the inhibitory subkernel identifies it as being fully inhibitory. It is a consequence of the rearming requirement in the trigger algorithm. It disappeared when a pure integrate-and-fire trigger model was used[2]. It returned whenever an absolute refractory period was incorporated into the integrate-and-fire model. It also returned when the integrate-and-fire model was replaced by a two-time-constant version (a Rashevsky–Monnier–Hill model, see MacGregor and Lewis, 1977). The latter includes accommodation- an adjustment of threshold in response to the history of the input to the spike trigger, and fires only on positive slopes. Taking the origin to be the present instant of time (the present sampling interval), one can interpret the top panel of Fig. 3 as follows: The probability of a spike occurring during the present instant is increased by any stimulus component that occurred between approximately 7 and 12 ms ago and whose short-term autocorrelation matched the temporal patterns of the counterdiagonals through the excitatory patch. The probability of such a spike is decreased by any stimulus component that occurred between approximately 12 and 16 ms ago and whose short-term autocorrelation matched the temporal pattern of the counterdiagonals through the inhibitory patch.

The graph of the power spectral difference function, $\Delta S(T,\omega)$, allows one to carry this interpretation into the frequency domain. The top panel of Fig. 4 shows $\Delta S(T,\omega)$ computed according to the algorithm described by Eq. 13–16 for the Wiener kernel in the upper panel of Fig. 3. The half-window, M, for the computation was 30 intervals. The excitatory patch in the kernel has been transformed into a region of positive power spectral difference (on average, those spectro-temporal components were present in the noise, when spikes occurred at time 0). The inhibitory patch has been transformed into a region of negative power spectral difference (on average, those spectro-temporal components were missing in the noise when spikes occurred at time 0). This graph implies that the probability of a spike occurring during the present instant (*now*) is increased by any stimulus components that occurred between approximately 7 and 12 ms ago and had spectral energy in the vicinity of 625 Hz. It also implies that if those same spectral components were present between approximately 12 and 16 ms ago they would reduce the probability of a spike occurring *now*. One can interpret this as follows: occurrence of excitatory spectral components at the earlier time tends to put the output of f_3

[2] The output of f_3 was integrated with respect to time. When the integral reach a designated threshold value, a spike occurred, the value of the integral was set to zero, and the integration process continued.

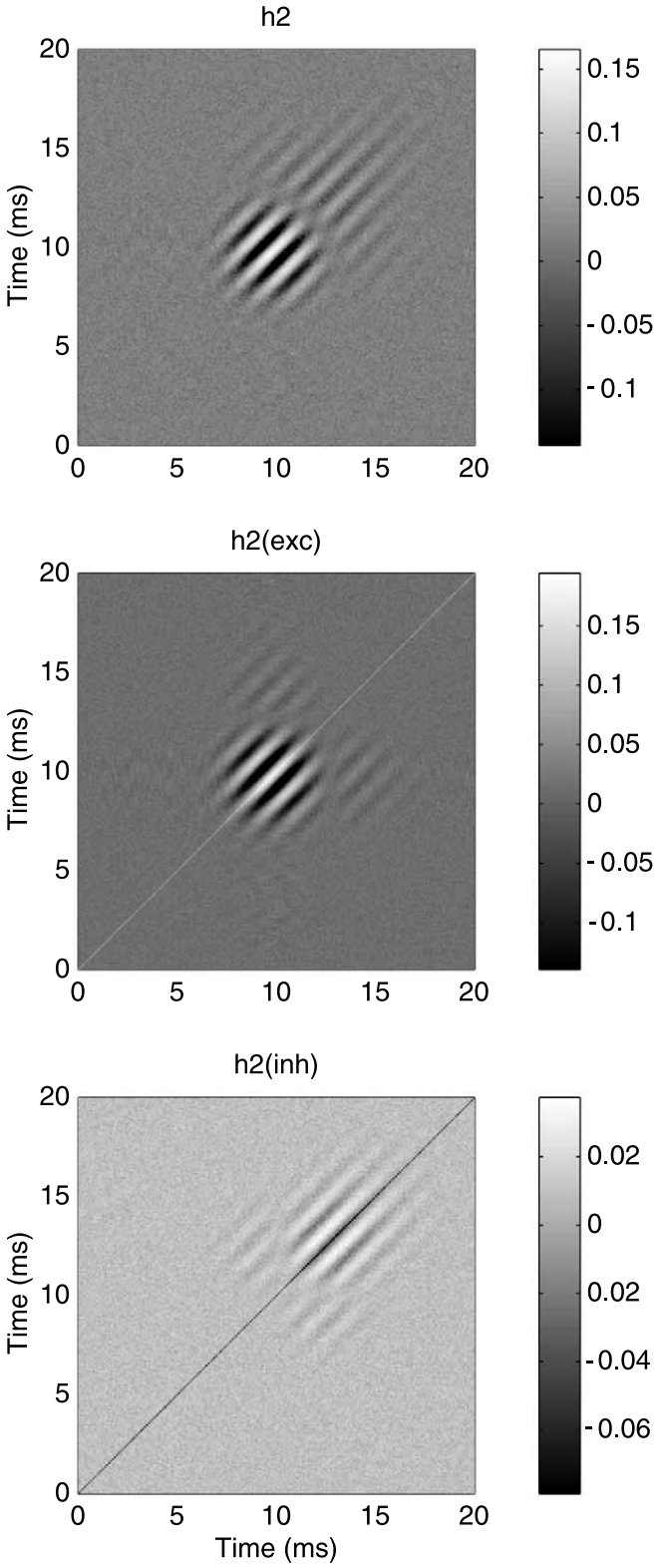

Fig. 3. Second-order Wiener kernel (h_2) and its excitatory and inhibitory subkernels, derived for Model I (excitatory path only).

above threshold, perhaps producing a spike. Whether or not the spike occurs, the sluggish nature of f_3 prevents that output from falling rapidly enough to allow the trigger to rearm in time for intervening excitation to produce a spike *now*. Thus the residual reduction in probability of spiking is consistent with the trigger algorithm of Model I. It would be consistent as well with the presence of absolute refractoriness or accommodation (or both), as our results with alternative trigger models imply (see also the discussion corresponding to fig. 9 in Eggermont, 1993). In fact, the rearming requirement is analogous to refractoriness and is inherent in the Hodgkin–Huxley model. The residual reduction in probability also should be consistent with the presence of adaptation, which we did not include in any of our simulations.

For a second-order Wiener kernel with a single excitatory patch of parallel diagonal line segments, estimates of the underlying filter function (f_1 in this case) are given by the two highest-ranking singular vectors that are tuned and have positive eigenvalues (Yamada, 1997; Lewis et al., 2002). These vectors always are 90° ($\pi/2$ rad) out of phase with one another and thus form a quadrature pair. In this case they are the first - and second-ranking singular vectors (SV1 and SV2, respectively). The third panel of Fig. 5 shows these vectors plotted with f_1. The upper two panels show the corresponding DFTs. It can be seen that the shapes of the tuned waveforms in SV1 and SV2 are very similar to that of f_1, but that they delayed slightly. The delay reflects the trigger algorithm coupled with the dynamics of the low-pass filter. Such delays were present in all of our simulations. Where the amplitude components of the DFTs of SV1 and SV2 were above the noise floor, they match very well that of f_1. Over the frequency range in which that occurs (bottom panel of Fig. 5), the differences, $\phi\{SV1\}$-$\phi\{f_1\}$ and $\phi\{SV2\}$-$\phi\{f_1\}$, between the phase components of the DFTs of SV1 and

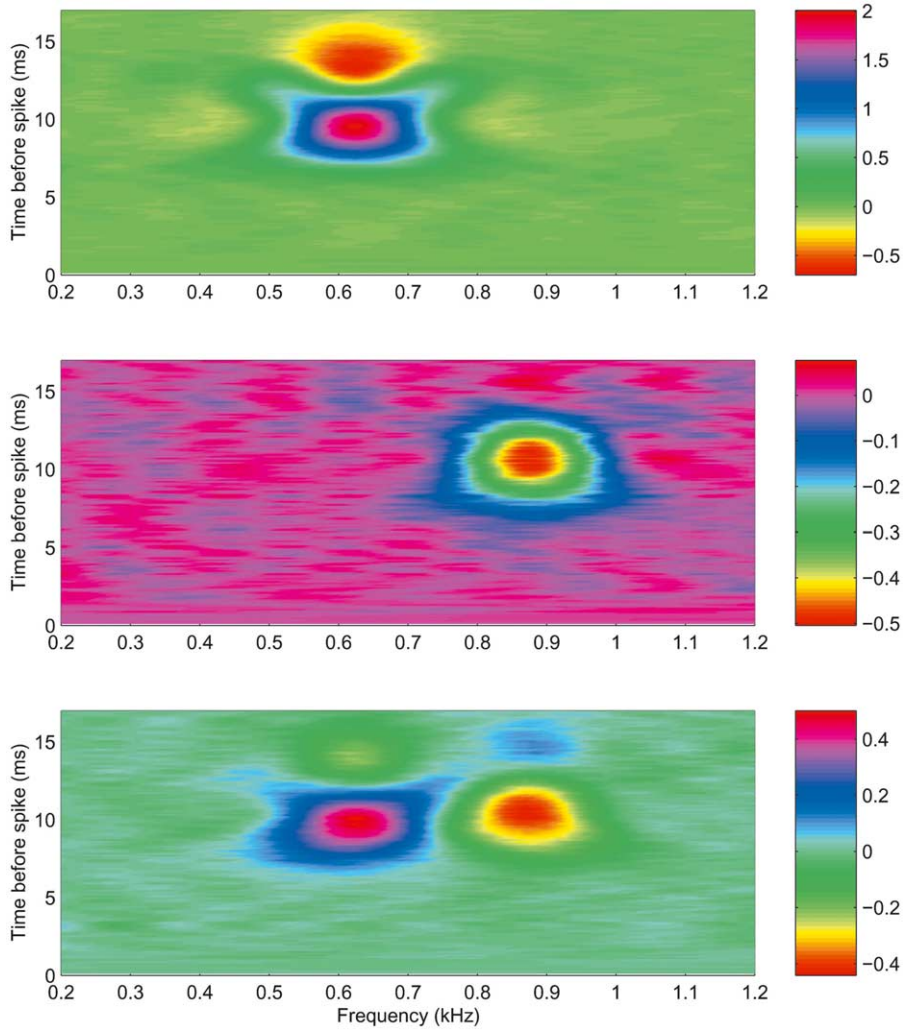

Fig. 4. Power spectral difference functions (spectro-temporal receptive fields) computed for the three configurations of the model in Fig. 1. Top panel: Model I (excitation). Middle panel: Model II (inhibition). Bottom panel: Model III (excitation and inhibition).

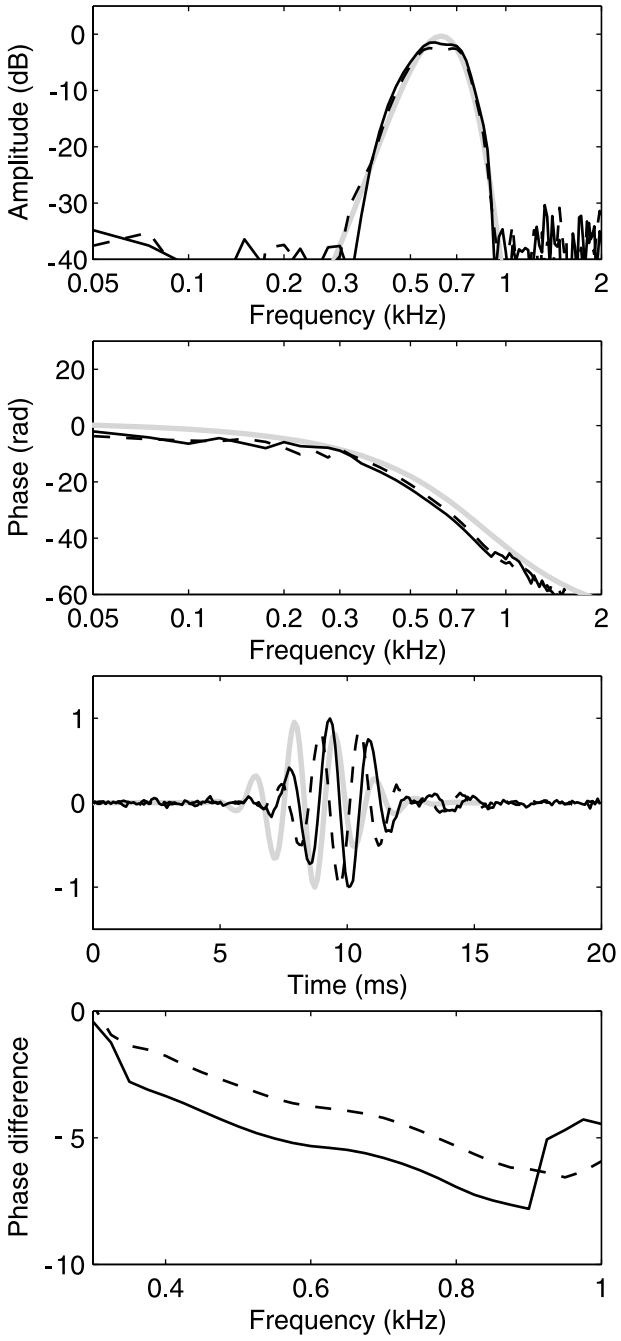

Fig. 5. Top three panels: first-ranking singular vector (SV1, solid black line) and second-ranked singular vector (SV2, dashed black line) of the Wiener kernel (h_2) in Fig. 3, along with the band-pass filter function f_1 (gray line). The second panel shows the phase functions, $\phi\{SV1\}$, $\phi\{SV2\}$ and $\phi\{f_1\}$. Bottom panel: $\phi\{SV1\}$-$\phi\{f_1\}$ (black line) and $\phi\{SV2\}$-$\phi\{f_1\}$ (dashed black line).

SV2 and that of the DFT of f_1 exhibit negative slopes of approximately 0.0013 cycle per Hz, implying a time delay of approximately 1.3 ms between the corresponding tuned waveforms. Over the same frequency range $\phi\{SV1\}$-$\phi\{SV2\}$ is within 1% of $\pi/2$ rad (the vertical distance between the lines in the bottom panel of Fig. 5) reflecting the quadrature relationship between SV1 and SV2.

For a second-order Wiener kernel with a single inhibitory patch of parallel diagonal line segments, estimates of the underlying filter function are given by the two highest-ranking singular vectors that are tuned and have negative eigenvalues (Yamada, 1997; Lewis et al.,

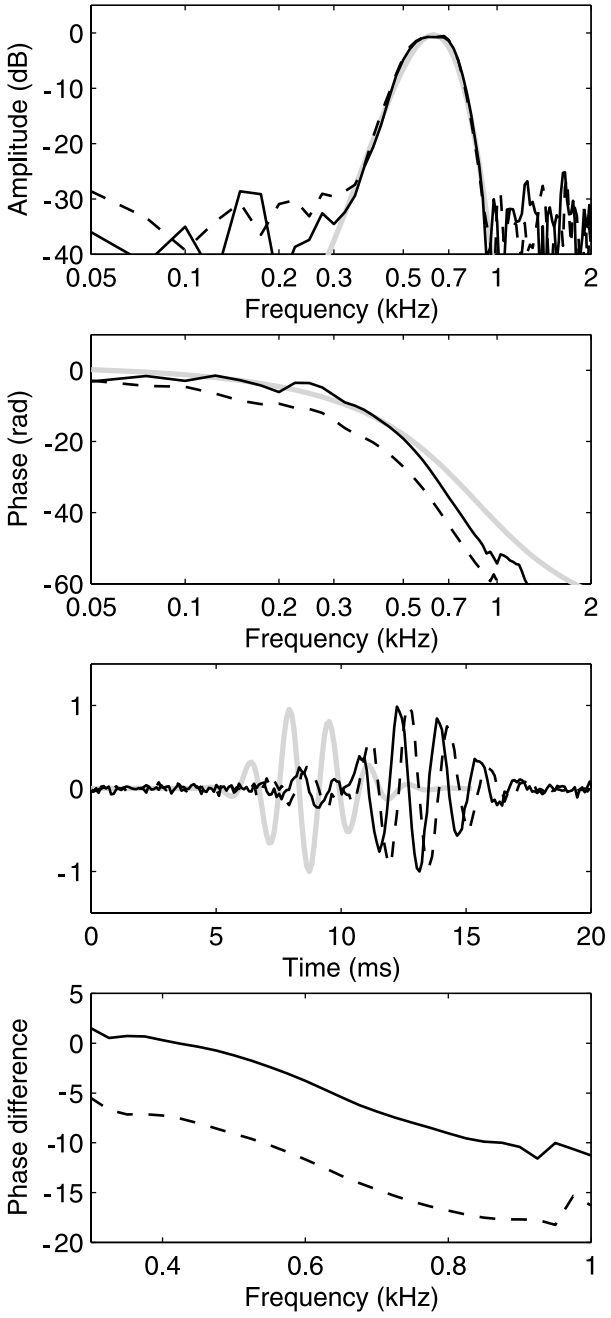

Fig. 6. Top three panels: third-ranking singular vector (SV3, solid black line) and fourth-ranked singular vector (SV4, dashed black line) of the Wiener kernel (h_2) in Fig. 3, along with the band-pass filter function f_1 (gray line). The second panel shows the phase functions, $\phi\{SV3\}$, $\phi\{SV4\}$ and $\phi\{f_1\}$. Bottom panel: $\phi\{SV3\}$-$\phi\{f_1\}$ (black line) and $\phi\{SV4\}$-$\phi\{f_1\}$ (dashed black line).

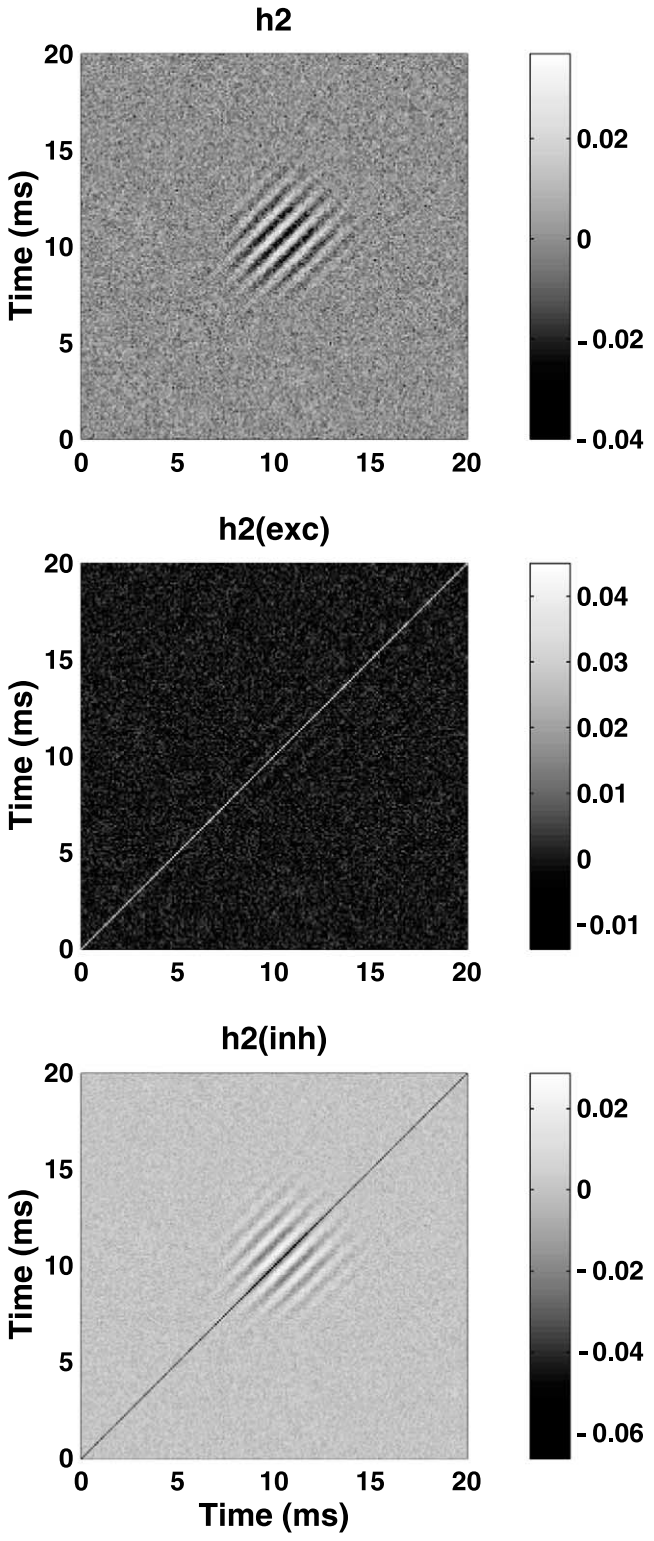

Fig. 7. Second-order Wiener kernel (h_2) and its excitatory and inhibitory subkernels, derived for Model II (inhibition).

2002). The bottom panel of Fig. 6 shows these vectors (SV3 and SV4 in this case) plotted with f_1. Again, the shapes of the tuned waveforms in these singular vectors are very similar to that of f_1, and, above their noise floors, the amplitude components of their DFTs match very well that of f_1. The waveforms in SV3 and SV4, however, are shifted by approximately 4 ms relative to that of f_1. This is consistent with the shift of the inhibitory patch relative to the excitatory patch in the original Wiener kernel. Where the amplitude tuning curves of SV3 and SV4 are above their noise floors, the slopes of the phase differences, $\phi\{SV3\}-\phi\{f_1\}$ and $\phi\{SV4\}-$

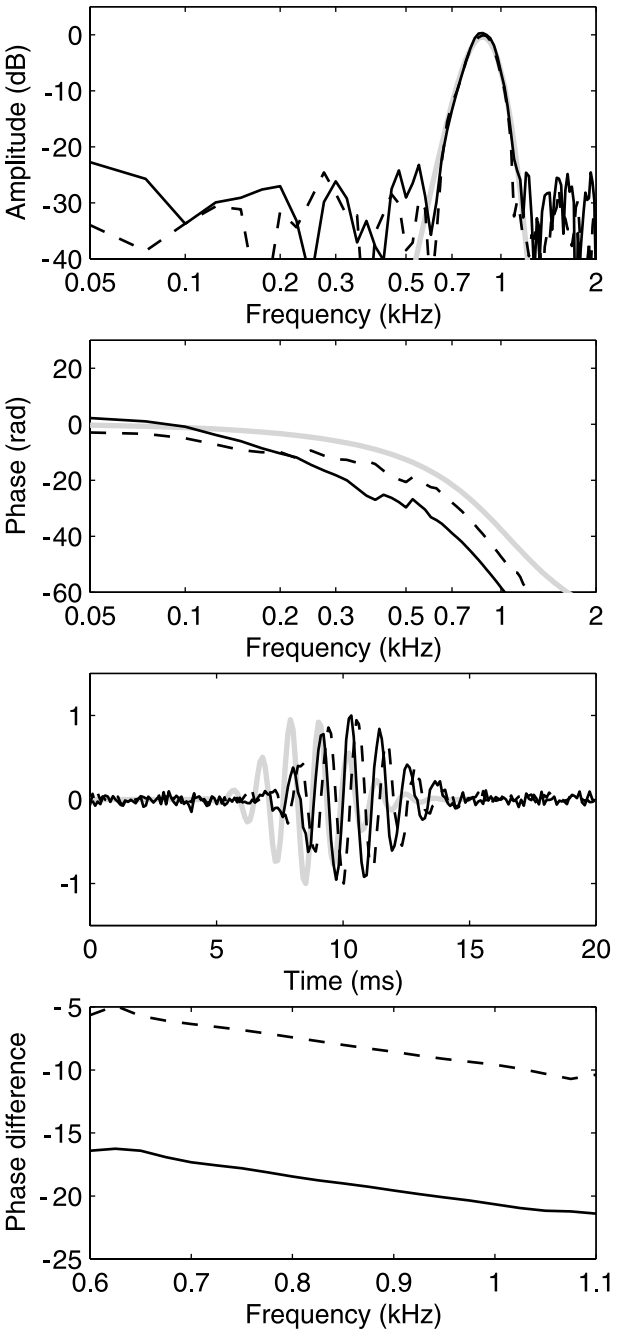

Fig. 8. Top three panels: SV1 (solid black line) and SV2 (dashed black line) of the Wiener kernel (h_2) in Fig. 7, along with f_2 (gray line). Bottom panel: $\phi\{SV1\}$-$\phi\{f_2\}$ (black line) and $\phi\{SV2\}$-$\phi\{f_2\}$ (dashed black line).

$\phi\{f_1\}$, imply a similar delay (bottom panel, Fig. 6). Over that same frequency range, the phase difference $\phi\{SV3\}$–$\phi\{SV4\}$ (vertical distance between the lines in the bottom panel of Fig. 6) is within 1% of $5\pi/2$ rad. With noisy waveforms such as these, the unwrapping of DFT phase data below the well-tuned region (below 0.3 kHz in this case) produces arbitrary offsets of $2n\pi$ rad (where n is an integer). The remainder difference, $\pi/2$ rad, reflects the quadrature relationship between SV3 and SV4.

4.2. Inhibition against background noise

For Model II, Fig. 7 shows the second-order Wiener kernel (top panel) along with its excitatory and inhibitory subkernels. This 200×200-element Wiener kernel was derived from reverse correlation taken over 30 394 spikes (average simulated spike rate equal to approximately 50.7 sp/s over a simulated period of 600 s). The single patch of parallel diagonal line segments represents inhibitory tuning. It is reproduced in the inhibitory subkernel (which is reconstructed from all of the singular vectors with negative eigenvalues). In the power spectral difference function (middle panel of Fig. 4), again computed with $M=30$, the inhibitory patch has been transformed into a region of negative power spectral difference. One can interpret this power spectral difference function as follows: the probability of occurrence of a spike during the present instant is decreased by any stimulus components that occurred between 7 and 14 ms before and had spectral energy in the vicinity of 875 Hz. Once again, we can estimate the underlying (suppressive) filter function, f_2 in this case, by computing the two highest-ranking tuned singular vectors with negative eigenvalues (SV1 and SV2 in this case). The comparison of these waveforms with f_2 (Fig. 8) is representative of our simulations with inhibition only. Above the noise floors of the singular vectors, the amplitude spectra matched well. The tuned parts of the waveforms themselves matched well, but those of the singular vectors were delayed slightly with respect to that of f_2. The same delay was reflected in the phase differences $\phi\{SV1\}$–$\phi\{f_2\}$ and $\phi\{SV2\}$–$\phi\{f_2\}$. Where the DFTs of SV1 and SV2 were well above the noise floor. the difference between $\phi\{SV1\}$ and $\phi\{SV2\}$ was within 1% of 4π–$\pi/2$. This again reflects the $2n\pi$ ambiguity in phase and the quadrature relationship between SV1 and SV2.

4.3. Excitation and inhibition against background noise

For Model III, Fig. 9 shows the second-order Wiener kernel (top panel) along with its excitatory and inhibitory subkernels. This 200×200-element Wiener kernel was derived from reverse correlation taken over 46 292 spikes (average simulated spike rate equal to approximately 77.2 sp/s over a simulated period of 600 s). Once again, we find variously delayed estimates of the band-pass filter functions f_1 and f_2 among the highest-rank-

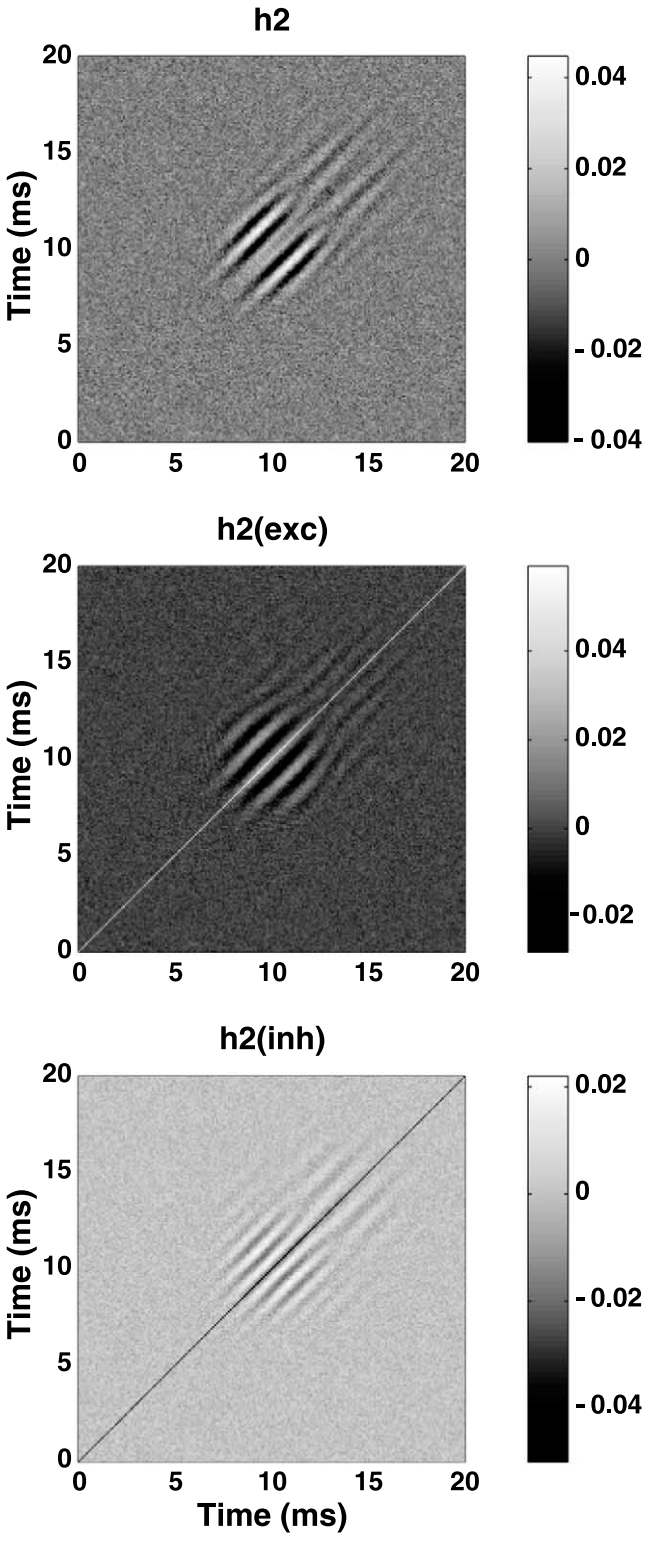

Fig. 9. Second-order Wiener kernel (h_2) and its excitatory and inhibitory subkernels, derived for Model III (excitation and inhibition).

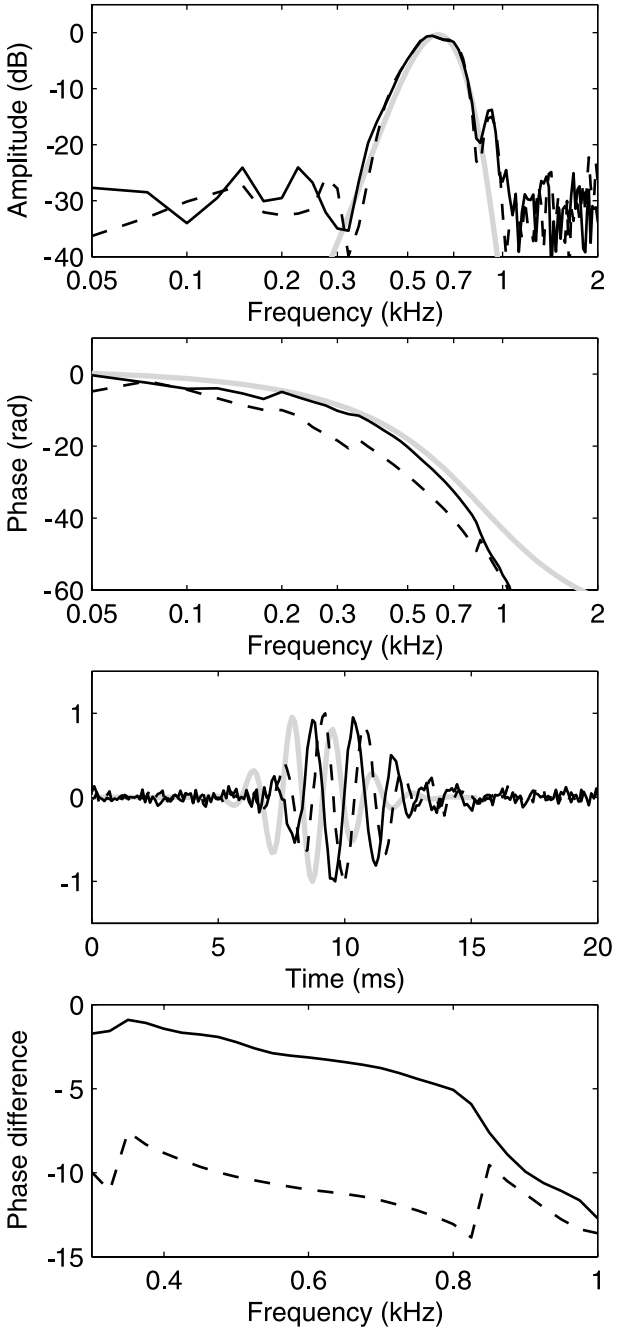

Fig. 10. Top three panels: SV1 (solid black line), SV2 (dashed black line) of the Wiener kernel (h_2) in Fig. 9, along with the band-pass filter function f_1 (gray line). Bottom panel: $\phi\{SV1\}$-$\phi\{f_1\}$ (black line) and $\phi\{SV2\}$-$\phi\{f_1\}$ (dashed black line). From 0.4 to 0.8 kHz, the vertical distance between the two lines in the bottom panel is within 1% of $2\pi+\pi/2$ rad. The slopes of those lines imply a delay of approximately 1.0 to 1.5 ms.

ing, tuned singular vectors (Figs. 10, 11 and 12). The clean separation of the suppressive filter function from the delayed inhibitory effects of the excitatory filter function (owing to the trigger algorithm) in Figs. 11 and 12 is not representative of all of our simulations. The effects of the suppressive filter and the delayed inhibitory effects of the excitatory filter often are merged into single singular vectors; but they nonetheless are represented clearly in the power spectral-difference functions. The bottom panel of Fig. 4 shows the power spectral-difference function computed according to the algorithm described by Eqs. 13–16 for the Wiener ker-

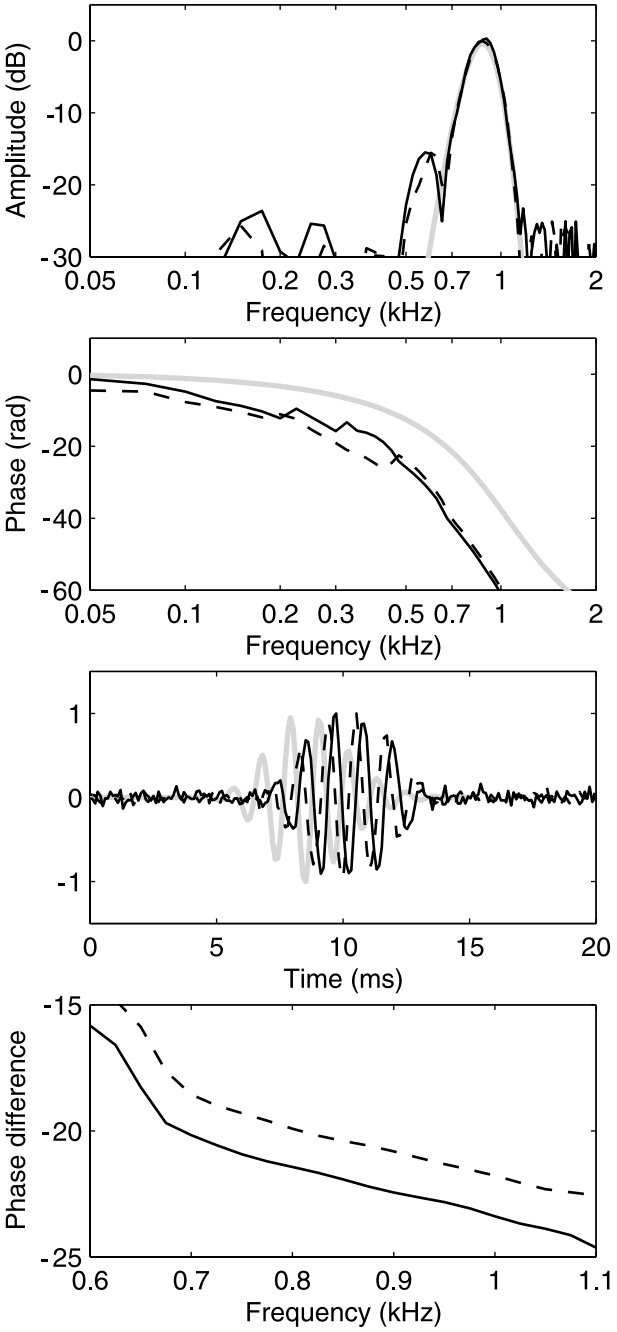

Fig. 11. Top three panels: SV3 (solid black line), SV4 (dashed black line) of the Wiener kernel (h_2) in Fig. 9, along with the band-pass filter function f_2 (gray line). Bottom panel: $\phi\{SV3\}$-$\phi\{f_2\}$ (black line) and $\phi\{SV4\}$-$\phi\{f_2\}$ (dashed black line). From 0.7 to 1.0 kHz, the vertical distance between the two lines in the bottom panel is within 1% of $\pi/2$ rad. The slopes of those lines imply a delay of approximately 1.5 ms.

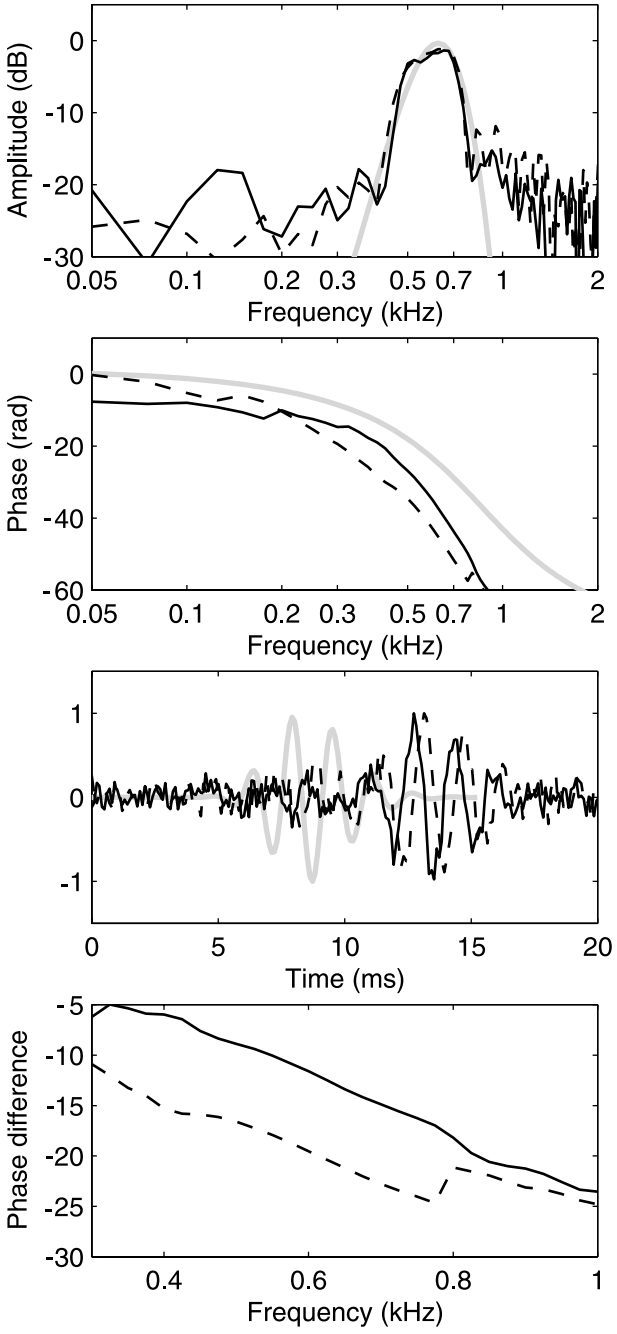

Fig. 12. Top three panels: SV5 (solid black line), SV6 (dashed black line) of the Wiener kernel (h_2) in Fig. 9, along with the band-pass filter function f_1 (gray line). Bottom panel: $\phi\{SV5\}$-$\phi\{f_1\}$ (black line) and $\phi\{SV6\}$-$\phi\{f_1\}$ (dashed black line). From 500 to 700 Hz, the vertical distance between the two lines in the bottom panel is 2π–$\pi/2$ rad. The slopes of those lines imply a delay of approximately 5 ms.

nel in the upper panel of Fig. 9 ($M = 30$). Consistent with all of our simulations, with various threshold algorithms and filter functions variously spaced in the frequency domain, the effects of the excitatory filter are clearly separated from those of the suppressive filter.

4.4. Separation of the spectro-temporal receptive fields for excitation and inhibition/suppression

In the power spectral difference functions of Fig. 4, positive values correspond to spectro-temporal components associated with excitation; negative values correspond to spectro-temporal components associated with inhibition or suppression. In the lower panel of Fig. 13, the negative values in the bottom panel of Fig. 4 are compressed to zero, leaving a plot of the positive values only, i.e., a plot of the excitatory spectro-temporal receptive field for Model III. An alternative procedure is the application of the algorithm described by Eqs. 13–16 directly to the corresponding excitatory subkernel (middle panel of Fig. 9). The result is shown in the upper panel of Fig. 13. In the bottom panel of Fig. 14, the positive values in the bottom panel of Fig. 4 are compressed to zero, leaving a plot of the inhibitory/suppressive receptive field for Model III. The result of applying the algorithm described by Eqs. 13–16 directly to the corresponding inhibitory subkernel (bottom panel of Fig. 9) is shown in the upper panel of Fig. 14. Because no element in the bottom panel of Fig. 4 can be both positive and negative at the same time, the excitatory and inhibitory/suppressive receptive fields depicted in the bottom panels of Figs. 13 and 14, respectively, do not overlap. Those in the upper panels of Figs. 13 and 14, on the other hand, overlap conspicuously at their margins. Given the fidelity with which the dominant eigenvectors (highest-ranking singular vectors) of the excitatory and inhibitory subkernels represent the original filter functions in this simulation, one might conclude reasonably that the subkernels represent the system dynamics well. It follows that the overlapping regions in the upper panels of Figs. 13 and 14 represent spectro-temporal areas of competition between excitation and inhibition/suppression. The lower panels reflect the outcomes of the competition.

4.5. Summary of simulation results

In both the upper and lower panels of Fig. 13, the principal excitatory receptive field is centered approximately at 625 Hz and 10 ms. The corresponding excitatory filter function (Fig. 2) was a 625-Hz gammatone with its envelope centered at approximately 9 ms. Thus, just as we saw in Fig. 10, the second-order Wiener kernel reflects a slightly delayed version of the excitatory filter function. In the lower panel of Fig. 13, the temporal spread of the excitatory receptive field in the vicinity of 625 Hz is approximately 6 ms. In the upper panel it is approximately 7 ms, which matches well the spread of the excitatory filter function in Fig. 2.

In both the upper and lower panels of Fig. 14, the principal inhibitory/suppressive receptive field is cen-

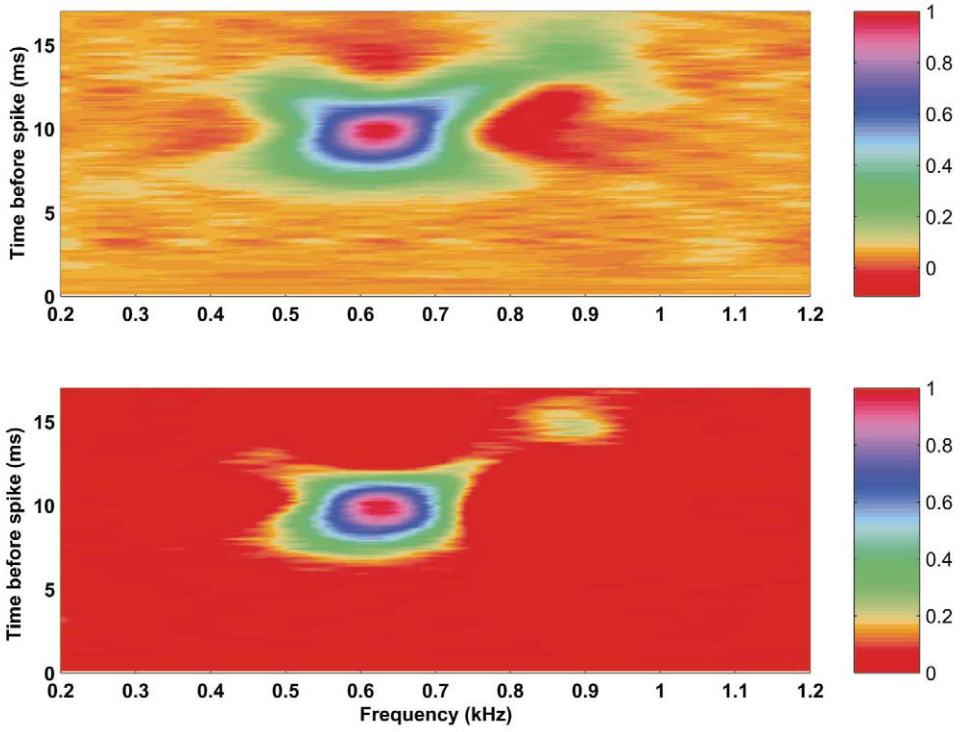

Fig. 13. For Model III, comparison of excitatory spectro-temporal receptive fields derived directly from the entire second-order Wiener kernel (lower panel) and from the excitatory subkernel (upper panel). Upper panel: power spectral difference function derived by application of the algorithm described by Eqs. 13–16 ($M = 30$) to the excitatory subkernel (middle panel of Fig. 9). Lower panel: power spectral difference function derived by application of the algorithm ($M = 30$) to the entire kernel (top panel of Fig. 9). The power spectral difference function in the lower panel was truncated at zero to show positive values only. For each panel, the entire power spectral difference function was normalized to yield a maximum absolute value of 1.0.

tered approximately at 875 Hz and 10.5 ms. The corresponding inhibitory filter function (Fig. 2) is an 875-Hz gammatone with its envelope centered at approximately 9 ms. Thus, just as we saw in Fig. 11, the second-order Wiener kernel reflects a slightly delayed version of the inhibitory filter function. In the lower panel of Fig. 14, the temporal spread of the inhibitory receptive field in the vicinity of 875 Hz is approximately 5 ms. In the upper panel it is approximately 7 ms, which matches well the spread of the inhibitory filter function in Fig. 2.

In all four panels of Figs. 13 and 14, one sees effects of stimulus history prior to the times of the principal receptive fields. In Fig. 14, there is an inhibitory/suppressive historic effect centered at 625 Hz, just as there was in the upper panel of Fig. 4 (for Model I). In Fig. 13, there is an excitatory historic effect centered at 875 Hz. This was not present in Model II. Both historic effects vanished when a pure integrate-and-fire trigger model was used. Both returned when an absolute refractory period was added to the integrate-and-fire model, or when that model was replaced with the two-time-constant model (incorporating accommodation). In the simulation represented here, both historic effects reflect the requirement that the trigger be armed before it can fire.

5. Examples from the frog auditory papillae

The biology of the frog auditory papillae is reviewed extensively in Lewis and Narins (1999). The frog AP is a tonotopically organized sensor with best excitatory frequencies (BEFs) beginning at approximately 100 Hz and ranging three to four octaves upward from there. In the American bullfrog (*R. catesbeiana*), afferent amphibian–papillar axons with BEFs up to approximately 600 Hz exhibit suppression. Responses of those axons to tones at BEF are reduced by the presence of a second tone at a higher frequency (Frishkopf and Goldstein, 1963; Feng et al., 1975; Capranica, 1976). The most effective frequency (best suppressive frequency) for the second tone was found to be between 300 and 800 Hz higher than BEF. In adult animals, the frog basilar papilla (BP) is tuned to a narrow range of frequencies – all units exhibiting nearly the same BEF (Ronken, 1990). In *R. catesbeiana* that BEF lies between 1200 and 1400 Hz (Frishkopf et al., 1968; Feng et al., 1975; Capranica, 1976). Suppression has not been found in frog basilar–papillar units.

Figs. 15–17 show power spectral difference functions (spectro-temporal receptive fields) computed for two AP units (Figs. 15,16) and one BP unit (Fig. 17) from

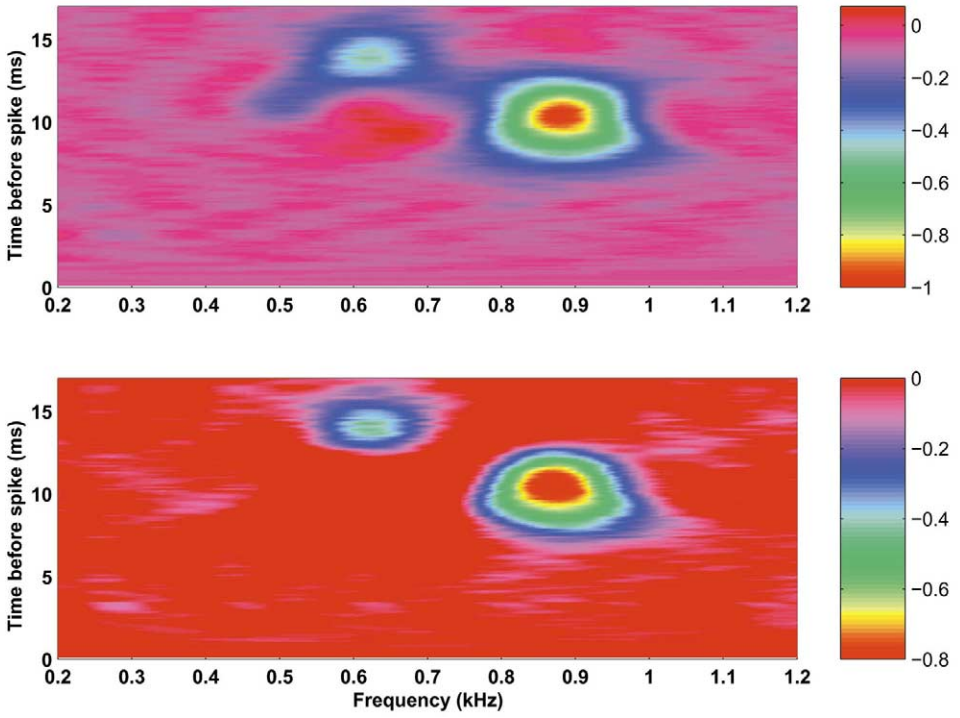

Fig. 14. For Model III, comparison of inhibitory spectro-temporal receptive fields derived directly from the entire second-order Wiener kernel (lower panel) and from the inhibitory subkernel (upper panel). Upper panel: power spectral difference function derived by application of the algorithm described by Eqs. 13–16 ($M = 30$) to the inhibitory subkernel (bottom panel of Fig. 9). Lower panel: power spectral difference function derived by application of the algorithm ($M = 30$) to the entire kernel (top panel of Fig. 9). The power spectral difference function in the lower panel was truncated at zero to show negative values only. Normalization as in Fig. 13.

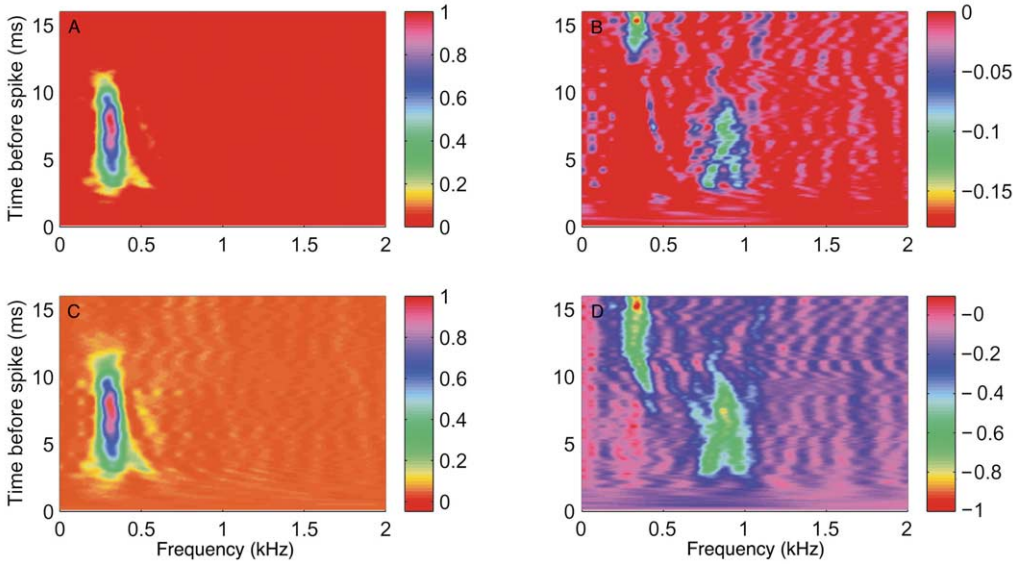

Fig. 15. Excitatory and inhibitory spectro-temporal receptive fields for bullfrog AP unit 042895 #7 (see Figs. 5–9 in Lewis et al., 2002). A: Excitatory power spectral difference function derived from entire second-order Wiener kernel. B: Inhibitory/suppressive power spectral difference function derived from entire second-order Wiener kernel. C: Excitatory power spectral difference function derived from the excitatory subkernel. D: Inhibitory/suppressive power spectral difference function derived from the inhibitory subkernel. In each case, $M = 40$ (selected to span tuned pattern in kernel). Normalization as in Fig. 13.

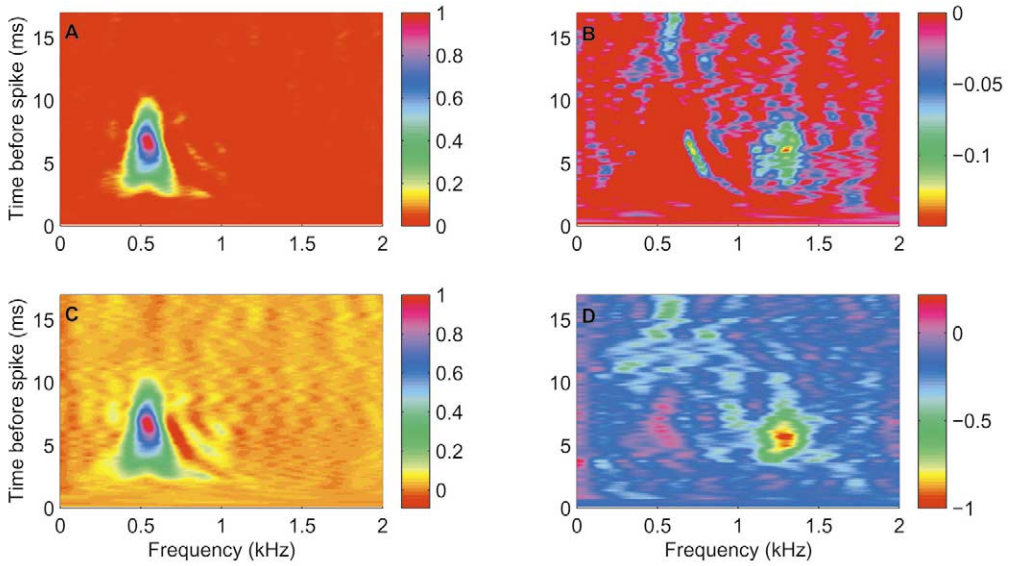

Fig. 16. Excitatory and inhibitory spectro-temporal receptive fields for bullfrog AP unit 041596 #2 (see Figs. 10–14 in Lewis et al., 2002). A: Excitatory power spectral difference function derived from entire second-order Wiener kernel. B: Inhibitory/suppressive power spectral difference function derived from entire second-order Wiener kernel. C: Excitatory power spectral difference function derived from the excitatory subkernel. D: Inhibitory/suppressive power spectral difference function derived from the inhibitory subkernel. In each case, $M = 30$ (selected to span tuned pattern in kernel). Normalization as in Fig. 13.

R. catesbeiana. For the same two AP units, the second-order Wiener kernels, along with their excitatory and inhibitory subkernels and highest-ranking singular vectors, were presented in Lewis et al., 2002. The functions displayed in the upper panels of all three figures were derived by application of the algorithm described by Eq. 13–16 directly to the entire kernel (in the same manner as those of the lower panels in Figs. 13 and 14). The functions displayed in lower panels were derived by application of the algorithm to the excitatory and inhibitory subkernels. The tuning and timing of excitation and inhibition/suppression represented in Figs. 15 and 16 are thoroughly consistent with those implied by the subkernels and highest-ranking singular

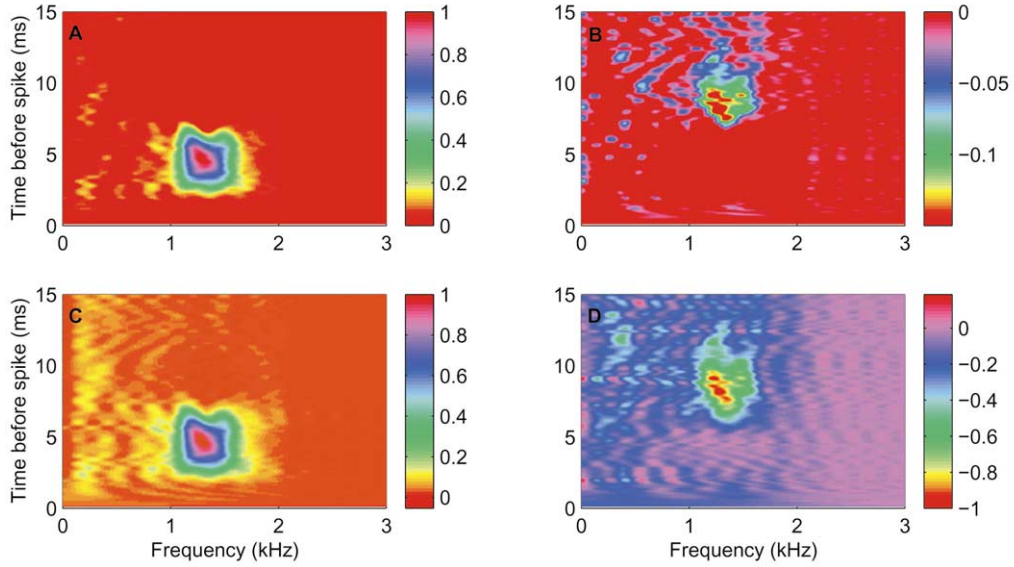

Fig. 17. Excitatory and inhibitory spectro-temporal receptive fields for bullfrog BP unit 050594 #1. A: Excitatory power spectral difference function derived from entire second-order Wiener kernel. B: Inhibitory/suppressive power spectral difference function derived from entire second-order Wiener kernel. C: Excitatory power spectral difference function derived from the excitatory subkernel. D: Inhibitory/suppressive power spectral difference function derived from the inhibitory subkernel. In each case, $M = 20$ (selected to span tuned pattern in kernel). Normalization as in Fig. 13.

vectors presented in the earlier paper (Lewis et al., 2002). The tuning and timing represented in Fig. 17 is thoroughly consistent with that of the highest-ranking singular vectors of the underlying Wiener kernel and subkernels (see Yamada and Lewis, 1999, for kernels and singular vectors from similar units).

All three units exhibit inhibitory/suppressive history effects (identified as such by exhibiting the same tuning as the excitation but occurring earlier in the spectro-temporal receptive field). This likely arises not only from the properties of the spike trigger in the initial segment of the afferent axon, but also from adaptation, which has been observed in both AP units and BP units from *R. catesbeiana* (Megela and Capranica, 1981; Megela, 1984). As expected from their BEFs (approximately 300 Hz and 500 Hz), both AP units exhibit suppression. The best suppressive frequencies were approximately 800 Hz and 1250 Hz. The latter is 250 Hz higher than the highest BEF (1000 Hz) of the *R. catesbeiana* AP. Interestingly, some spontaneous otoacoustic emissions also occur in frogs at frequencies outside the BEF ranges of the AP and BP (Van Dijk et al., 1989, 1996). We would point out that for these three units the width of the principal excitatory receptive field (along the frequency axis) increases with BEF; the height of that receptive field (along the time axis) decreases with increasing BEF. This result is representative of our results from this frog species and from *R. esculenta*. It is complicated, however, by the fact that both width and height are dependent on the stimulus level (the root-mean-square amplitude of the stimulus noise) and on animal temperature. Using spectro-temporal receptive fields derived with the algorithm described by Eq. 13–16, we shall explore these dependencies in a subsequent paper.

6. Conclusions

The spectro-temporal receptive field (power spectrum difference function) provides a clear, unified picture of the tuning and timing of excitation and inhibition/suppression in auditory afferent axons. For most hearing researchers, that picture probably is more easily interpretable than the purely temporal image provided by the second-order Wiener kernel, its subkernels and singular vectors. Derived directly from the entire second-order Wiener kernel, the spectro-temporal receptive field reflects the net results of competition between excitation and inhibition/suppression. Derived separately from the excitatory and inhibitory subkernels, the excitatory and inhibitory/suppressive receptive fields reflect, separately, the phenomena participating in that competition.

Acknowledgements

We are grateful to Walter Yamada. While working in the laboratory of Dr. Lewis, he obtained the frog data used in this paper and pioneered the work on singular-value decomposition of the second-order Wiener kernel. We are grateful as well to two anonymous reviewers who were puzzled by the garbled phase data presented in the original draft of this paper. In response to their puzzlement, we refined the phase data (by zero-padding the ends of the original waveforms prior to DFT), allowing us to construct (and interpret) the phase-difference graphs presented here, in the final version of the paper. P.v.D.'s research is supported by the Netherlands Organization for Scientific Research (NWO) and the Heinsius Houbolt Foundation.

References

Billings, S.A., Fakhouri, S.Y., 1978. Identification of a class of non-linear systems using correlation analysis. Proc. IEE 125, 691–697.
Bracewell, R., 1986. The Fourier Transform and its Applications 2nd Edition. McGraw Hill, New York.
Capranica, R.R., 1976. Morphology and physiology of the auditory system. In: Llinas, R., Precht, W. (Eds.), Frog Neurobiology. Springer, Berlin, pp. 551–575.
Carney, L.H., Yin, C.T., 1988. Temporal coding of resonances by low-frequency auditory nerve fibers: single-fiber responses and a population model. J. Neurophysiol. 60, 1653–1677.
de Boer, E., 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., Kuyper, P., 1968. Triggered correlation. IEEE Trans BME 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.H.J., Johannesma, P.I.M., 1983a. Quantitative characterization procedure for auditory neurons based on the spectro-temporal receptive field. Hear. Res. 10, 167–190.
Eggermont, J.J., Aertsen, A.M.H.J., Johannesma, P.I.M., 1983b. Prediction of the responses of auditory neurons in the midbrain of the grass frog based on the spectro-temporal receptive field. Hear. Res. 10, 167–190.
Eggermont, J.J., Johannesma, P.I.M., Aertsen, A.M.H.J., 1983c. Reverse-correlation methods in auditory research. Q. Rev. Biophys. 16, 341–414.
Evans, E.F., 1989. Cochlear filtering: a view seen through the temporal discharge patterns of single cochlear nerve fibers. In: Wilson, J.P., Kemp, D.T. (Eds.), Cochlear Mechanisms, Structure, Function and Models. Plenum, New York, pp. 241–250.
Feng, A.S., Narins, P.M., Capranica, R.R., 1975. Three populations of primary auditory afferent axons in the bullfrog (*Rana catesbeiana*): their peripheral origins and frequency sensitivities. J. Comp. Physiol. 100, 221–229.
Frishkopf, L.S., Copranica, R.R., Goldstein, M.H., Jr., 1968. Neural coding in the bullfrog's auditory system – a teleological approach. Proc. IEEE 56, 969–980.
Frishkopf, L.S., Goldstein, M.H., Jr., 1963. Responses to acoustic

stimuli from single units in the eighth nerve of the bullfrog. J. Acoust. Soc. Am. 35, 1219–1228.

Hermes, D.J., Aertsen, A.M.H.J., Johannesma, P.I.M., Eggermont, J.J., 1981. Spectro-temporal characteristics of single units in the auditory midbrain of the lightly anaesthetized grass frog (*Rana temporaria* L.) investigated with noise stimuli. Hear. Res. 5, 147–178.

Hunter, I.W., Korenberg, M.J., 1986. The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. Biol. Cybern. 55, 135–144.

Lee, Y.W., Schetzen, M., 1965. Measurement of the Wiener kernels of a nonlinear system by cross-correlation. Int. J. Control 2, 237–254.

Lewis, E.R., Henry, K.R., Yamada, W.M., 2002. Tuning and timing of excitation and inhibition in primary auditory nerve fibers. Hear. Res. 171, 13–31.

Lewis, E.R., Narins, P.M., 1999. The acoustic periphery of amphibians. In: Fay, R.R., Popper, A.N. (Eds.), Comparative Hearing: Fish and Amphibians. Springer, New York, pp. 101–154.

MacGregor, R.J., Lewis, E.R., 1977. Neural Modeling. Plenum, New York.

Marmarelis, P.Z., Marmarelis, V.Z., 1978. Analysis of Physiological Systems: The White Noise Approach. Plenum, New York.

Megela, A.L., 1984. Diversity of adaptation patterns in responses of eighth nerve fibers in the bullfrog, *Rana catesbeiana*. J. Acoust. Soc. Am. 75, 1155–1162.

Megela, A.L., Capranica, R.R., 1981. Response patterns to tone bursts in peripheral auditory systems of anurans. J. Neurophysiol. 46, 465–478.

Møller, A.R., 1986. Systems identification using pseudorandom noise applied to a sensorineural system. Comp. Math. Appl. 12A, 803–814.

Papoulis, A., 1977. Signal Analysis. McGraw-Hill, New York.

Ronken, D.A., 1990. Basic properties of auditory nerve responses from a 'simple' ear: the basilar papilla of the frog. Hear. Res. 47, 63–82.

Schetzen, M., 1980. The Volterra and Wiener Theories of Nonlinear Systems. John Wiley and Sons, New York.

Van Dijk, P., Narins, P.M., Wang, J., 1996. Spontaneous otoacoustic emissions in seven frog species. Hear. Res. 101, 102–112.

Van Dijk, P., Wit, H.P., Segenhout, J.M., 1989. Spontaneous otoacoustic emissions in the European edible frog (*Rana esculenta*): spectral details and temperature dependence. Hear. Res. 42, 273–282.

Van Dijk, P., Wit, H.P., Segenhout, J.M., Tubis, A., 1994. Wiener kernel analysis of inner-ear function in the American bullfrog. J. Acoust. Soc. Am. 95, 904–919.

Van Dijk, P., Wit, H.P., Segenhout, J.M., 1997a. Dissecting the frog inner ear with Gaussian noise. I. Application of high-order Wiener-kernel analysis. Hear. Res. 114, 229–242.

Van Dijk, P., Wit, H.P., Segenhout, J.M., 1997b. Dissecting the frog inner ear with Gaussian noise. II. Temperature-dependence of inner-ear function. Hear. Res. 114, 243–251.

Van Stokkum, I.H.M., Johannesma, P.I.M., Eggermont, J.J., 1986. Representation of time-dependent correlation and recurrence time functions. Biol. Cybernet. 55, 17–24.

Victor, J., Shapley, R., 1980. A method of nonlinear analysis in the frequency domain. Biophys. J. 29, 459–484.

Yamada, W.M., 1997. Second-order Wiener Kernel Analysis of Auditory Afferent Axons of the North American Bullfrog and Mongolian Gerbil Responding to Noise. Doctoral dissertation, Graduate Group in Neurobiology, University of California, Berkeley.

Yamada, W.M., Lewis, E.R., 1999. Predicting the temporal responses of non-phase-locking bullfrog auditory units to complex acoustic waveforms. Hear. Res. 130, 155–170.