Neural Model for Physiological Responses to Frequency and Amplitude Transitions Uncovers Topographical Order in the Auditory Cortex

Alon Fishbach,[1] Yehezkel Yeshurun,[2] and Israel Nelken[3]

[1]*Department of Physiology, Northwestern University, Chicago, Illinois 60611;* [2]*Department of Computer Science, Tel-Aviv University, Ramat-Aviv, Tel-Aviv 69978; and* [3]*Department of Physiology, Hebrew University–Hadassah Medical School and the Interdisciplinary Center for Neural Computation, Hebrew University, Jerusalem 91120, Israel*

Fishbach, Alon, Yehezkel Yeshurun, and Israel Nelken. Neural model for physiological responses to frequency and amplitude transitions uncovers topographical order in the auditory cortex. *J Neurophysiol* 90: 3663–3678, 2003. First published August 27, 2003; 10.1152/jn.00654.2003. We characterize primary auditory cortex (AI) units using a neural model for the detection of frequency and amplitude transitions. The model is a generalization of a model for the detection of amplitude transition. A set of neurons, tuned in the spectrotemporal domain, is created by means of neural delays and frequency filtering. The sensitivity of the model to frequency and amplitude transitions is achieved by applying a 2-dimensional rotatable receptive field to the set of spectrotemporally tuned neurons. We evaluated the model using data recorded in AI of anesthetized ferrets. We show that the model is able to fit the responses of AI units to variety of stimuli, including single tones, delayed 2-tone stimuli and various frequency-modulated tones, using only a small number of parameters. Furthermore, we show that the topographical order in maps of the model parameters is higher than in maps created from response indices extracted directly from the responses to any single stimulus. These results suggest a possible ordered organization of a simple rotatable spectrotemporal receptive field in the mammalian AI.

INTRODUCTION

Fast frequency and amplitude transitions are important features of the auditory input, and many studies have demonstrated the sensitivity of the auditory system to these transients (e.g., for amplitude transitions: Eggermont 1993; Kitzes et al. 1978; Phillips 1988; Rees and Møller 1983; Schreiner and Langner 1988a; Suga 1971; and for frequency transitions: Heil et al. 1992a–c; Kowalski et al. 1995; Mendelson et al. 1993; Nelken and Versnel 2000; Phillips et al. 1985; Schreiner and Mendelson 1990; Shamma et al. 1993; Tian and Rauschecker 1994, 1998).

Many physiological studies suggest that both the timing and the strength of neural responses for amplitude and frequency changes are correlated with the dynamics of the physical change (Barth and Burkard 1993; Heil 1997a,b; Heil and Irvine 1996, 1997, 1998a,b; Nelken and Versnel 2000; Phillips 1988, 1998; Phillips and Burkard 1999; Phillips et al. 1995). These results led us recently to propose a neural model for the detection of amplitude transients. The basic operation of this model is the calculation of the smoothed time derivative of the log-compressed envelope of the stimulus (Fishbach et al. 2001). In that model, a standard neural representation of the auditory input is being progressively delayed by a sequence of neurons that form a "delay layer." The time-derivative computation is implemented as a weighted sum of the activity along the delay layer, with weights that form an ON-OFF receptive field. In that study we have shown that the model is able to reproduce and predict physiological responses to amplitude transients collected at multiple levels of the auditory pathways using a variety of experimental procedures. In addition, we have demonstrated the ability of the model to reproduce the effect of amplitude transients on several psychoacoustical phenomena.

To account for neural responses to frequency transitions we add, in the present work, a spectral dimension to the temporal delay layer by forming an array of neurons, which differ in their best frequencies and temporal delays. The sensitivity of the model to frequency and amplitude transitions is achieved by applying a 2-dimensional receptive field to the spectrotemporally tuned neurons. The receptive field, which in its basic form is a separable function of a Gaussian in the spectral dimension and a 1st-order derivative of a Gaussian in the temporal dimension, can be rotated in the spectrotemporal plane. We compare the model responses to published responses of primary auditory cortex (AI) units to a variety of stimuli such as tone bursts, 2-tone complexes, and linear and exponential FM sweeps. Our results suggest that the model is able to fit the responses of AI units to these stimuli, adjusting only 4 parameters for each unit. Moreover, we demonstrate that the model parameters have a more ordered topographical organization than the standard parameters that are extracted from responses of the units to a single class of stimuli.

These results, along with the topographical order that is being revealed by the 4 parameters of the model, suggest that a simple well-ordered rotatable spectrotemporal receptive field (STRF) may capture some fundamental aspects of the computation performed by neurons in the mammalian AI.

METHODS

Neural model principles

The model is a generalization of a model for the detection of amplitude transitions, which is described in detail elsewhere (Fishbach et al. 2001). In short, the basic operation of that model is the

---

calculation of the first-order time derivative of the log-compressed envelope of the stimulus. To be able to compute the time derivative, we hypothesize the existence of a temporal delay dimension, along which the stimulus is progressively delayed. We construct the delay dimension using a layer of neurons with ascending membrane time constants (τ_m); each neuron is modeled by a standard version of the integrate-and-fire model (I&F). Our I&F makes use of a kernel function in the form

$$K(x) = \frac{1}{\tau_m^2} \; xe^{-x/\tau_m} \quad x \geq 0 \qquad (1)$$

where x represents the time elapsed since the occurrence of a synaptic input. The kernel function, when convolved with the neuron's presynaptic input, determines its postsynaptic potential (Gerstner 1999a). Higher membrane time-constant values induce greater delay in the neuron's response (Agmon-Snir and Segev 1993). The output of the delay layer neurons converges to a single output neuron by a set of connections with various efficacies that reflects a receptive field of a Gaussian derivative. This combination of excitatory and inhibitory connections embodies the operation of time-derivative computation.

In the current model, we add a spectral dimension to the temporal delay layer by forming an array of neurons, which differ in their best frequencies and temporal delays. Accordingly, the one-dimensional (1D) temporal receptive filed is replaced by a 2-dimensional (2D) STRF.

Figure 1 presents a schematic diagram of the model and the flow of data along the different model components. Exact implementation details are given in the APPENDIX. An example of a tone burst with linear ramps, which was one of the inputs tested, is displayed in Fig. 1A.

NEURAL REPRESENTATION. The neural representation (NR) (Fig. 1B) is a simplified approximation of the expected excitatory representation of sound by neurons of a subcortical auditory station; e.g., type II and type III neurons in the ventral cochlear nucleus (Young and Voigt 1982) or type V and type I neurons in the inferior colliculus (IC) (Ramachandran et al. 1999). This representation is generated using a simple preprocessing stage that includes band-pass filtering, demodulation to extract the temporal envelope, nonlinear compression, and low-pass filtering. A bank of 81 band-pass filters are used; the filters' center frequencies f_i are logarithmically spaced around the middle frequency (MF) of the model, which is adjustable. The frequency response magnitude of the ith band-pass filter in dB is given by

$$N_i(x) = -T \left| \log_2 \left(\frac{x}{f_i} \right) \right| \quad f_i = MF \times 2^{0.05(i-40)} \quad i = 1, 2, \ldots, 81 \qquad (2)$$

where T determines the filter bandwidth. The low-pass filtering is achieved by convolving the log-envelope with an alpha kernel function (Eq. 1) with a time constant (τ_m) in the millisecond range.

FREQUENCY-DELAY LAYER. The preprocessed input is fed to the frequency-delay layer of the model, which consists of an array of neurons $U_{i,j}$ with ascending characteristic frequencies (f_i) in one dimension, and ascending membrane time constants (τ_j) in the 2nd dimension. Each unit $U_{i,j}$ receives input from only one NR filter N_i, and represents a population of neurons with identical characteristics. $U_{i,j}$ is modeled as an analog variable by convolving the neuronal representation $N_i(t)$ with a kernel (Gerstner 1999b) whose time constant is τ_j. I&F kernel functions and membrane time-constant values are shown for several units (Fig. 1C). The membrane potential of each neuron in the frequency-delay layer is then saturated using a sigmoidal function

$$S(x) = F_{\max} \left\{ \frac{2}{[1 + e^{-(x/D)}] - 1} \right\} \qquad (3)$$
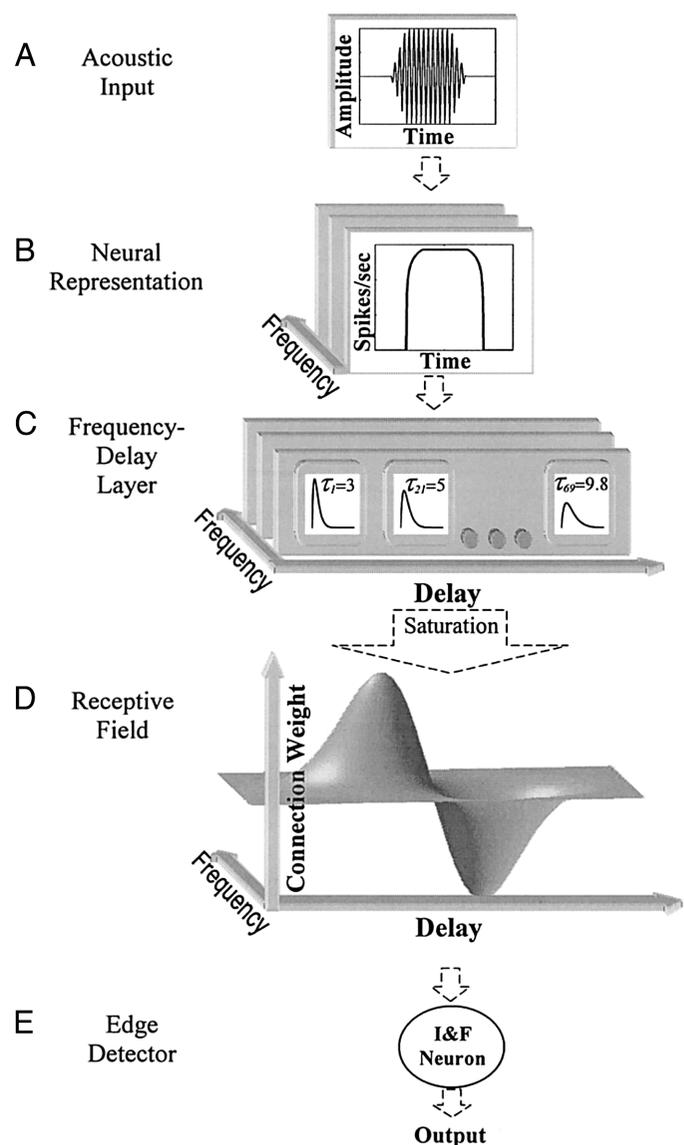
FIG. 1. Schematic diagram of model. See detailed explanation in METHODS.

where $F_{\max}$ is the maximal instantaneous output firing rate (225 spikes/s) and D is a scaling factor that determines the dynamic range of the transformation.

RECEPTIVE FIELD. The frequency-delay layer neurons are connected to a single neuron using inhibitory and excitatory connections with various efficacies (Fig. 1D) that reflect the receptive field shape. For each frequency channel, the receptive field has the same shape as the receptive field of the purely temporal model (Fishbach et al. 2001), that is, an ON-OFF weighting of the delayed versions of the stimulus envelope, implementing a derivative operation on the temporal envelope. However, the delays over which the derivative operation is performed are shifted successively along the delay dimension as a function of frequency. This shift is modeled by the term R_τ (Eq. 4). Finally, the results from each frequency channel are summed up together with weights that depend on the distance (in octaves) from the MF of the receptive field

$$R(i, j) = R_f(f_i)R_\tau(\tau_j, f_i) \qquad (4)$$

where

$$R_f(f_i) = \frac{1}{\sigma_f} e^{-(F/\sigma_f)^2} \quad \text{where} \quad F = \log_2(f_i/MF)$$

$$R_\tau(\tau_j, f_i) = \frac{2}{\sigma_\tau^3} t e^{-(t/\sigma_\tau)^2} \quad \text{where} \quad t = \tau_j - (\mu + \alpha F)$$

and where μ is the mean delay of the receptive field, σ_f is the frequency integration bandwidth, σ_τ is the width of the delay time window, and α is the slope of the frequency-dependent shifts along the delay dimension (in octaves/ms) that produce the sensitivity to the direction of frequency-modulated (FM) tones. We use this shearing transformation of the receptive field instead of a simple rotation transformation because the latter would destroy the temporal characteristics of the STRF in each individual frequency channel. However, because this transformation resembles a rotation transformation in the context of our receptive field, α is termed throughout this study as the pseudo-rotation parameter. Note that the *MF* of the model is not necessarily equal to the best frequency (BF) of the model, as defined by the frequency of a tone burst that yields the lowest intensity threshold; high values of α and σ_f may yield a BF that significantly deviates from the STRF's middle frequency.

EDGE-DETECTOR NEURON. The model's output neuron (Fig. 1*E*) is a single I&F neuron with a noisy integration and membrane time constant τ_{ed} (Gerstner 1999a). Because the model demonstrates sensitivity to sharp transitions in amplitude and frequency, the model's output neuron is referred to as an edge-detector neuron.

PARAMETERS OF THE MODEL. The responses of the model are adjusted to fit the response of a specific neuron by adjusting 4 parameters. Two of the parameters relate to the STRF: α, the pseudo-rotation parameter, and σ_f, the frequency integration bandwidth (*Eq. 4*). The other 2 parameters are the slope (*T*) of the symmetrical band-pass filter of the neural representation N_i (*Eq. 2*), and *D*, the scaling factor of the delay layer saturation transformation (*Eq. 3*). In addition, the middle frequency of the model (*MF*) and the threshold of the edge-detector neuron were varied. However, these parameters were not manipulated independently; instead, their values were always set to best approximate the threshold and the BF of the neuron that was fitted after the values of the 4 adjustable parameters are set. There are 8 additional fixed parameters of the model, 4 of which are parameters of the I&F model. These parameters and their values are listed in full in the APPENDIX. Their specific values have only minor or redundant effect on the responses of the model.

Simulation paradigms and analysis

We evaluated the adequacy of the model to match the responses of AI units by matching the responses of the model with the reported responses of a neuron to a variety of stimuli, using a single set of parameters per neuron. We used the data of Nelken and Versnel (2000), who recorded multiunit clusters in the AI of barbiturate-anesthetized ferrets. To attain a reliable comparison between the experimental and the simulated results, we simulated the responses of the model to the same set of stimuli and used the same data analysis methods as in Nelken and Versnel (2000).

STIMULI. The stimuli are described in detail by Nelken and Versnel (2000). In short, pure tones, of 200-ms duration and 10-ms rise/fall time, were used to characterize the BF of the model (in octaves relative to the model *MF*), and to obtain a rate-level function at the BF. Two-tone stimuli consisted of a 1st tone, which varied in frequency, followed 50 ms later by a 2nd tone at BF and intensity of 20 dB above the model BF threshold; both tones were of 200-ms duration and 10-ms rise/fall time. This paradigm was used to estimate the excitatory and inhibitory response regions of the model.

Nelken and Versnel used 6 different FM stimulation paradigms in 3 pairs. Within each pair, the fine structure of the rate of change of frequency was manipulated, being either linear with frequency (linear

FM sweeps) or linear with log frequency (logarithmic FM sweeps, sometimes also called exponential FM sweeps in the literature). Between pairs of FM paradigms, the coarse structure of the frequency trajectory was manipulated. The 1st FM frequency trajectory started with a fixed low frequency followed by an upward frequency sweep to a fixed high-frequency, which was then followed by a downward sweep back to the low frequency. The sweep range was typically 4 octaves (oct) and its center frequency was chosen around the BF. The total stimulus duration was 500 ms. These paradigms were called by Nelken and Versnel LH-lin (for the linear version) and LH-log (for the logarithmic version). In the 2nd pair of FM paradigms, the stimulus started at the high frequency, swept to the lower one, and then up again. These paradigms were annotated as HL-lin and HL-log. In the last pair of FM paradigms, the stimulus started at the low frequency for 10 ms and than swept up to the high frequency on one trial and on the next trial the stimulus started at the high frequency for 10 ms and then swept to the low frequency. Thus there was a period of silence of about 1 s between the upward and downward sweeps. These paradigms were annotated as N-lin and N-log for the linear and logarithmic versions, respectively. For all the FM paradigms, 10 FM velocities were used, equally distributed between 0.2 and 2 kHz/ms for the linear variants, and between 30 and 300 oct/s for the logarithmic variants.

The responses of the model to the LH and the HL paradigms used by Nelken and Versnel (2000) are identical, which is inconsistent with the experimental results. This is attributed to the fact that the model integrates stimulus information over a time window of at most a few tens of milliseconds, and therefore responds to a frequency transition regardless of the stimulus history that precedes its integration time window. In contrast, cortical activity seemed to undergo a fast adaptation during the stimulation period, presumably because the upward and downward sweeps were presented as part of a continuous sound in these paradigms. Because of this discrepancy between the experimental and the simulated results, the model was tested with N-log and N-lin stimuli only. In cases for which these data were missing, we used the neural responses for the upward sweep of the LH stimulus and the responses to the downward sweep of the HL stimulus. This method of analyzing responses to FM of similar shape was also used by Heil et al. (1992a).

DATA ANALYSIS. The data analysis methods are described in full in Nelken and Versnel (2000), and are reviewed here only briefly. Two-tone responses were summarized as the *M* index, measuring the asymmetry of the side-band inhibition (Shamma et al. 1993)

$$M = \frac{R_{>BF} - R_{<BF}}{R_{>BF} + R_{<BF}}$$

where $R_{>BF}$ and $R_{<BF}$ are the total number of spikes elicited by both tones when the 1st tone is above BF and below BF, respectively.

FM responses were summarized using 2 indices; the 1st is the directional sensitivity (Shamma et al. 1993)

$$C = \frac{R_\downarrow - R_\uparrow}{R_\downarrow + R_\uparrow}$$

where $R_\uparrow$ and $R_\downarrow$ represent the total number of spikes evoked by upward and downward sweeps, respectively.

The second FM index is the velocity sensitivity, which is the weighted average of all the FM velocities, each weighted by the total number of spikes it evoked

$$V = \frac{\sum |x| R(x)}{\sum R(x)}$$

where *x* runs over all velocities tested, and *R(x)* is the total number of spikes evoked by the sweep at that velocity. The directional and velocity sensitivity for the log FM stimuli are annotated as *C-log* and

V-log, respectively, and for the linear FM stimuli they are annotated as *C-lin* and *V-lin*.

Standard deviations of response indices were estimated using boot-strap by sampling with repetitions the responses to each experimental condition. The resulting fictive data sets were summarized using the response indices (*M*, *C*, and *V*). This process was repeated 100 times. The SD for each index was derived from the 25–75% percentile interval of the bootstrap distribution by comparing this interval to the 25–75% percentile interval of a normal distribution with SD = 1. This indirect method was used to overcome possible effects of outliers of the bootstrap distribution.

A brute force search was used to find the parameter set (α, σ_f, T, D) that minimizes the target functionϕ, given by

$$\phi = 0.1\phi_R + 0.9\phi_I$$

where ϕ_R measures the fit error of the model with regard to the raw responses and is given by

$$\phi_R = \sum_{k \in \{\text{stimuli}\}} \sqrt{\frac{1}{N_k} \sum_{i_k=1}^{N_k} (R_k^M(i_k) - R_k^E(i_k))^2}$$

where R_k^M and R_k^E are the total number of spikes elicited by the *k*th stimulus by the model and experimental multiunit cluster, respectively; and N_k is the number of different experimental conditions used for the *k*th stimulus. ϕ_I measures the model's fit error with regard to the response indices (*M*, *C*, *V*, *C-lin*, *V-lin*) and is given by

$$\phi_I = \sqrt{\sum_{k=\{\text{indices}\}} \left[\frac{1}{a_k} (I_k^M - I_k^E)\right]^2}$$

where I_k^M and I_k^E are the model and experimental response indices, respectively, and a_k is the SD of I_k^E. Sensitivity analysis showed that the shape of the target function around its minima was reasonably compact, leading to well-defined parameter values.

Functional maps of the distribution of experimental response indices and of the model parameters on the cortical surface were produced as follows. A Voronoi tessellation of the cortical surface around the penetration points was performed and each of the polygons was shaded according to the value of the point in its center.

The tendency of the experimental response indices and of the model parameters to cluster was measured using a procedure first suggested by Nelken and Versnel (2000) and improved in Rotman et al. (2001). For every pair of penetrations, the topographical distance (d_i) and the absolute difference in their parameter values (p_i) were calculated, thus forming a sequence of $n \times (n - 1)/2$ pairs $S = \{d_i, p_i\}$, where *n* is the total number of penetrations. Next, the pairs were divided into 10 overlapping subsets S_k, $S_1 \subset S_2 \subset \cdots \subset S_{10} = S$, such that S_k contains all the pairs for which d_i is smaller than or equal to the $k \times 10\%$ percentile of all the distances $\{d_i\}$. We denote by $\{d_i^k, p_i^k\}$ pairs such that $\{d_i^k, p_i^k\} \in S_k$.

When a map is highly clustered, it is expected that for small values of *k*, $\{p_i^k\}$ will be small on the average (except for possible large values that may originate at cluster borders), whereas for large values of *k*, $\{p_i^k\}$ will consist of both small and large values. To quantify this, the 80% percentile of $\{p_i^k\}$ is calculated for each *k*. The 80% percentile is used because it captures the behavior of the majority of $\{p_i^k\}$ but is not corrupted by occasional large values of $\{p_i^k\}$ expected at cluster borders. Confidence intervals for this measure were calculated using bootstrap, by randomly distributing the parameter values between the penetrations so that no clustering was possible except by chance.

Statistical tests are considered as significant for $P < 0.05$. In some cases, exact *P* values are reported as a rough indication for the strength of the results.

RESULTS

Responses to tone bursts and comparison with the 1D model

The main new features of the 2D model, which differentiate it from the 1D model, are the adjustable parameters of the model's STRF, α and σ_f. For the STRF in its basic form ($\alpha = 0$) the responses of the 2 models to any stimulus with stationary spectra are comparable. This implies that the 2D model is able to reproduce the responses of the 1D model to amplitude transitions for STRF with $\alpha = 0$. Because we use a pseudo-rotation that maintains the temporal characteristics at each frequency channel of the STRF (see *Eq. 4*), the 2D model reproduces the responses of the 1D model to amplitude transitions of spectrotemporally separable sounds for any value of α.

Responses to frequency sweeps

The model is capable of reproducing many of the physiological characteristics of responses of AI neurons to linear and exponential FM sweeps. In particular, the model is capable of showing various degrees of sensitivity to the velocity and to the direction of the FM sweep.

Figure 2 illustrates the way the model responds to FM sweeps and the effect of the STRF on the directional sensitivity of the model. The frequency trajectory of an upward N-log FM sweep is shown in Fig. 2*A*. For clarity, we consider a simplified configuration of the model, which consists of 2 NR filters (Fig. 2*B*), 2 frequencies by 3 delays (2 × 3) frequency-delay layer (Fig. 2*C*) and accordingly a 2 × 3 STRF (Fig. 2*D*). The best frequencies of the lower and upper NR filters are indicated in Fig. 2*A* by the lower and upper dashed lines, respectively. Given that the simplified NR filters are identical except for a translation on a log–frequency axis, their responses to the stimulus are identical except for a time shift that corresponds to the frequency shift between the center frequencies of the 2 filters divided by the FM velocity (Fig. 2*B*).

The simplified frequency-delay layer consists of 2 series of neurons with time constants of 3, 5, and 7 ms (Fig. 2*C*); each series is fed with the output of one NR filter. Figure 2*D* illustrates 3 different STRFs with different α values; $\alpha = 0$ (STRF1), $\alpha > 0$ (STRF2), and $\alpha < 0$ (STRF3). Note that the STRFs for $\alpha \neq 0$ are not rotations of the STRF with $\alpha = 0$: such rotations would cause the weights in all the "corners" of STRF2 and STRF3 to be nonzero. Instead, the weights for the 2 frequency channels are shifted with respect to each other, such that for $\alpha > 0$, higher-frequency channels have weights shifted to larger delay values, whereas for $\alpha < 0$, higher-frequency channels have weights shifted to smaller delay values. The responses of neurons at the 2 center frequencies, but with different delays, are identical except for the time shift that is caused by the NR filters. Zero or positive α (STRF1 and STRF2) preserve or enlarge the time shift caused by the upward frequency change of the stimulus. On the other hand, negative α (STRF3) compensates for this time shift, aligning the time derivatives calculated in each frequency series of the delay neurons. This synchronization causes the membrane potential of the edge-detector neuron to peak at a higher value. Thus for negative α, the total area of the membrane potential above a high enough threshold (which is monotonically related
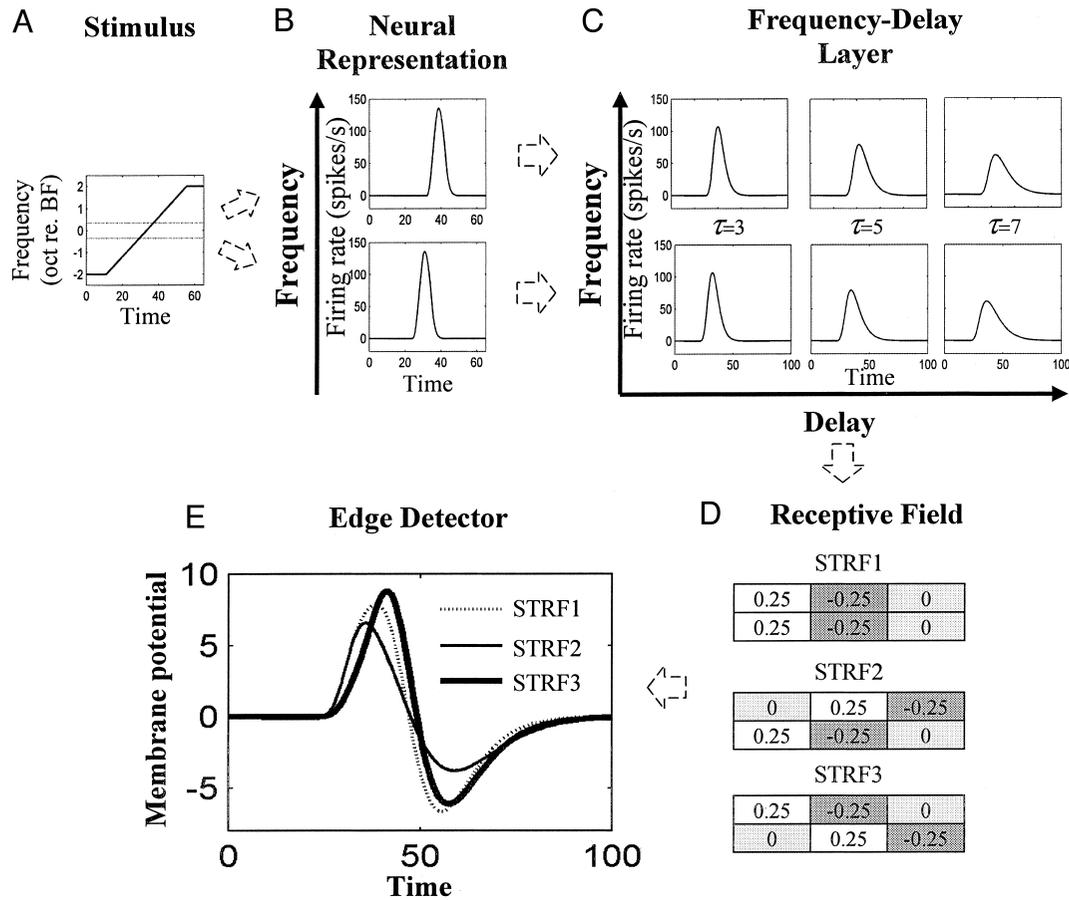
FIG. 2. Responses of model components to frequency transitions, as illustrated by cartoon configuration of model [2 neural representation (NR) filters and 2 × 3 delay neurons]. Frequency trajectory of an upward N-log frequency-modulated (FM) sweep is presented in *A*. Sweep evokes identical responses at NR filters (*B*), with a time delay that corresponds to difference in filters' center frequencies (dashed lines in *A*) divided by velocity of sweep. Responses of each NR filter are progressively delayed by frequency-delay neurons of same center frequency (*C*). Output of these neurons converges to edge-detector neuron (*E*) by connections that reflect spectrotemporal receptive field (STRF). Three examples of STRF with different *a* values are presented in *D*. Different *a* values do not change STRF *within* each frequency band; rather they affect temporal relation *between* frequency bands. Negative *a* (STRF3) introduces time delay that compensates for time delay originated at NR level. This causes STRF3 to better charge the edge-detector membrane (thick line in *E*) than STRF1 and STRF2.

to the expected number of spikes) will be larger (Fig. 2*E*). Although negative α generally yields model neurons that are sensitive to upward sweeps (*negative C* according to the definitions used here), the direction sensitivity of the model is a complex function of all parameters and, under some conditions, a downward-sensitive neuron can be generated with negative α. For example, a small value of *D*, the scaling factor of the delay layer saturation transformation (*Eq. 3*), combined with a narrow NR filter bandwidth (*Eq. 2*) and negative α, may yield a very short excitation of the edge detector in response to an upward FM sweep, generating only a small number of spikes. In such cases, a downward FM sweep will cause a more prolonged excitation period. Because of the refractoriness of the edge-detector neuron, such a set of parameters may produce more spikes for a downward than for an upward FM sweep.

An important property of the model, which is illustrated by Fig. 2, is that the detection of frequency and amplitude transitions uses the same basic computation, that is, calculating the time derivative of the envelope of a band-pass filtered version of the input. This computational principle may explain some

reported relations between the responses of cortical neurons to frequency transitions and amplitude transitions. Phillips et al. (1985) recorded the responses of cat AI units to narrow excursion FM sweeps. Each sweep consists of a 200-ms tone of a fixed frequency with which the sweep started, a linear FM sweep of 2 kHz, and then a 400-ms tone of a fixed frequency with which the sweep ended. The center frequency of the sweep was varied over a 4–9 kHz range centered at the unit's BF. Phillips et al. report that the directional sensitivity of AI neurons to narrow excursion FM sweeps depends on the center frequency of the sweep. In particular, they found a good congruence between the directional sensitivity, measured at some center frequency, and the slope of the frequency response map, measured at that frequency. Figure 3*A* illustrates this finding, as replotted from Phillips et al. (1985). Figure 3*B* demonstrates that the model reproduces this phenomenon very accurately. In the barbiturate-anesthetized cats, most responses to pure tones occur at sound onset. The model generates these responses as derivatives of the instantaneous amplitude envelope of the tones, and therefore the frequency tuning curve measures the amplitude of the temporal derivatives in each
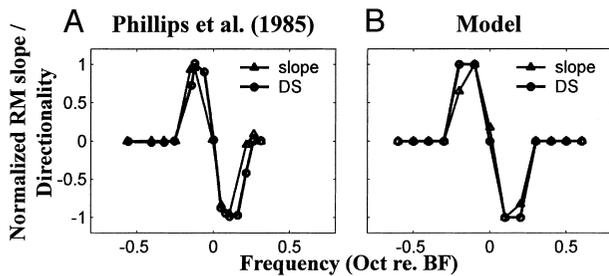
FIG. 3. Directional sensitivity for narrow-excursion FM sweep as a function of sweep center frequency is correlated with slope of frequency–response map of cat primary cortex (AI) neuron, as replotted from Phillips et al. (1985) (*A*). Model reproduces this phenomena (*B*) by suggesting a common mechanism for sensitivity of AI neurons to amplitude transitions (at tone onset) and for narrow-excursion frequency transition. Slope of frequency–response map is normalized to range [−1, 1].

frequency channel. The model accounts for the data of Phillips et al. (1985) by summing up such temporal derivatives in each frequency channel, thus linking the computations underlying the detection of frequency transitions and amplitude transitions in AI neurons.

Responses to two-tone complexes

The model responses to 2-tone complexes are suppressed when the OFF-BF tone is close to the model's BF, in agreement with the responses of AI neurons to these stimuli. Moreover, the model is able to show a full range of asymmetrical responses with respect to the frequency of the OFF-BF tone. Several factors are believed to contribute to the 2-tone suppression phenomenon along the auditory pathway. At the level of the auditory nerve the 2-tone suppression phenomenon is thought to be related to nonlinear effects of the basilar membrane, whereas at higher levels of the auditory pathway neural inhibitory mechanism between neuron with adjacent BFs are thought to enhance the suppression effect.

The simplified neural representation of the model does not include nonlinearities that can account for 2-tone suppression, and hence the model's 2-tone suppression is formed elsewhere. In fact, the model supplies an alternative explanation to the enhancement of the 2-tone suppression by higher levels of the auditory pathways. Figure 4 illustrates the way the model gives rise to the 2-tone suppression phenomenon. Figure 4, *A*, *B*, and *C* sketch 3 different scenarios of 2-tone stimuli used by Shamma and his collaborates (Kowalski et al. 1995; Shamma et al. 1993) and by Nelken and Versnel (2000). In all 3 cases the BF tone (T2) starts 50 ms after the OFF-BF tone (T1). The frequency of T1 is far from BF in Fig. 4*A*, close to BF in Fig. 4*B*, and on BF in Fig. 4*C*. For simplicity, we consider a simplified configuration of the model in which the only significant input to the model comes from the NR filter at the model *MF* (very small σ_f; see *Eq. 4*). The output of this NR filter to the 3 different 2-tone combinations is plotted in Fig. 4, *D*, *E*, and *F*. In Fig. 4*D* the T1 tone has no effect on the output of the NR filter because its frequency is far from BF. The response of the model is therefore attributed only to the delayed T2 tone. When the frequency of T1 is near (but not at) BF, its effect on the NR filter output is larger but still attenuated (Fig. 4*E*). However, because of the log compression of the neural representation the amplitude envelope of the combined tones is almost the same as of that of T2 when presented alone. After

the temporal differentiation performed by the STRF, the membrane potential of the edge-detection neuron will show 2 positive deflections (Fig. 4*H*), each smaller than the deflection caused by T2 alone (Fig. 4*G*). When the frequency of T1 is at BF, it is not attenuated at all by the NR filter, and because of the log compression of the NR, T2 has almost no effect on the output of the NR filter (Fig. 4*F*). This causes the model responses in this case (Fig. 4*I*) to be dominated by the responses to the T1 tone. These are similar to the model responses when the T1 frequency is far from BF (Fig. 4*G*), except for an advance in latency of 50 ms. Figure 4*J* shows a schematic diagram, replotted from Shamma et al. (1993), that shows typical responses of an AI neuron to such 2-tone stimuli. The recorded spikes are attributed to T2 or T1 according to their latencies (thin and thick lines, respectively). Taking into consideration that the simulated spikes are proportional to the integrated membrane potential of the edge-detector neuron above some positive threshold level (dashed line in Fig. 4,
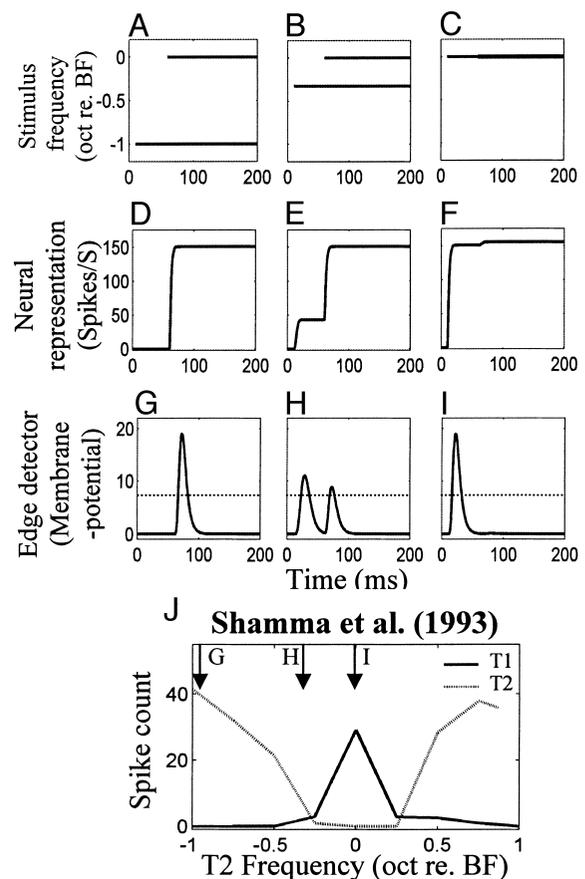
FIG. 4. Responses of model to delayed 2-tone complexes, as illustrated by cartoon configuration of model (one NR filter). Three typical 2-tone complexes are discussed, in which 1st tone (T1) is far from model's best fit (BF) (*A*), close to BF (*B*), and at BF (*C*). Frequency of 2nd tone is fixed at BF. Because of log-compression, responses of NR filter to steady-state complex differ only slightly from its responses to a single BF tone (*D*, *E*, and *F* after onset of T2). Time courses of NR filter responses to 3 scenarios differ because of frequency-dependent attenuation of T1. Membrane potential of edge-detector neuron (*G*, *H*, and *I*) reflects time-derivative computation performed by frequency-delay layer and STRF. Responses of model are proportional to integrated membrane potential above some positive threshold (dashed line in *G*, *H*, and *I*) and are consistent with typical responses of AI neurons to same stimuli [*J*, as replotted from Shamma et al. (1993)]. Arrows in *J* mark frequency of T2 used in corresponding simulated scenarios.
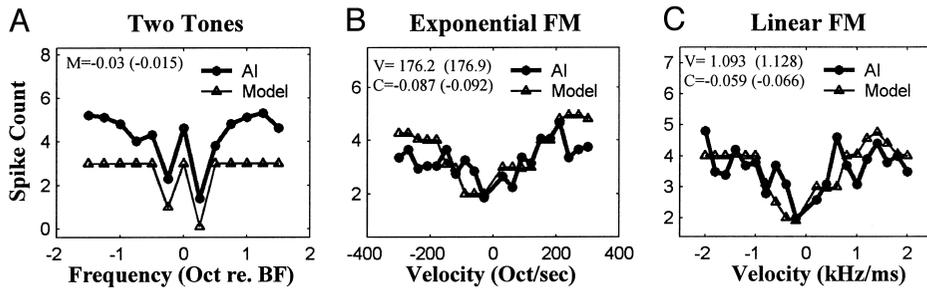
FIG. 5. Simulated vs. experimental responses [replotted from Nelken and Versnel (2000)] to 2-tone complexes (*A*), exponential FM sweeps (*B*), and linear FM sweeps (*C*). Experimental (simulated) response indices are specified within relevant subplot. Experimental data are of multicluster 21a of *experiment 166*.

G–I), it is clear that the model follows very closely both the timing and the strength of responses of AI neurons to 2-tone complexes (compare the simulated responses to T2 and T1 in Fig. 4, *G–I* with the experimental responses at T1 frequencies pointed by the arrows in Fig. 4*J*). The description of the model responses to 2-tone complexes becomes more complicated when the STRF integrates NR filters other than the *MF* filter. In these cases, α plays a role in determining the asymmetry of the suppression with respect to T2 frequency, by synchronizing responses across NR filters in a way similar to the effect α has on the directional sensitivity of the model (Fig. 2).

Evaluation of the model: fit of AI multiunit data

Nelken and Versnel (2000) recorded a total of 212 multiunit clusters in the AI of 6 ferrets. For the purpose of fitting the model responses to the cluster responses we considered only clusters whose responses to a sufficient stimuli set were recorded (single tones, 2-tone complexes, and at least one pair out of the 3 FM stimulation paradigms). After this screening process, we discarded data from 3 ferrets with <10 multiunit clusters per ferret. This left a total of 74 multiunit clusters from 3 ferrets (*experiments 162*, *166*, and *168* of Nelken and Versnel 2000). The responses of each multiunit cluster to the entire set of stimuli (see METHODS) were fitted by adjusting the 4 adjustable parameters of the model. Figure 5 demonstrates the agreement between the experimental responses of one AI cluster and the simulated responses of the model with one set of parameters that best fitted that cluster. Although there may be differences in the overall strength of response between the experimental and simulated responses (Fig. 5*A*), the model follows the general pattern of the experimental responses and matches reasonably well the experimental response indices that summarized them (*M*, *C*, and *V*).

Figure 6 shows the ability of the model to fit the 74 multiunit clusters by comparing the simulated versus experimental response indices (*M*, *C*, and *V*) that summarize the responses to 2-tone complexes and to linear and logarithmic FM sweeps. The vertical error bars represent the SD of the experimentally derived indices as estimated by bootstrap (see METHODS). The SD for the simulated data are not shown because they are arbitrarily determined by the integration noise of the edge-detector neuron (Fig. 1*E*). The model is able to fit well all of the experimental indices except for the linear velocity sensitivity (Fig. 6*D*), for which the model shows a small but consistent undershoot.

The distribution of the 74 values of each of the 4 model parameters used to fit the 74 multiunit clusters is shown in Fig. 7. The histograms of σ_f, *D*, and *T* (Fig. 7, *A*, *C*, and *D*, respectively) are plotted on a log-scale, whereas Fig. 7*B* plots

the histogram of the absolute values of α. The circles in each plot indicate the parameter values that divide each distribution to groups of equal size. The use of these values is described below. The σ_f and α parameters, describing the shape of the STRF, are unimodally distributed. The parameter *T*, which determines the frequency selectivity of the NR filters, is clearly bimodally distributed. The parameter *D*, which determines the saturation level of the delay neurons, also shows a tendency for 2 modes. The 4 adjustable parameters are pairwise uncorrelated except for *T* and σ_f, which are weakly correlated on a log–log scale ($r_{72} = 0.26$, $P < 0.05$).
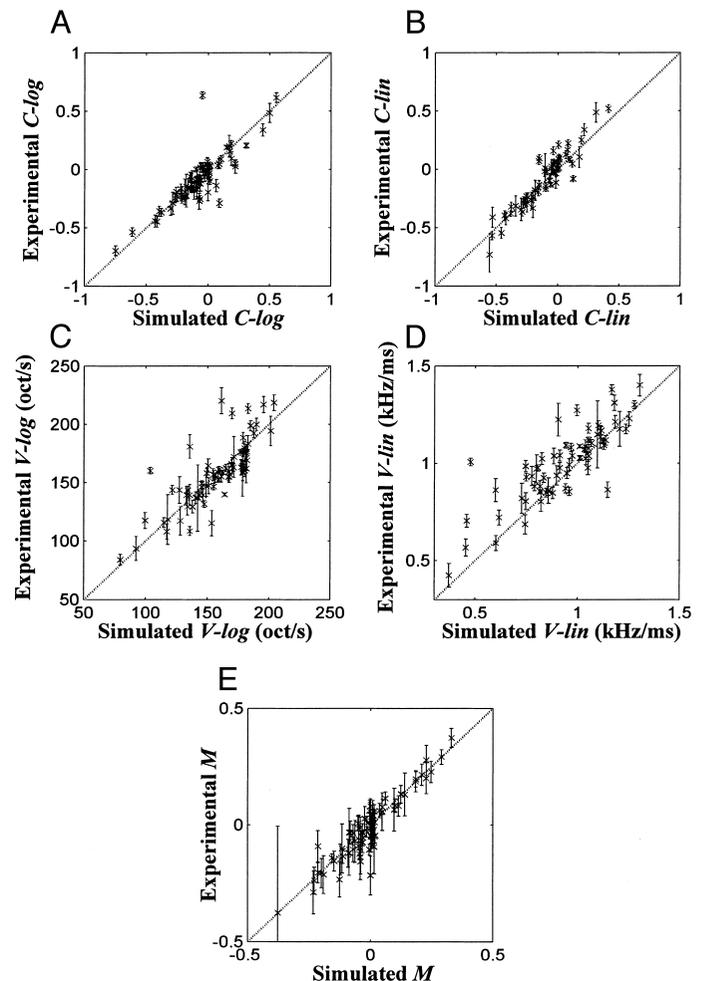
FIG. 6. Simulated vs. experimental response indices for set of 74 neurons matched by model. SD for experimental response indices (estimated by bootstrap) are plotted as vertical error bars on each data point.
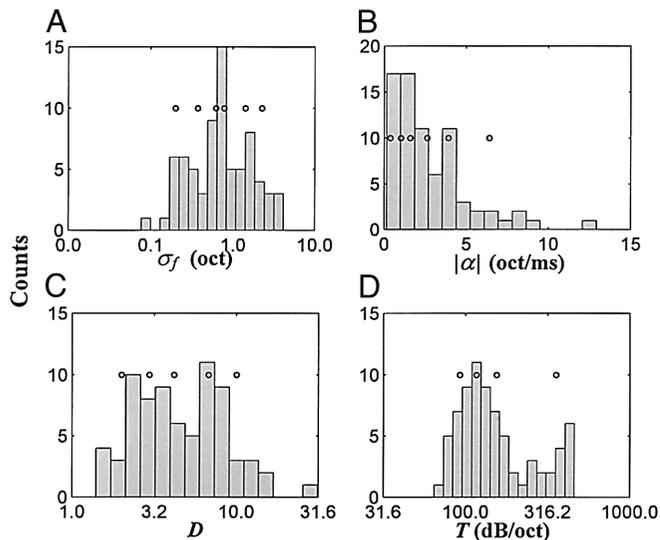
FIG. 7. *Distribution histograms* of 4 model parameters adjusted to fit experimental set of 74 neurons. Note that σ_f (A), D (C), and T (D) are plotted on a log-scale. Circles in each histogram mark parameter values chosen to form a canonical set of 720 model neurons used to investigate properties of model. These values divide each distribution to sets of equal size.

Correlation between response indices: experimental versus simulated data

Shamma et al. (1993) and Kowalski et al. (1995) recorded AI units in barbiturate-anesthetized ferrets and showed a weak but significant positive correlation between the *M* and the *C* response indices using 2-tone complexes and FM sweeps following the LH-log paradigm as defined by Nelken and Versnel (2000) (Fig. 8, *A* and *B*, respectively). Nelken and Versnel (2000) also reported a significant correlation between the *M* index and the *C* index measured using the LH-log paradigm, but this correlation was absent for *C* indices computed using the other FM paradigms. For the subset of experimental data used here the correlation between *M* and *C-log* is significant ($r_{72} = 0.35$, $P < 0.0023$), but the correlation was not significant for *C-lin*, reproducing the same pattern as that of the full

data. The simulated response indices show a small but significant positive correlation for *C-log* ($r_{72} = 0.39$, $P < 0.0006$) and a weaker, although still significant correlation, for *C-lin* ($r_{63} = 0.28$, $P < 0.05$).

Kowalski et al. (1995) recorded from the ferret AI and anterior field (AAF) and tested the relations between the unit excitatory bandwidth at a level of 20 dB above threshold (*BW20*) and the velocity sensitivity of the unit to logarithmic FM sweeps. Kowalski et al. (1995) found a significant correlation between these indices at the AAF but not at the AI (Fig. 8, *D* and *E*). The scatter plots relating BW20 and V-log had a wedge shape, so that in both cases *BW20* had a limiting effect on *V-log*: at higher values of *BW20* the dynamic range of *V-log* is smaller than at lower values of *BW20*. A similar phenomenon was reported by Mendelson et al. (1993) in cat AI and by Heil et al. (1992a) in Field L of chicks, and was found in the subset of experimental data used here as well as in the simulated data.

We tested whether these relationships between the response indices can be attributed to the model, or whether these correlations are attributable to some other mechanism, and the model reproduces these phenomena by merely adjusting its parameters to best match the experimental results. For this purpose we selected representative values of the 4 adjustable parameters of the model (6 values for σ_f, 6 for α, 5 for *D*, and 4 for *T*) and constructed a canonical set of 720 model neurons by using all the possible combinations of these values. The representative values of each parameter were chosen to divide the parameter distribution (as a result of the fitting procedure of the 74 multiunit clusters) to sets of equal size (see Fig. 7). Each model neuron in the canonical set was presented with the entire set of stimuli and the response indices were calculated as described in METHODS. This method enables us to test the properties of the model while minimizing the influence of the particular choice of parameters and maintaining a reasonable distribution of the parameters. Figure 8, *C* and *F* show the scatter plot of the *M* versus *C-log* and of the *BW20* versus *V-log* response indices, respectively, for the canonical set. Clearly, the model reproduces the experimental phenomena
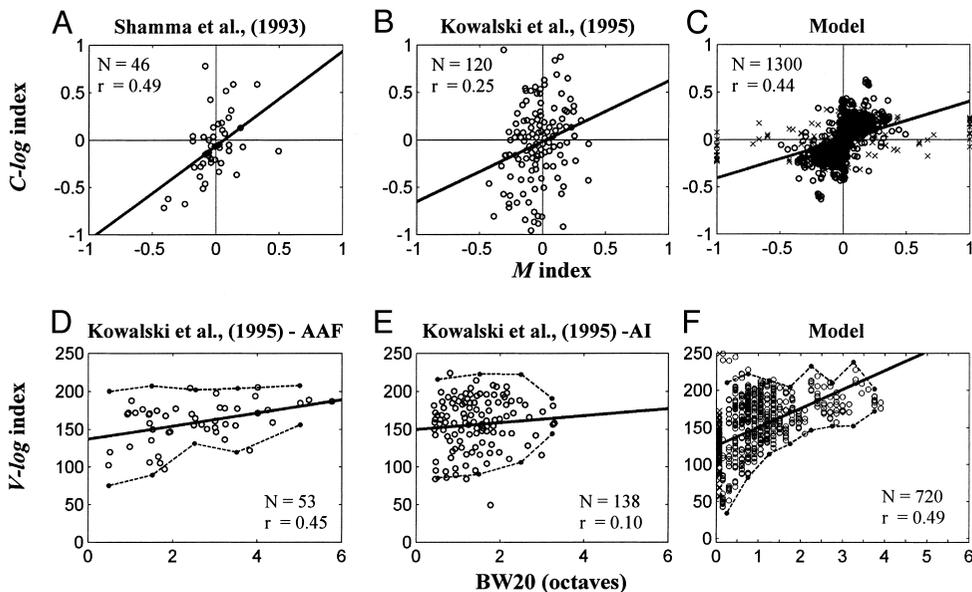
FIG. 8. Correlations between experimental response indices are reproduced by intrinsic properties of model without a particular choice of parameter. *M* vs. *C-log* as replotted from Shamma et al. (1993) (*A*), Kowalski et al. (1995) (*B*), and for simulated response indices (*C*). Regression line and correlations are plotted in the subplots. *D* and *E* replot experimental *BW20* vs. *V-log* as measured from ferret anterior field (AAF) (*D*) and AI (*E*) by Kowalski et al. (1995). Simulated *BW20* vs. *V-log* are plotted in *F*. Regression line as well as the V-log scatter around it (measured as 2 SDs of *V-log* index within each *BW20* octave) are plotted within the subplots. *C* and *F*: parameter combinations that yield nonphysiological responses are marked with crosses. Correlations between experimental indices are significant even when nonphysiological parameter combinations are excluded ($r_{1204} = 0.66$ and $r_{602} = 0.43$ for C and F, respectively).

regardless of the specific choice of parameters. Note that 118 out of the 720 parameter combinations in the canonical set yielded model neurons that had extremely narrow excitatory bandwidth and extreme asymmetry values (marked as crosses in Fig. 8, *C* and *F*). The significant correlations between the response indices in the canonical set was preserved when these nonphysiologically parameter combinations were excluded from the analysis.

The canonical set reproduced even finer features of the correlation structure of the experimental data. For example, when the canonical set was divided into 2 groups of equal size according to their velocity sensitivity *V-log*, the correlation between the *M* and *C-log* indices was higher for the group with the higher *V-log* values ($r_{592} = 0.3$ and $r_{704} = 0.71$, respectively). The same phenomenon held for the experimental response indices; the correlation between *M* and *C-log* for the low *V-log* group was relatively low ($r_{35} = 0.27$) and nonsignificant, whereas for the high *V-log* group the correlation was higher and significant ($r_{35} = 0.47$, $P < 0.0018$). It can be concluded that the model parameters form a correlation-free representation of the experimentally measured representation.

Topographical organization: response indices versus model parameters

One of the main results of Nelken and Versnel (2000) is that the functional maps of the *C* index and to a smaller extent the *V* index are sensitive to the fine details of the FM stimulus, and that in many cases these maps do not show significant clustering. Because each multiunit cluster can be described not only by the experimental *C*, *V*, and *M* values, but also by the values of the 4 model parameters that best approximate it, we could plot topographical maps of the model parameters and compare them with the experimentally measured ones. Figures 9 and 10 show functional maps that are derived from the 2 experiments (*162* and *166*, respectively) for which the largest number of multiunit clusters were fitted by the model. Figure 9 compares the functional maps of the pseudo-rotation parameter α with the functional maps for the 3 experimental indices that are most correlated with it, *M*, *C-log*, and *C-lin*, as derived from *experiment 162*.

As already noted by Nelken and Versnel (2000), the maps for the directional sensitivity are sensitive to the stimulation paradigm. As a result, the *C-log* and *C-lin* functional maps are different from each other. Furthermore, the functional maps derived from these experimentally measured parameters did not show smooth topographical organization, as confirmed by the statistical clustering analysis described in METHODS. These results confirm the conclusions of Nelken and Versnel (2000). In contrast, the functional map derived from the rotation parameter of the model as fitted to each cluster shows a significant clustering ($P < 0.05$) along the cortical surface.

Figure 10 shows similar results for *experiment 166*. The functional maps derived from the *C-log* and *C-lin* indices
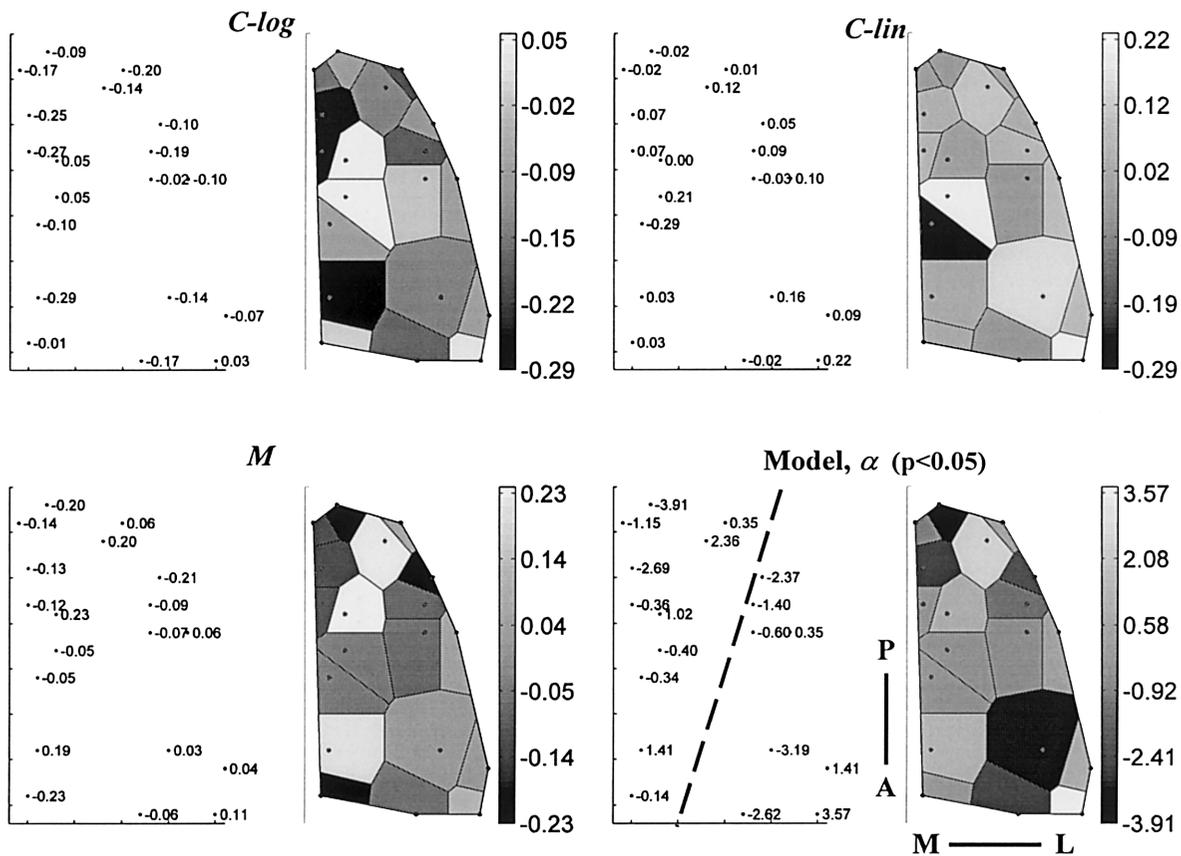
FIG. 9. Functional maps for experimental *C* and *M* response indices for *experiment 162*, and for pseudo-rotation parameter of model α, used to fit these neurons. Dashed line marks isofrequency axis. Each map is given as a raw map and as a gray-level map using a Voronoi tessellation of cortical surface around penetration points. Intertick interval in raw maps corresponds to a distance of 185 μm on cortical surface. Significance level of topographical clustering is given for maps that are significantly clustered.
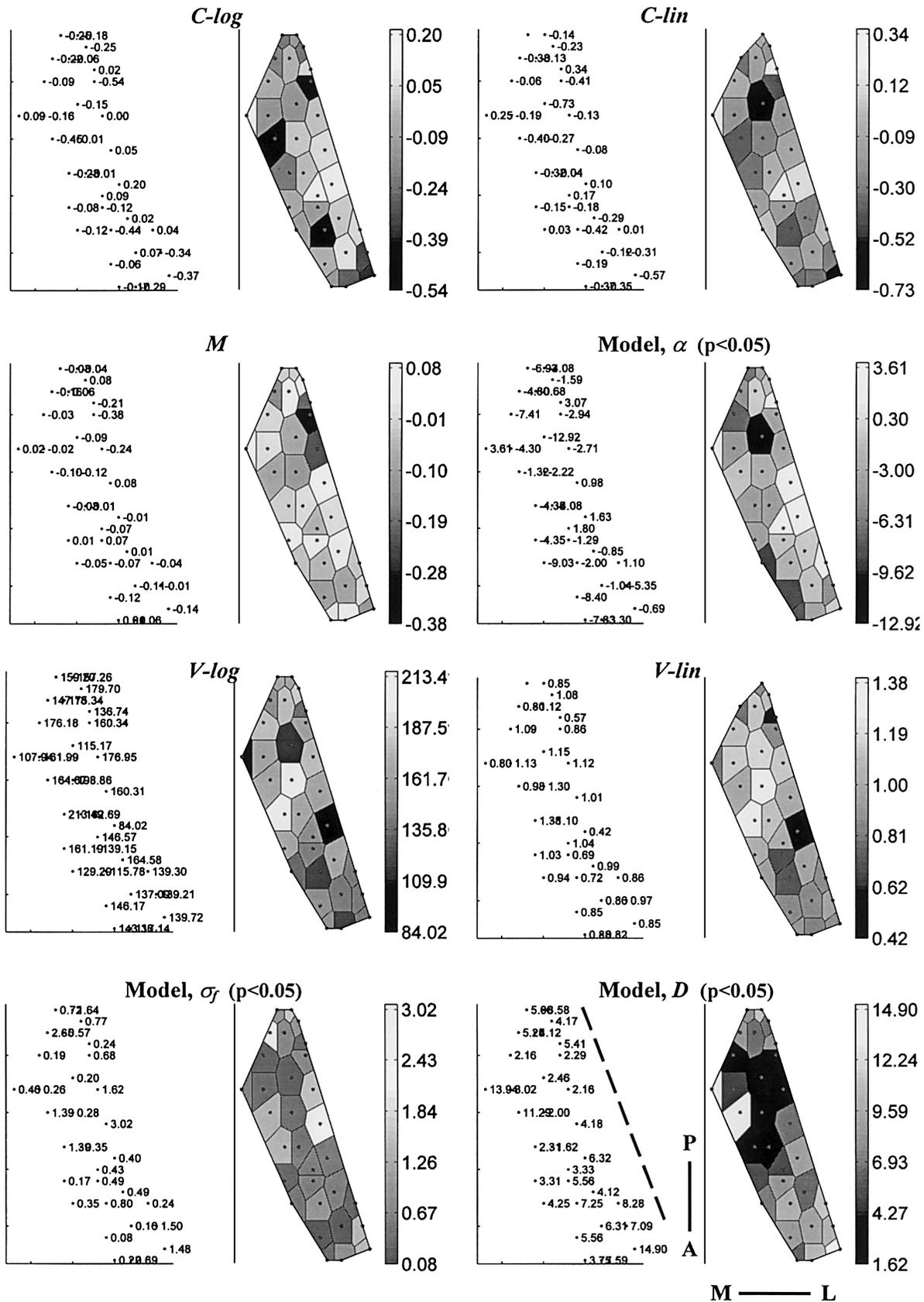
FIG. 10. Functional maps for experimental *C*, *V*, and *M* response indices for *experiment 166*, and for α, σ_f, and *D* parameters of model, used to fit these neurons. Same format as in Fig. 9. Intertick interval corresponds to a distance of 127 μm on cortical surface.
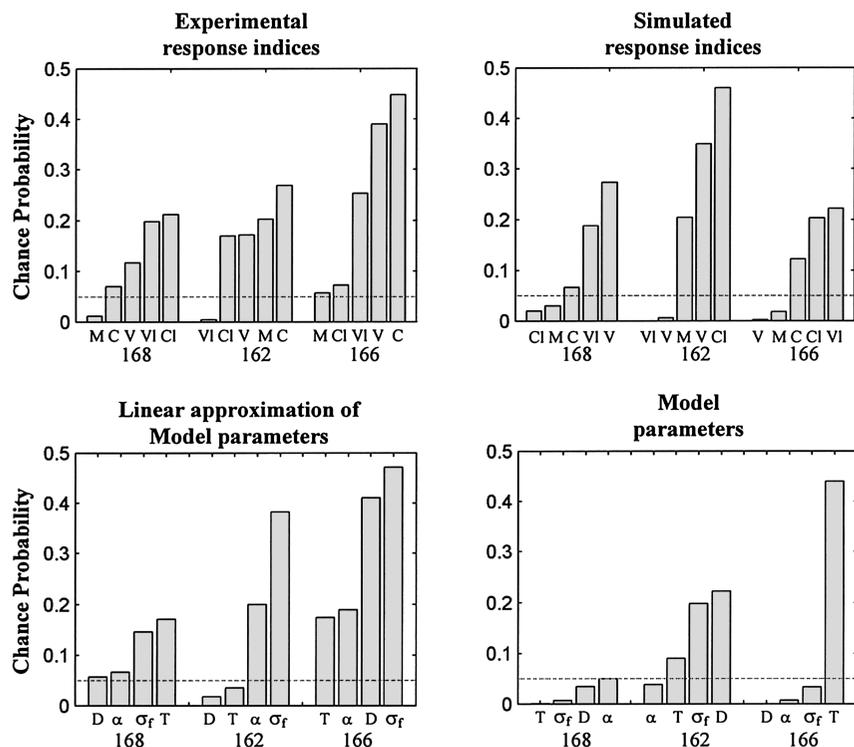
FIG. 11. Probability of a map to be clustered by chance (as calculated by bootstrap) for experimental response indices, simulated response indices, linearly estimated parameters of model, and for model parameters for 3 experiments that were analyzed. Significance level of 0.05 is marked by dashed line.

differ, although sharing some common features (low C values at the posterior part and at the anterior-lateral edge of the mapped region). In contrast, the functional maps derived from the V-log and V-lin indices show much greater similarity, a phenomenon previously described by Nelken and Versnel (2000). None of the functional maps derived from the M, C, and V indices was significantly clustered at a significance level of 0.05; however, the C-lin and M maps showed borderline clustering ($P < 0.075$). As in *experiment 162*, the functional maps derived from the model parameters fitted to each cluster are much more organized. In fact, 3 (α, σ_f, and D) out of the 4 parameters of the model are significantly clustered on the cortical surface.

Figure 11 plots the probability of a map to be clustered by chance for the experimental and simulated response indices (derived for each location from the model fitted to the responses at that location) as well as for the model parameters and the linearly approximated model parameters (the latter are described below). The P values, calculated using the procedure described in METHODS, are also presented in Table 1. Figure 11 and Table 1 demonstrate that the number of significantly clustered maps for the model parameters outnumbers the significantly clustered maps induced by any of the other indices or parameters. In total, only 2 out of 15 maps computed for the 5 experimental response indices (C-log, C-lin, V-log, V-lin, and M) in the 3 experiments are significantly clustered (V-lin of *experiment 162*, and M of *experiment 168*). For the simulated response indices, 5 of the 15 maps are significantly clustered. Note that the simulated indices are already "smoothed" in the sense that all the data recorded from each cluster are considered in their calculation. The highest number of clustered maps, however, was found among the maps derived from model parameters: 8 out of the 12 maps computed for the 4 model parameters are significantly clustered.

The model parameters are the result of a nonlinear transformation from the response space to the model parameter space, which, as described in the previous section, decorrelates the experimental indices. The lack of analytical expression for this transformation impairs the ability to track down the origin of the improved topographical organization of the model parameters. In principle, there are 3 explanations as to why the model parameters show more clustering than the experimentally derived indices: across-cluster smoothing; within-cluster smoothing; or some real physiological process that is captured by the model and is not captured well by any of the experimentally derived indices.

The improved topographical organization of the model parameters with respect to the experimental indices does not arise from smoothing of the parameters across neighboring multiunit clusters. This is attributed the fact that the parameters of the model were separately fitted to the experimental data of each

TABLE 1. *Topographically clustered indices and model parameters*

	162 (*n* = 19)	166 (*n* = 38)	168 (*n* = 17)
Significantly clustered experimental indices	*V*-lin		*M*
Significantly clustered simulated index	*V*-lin	*M* *V*-log	*C*-lin *M*
Significantly clustered approximated parameters	*D* *T*		
Significantly clustered model parameters	α	*D* α σ_f	*T* σ_f *D* α
Model mean (SD) fit error	28.1 (10.9)	20.0 (13.5)	18.5 (7.7)

multiunit cluster, without any information regarding the topographical location of the cluster or its neighboring clusters.

The second explanation for the improved topographical organization of the model parameters is that the experimentally derived indices are in fact topographically organized, but noisy, resulting in an apparent loss of order. According to this explanation, the fitting procedure smooths the experimental response indices *within* each cluster (e.g., by averaging the *C* and *V* indices across the logarithmic and linear paradigms), causing the simulated response indices to become more topographically organized than the original response indices. Although the simulated response indices generally showed a more ordered topographical distribution than the experimental indices, this explanation is probably incomplete given that the topographical organization of the simulated response indices was worse than that of the model parameters (Fig. 11 and Table 1). To study this question further, we approximated the model parameters (fitted using the nonlinear procedure described above) using a linear regression on the measured response indices (*M*, *C-log*, *C-lin*, *V-log*, *V-lin*), resulting in a global linear approximation to the nonlinear fitting procedure. Within-cluster smoothing would predict that the resulting "linearized" model parameters would also show significant clustering. The model parameters, fitted by the nonlinear procedure, were only moderately correlated with the linearized model parameters (α, $r^2 = 0.484$; σ_f, $r^2 = 0.11$; D, $r^2 = 0.357$; T, $r^2 = 0.129$). Most important, the linearized model parameters failed to produce significantly clustered maps (Fig. 11 and Table 1).

Table 1 also shows that the experiment for which only one model parameter is significantly clustered (*experiment 162*) is the experiment for which the model fits the experimental data less well than for the other 2 experiments. The fit error for the multiunit clusters of *experiment 162* are significantly higher than those of *experiment 166* ($t_{55} = 2.27$, $P < 0.015$) and from those of *experiment 168* ($t_{34} = 3.05$, $P < 0.0022$). The fit errors for the multiunit clusters of *experiments 166* and *168* do not differ significantly. Notably, *experiment 162* is the only experiment out of the 3 reported here that was not tested with the N-log and N-lin paradigms by Nelken and Versnel (2000). Therefore for all the multiunit clusters of *experiment 162* the model responses were matched with the neural responses for the upward portion of the LH stimulus and with the responses to the downward portion of the HL stimulus, as described in METHODS. The only difference between the N-log/N-lin paradigms and the first half of the LH/HL paradigms is the duration of the fixed tone at the beginning of the stimulus. For the N-log/N-lin this duration is fixed to 10 ms, whereas for the LH-lin/HL-lin it is set to 75 ms and for the LH-log/HL-log the duration varies in the range of 10–90 ms in accordance with the sweep velocity. As noted by Nelken and Versnel (2000), this difference by itself already gave rise to a measurable amount of adaptation in the neural responses to the following frequency sweep.

It is possible that the temporal proximity between the beginning of the stimulus and the start of the FM sweep makes the N-log/N-lin paradigms more suitable for the model fitting, in that the model lacks components that can account for adaptation effects. Therefore it is possible that the lack of N-log/N-lin data for *experiment 162* accounts for the modest ability of the model to fit the responses and to show significant parameter clustering for that experiment.

DISCUSSION

We present in the current study a neural model for the detection of frequency and amplitude transitions. The basic computational operation of the model is a delayed integration of the amplitude time-derivative information across neighbor-frequency channels. This computation is carried out by applying a rotatable 2D receptive field to a spectrotemporal representation of the acoustic information. The spectrotemporal representation is spanned by an array of neurons that differ in their best frequencies and temporal delays. We show that the model is able to reproduce the responses of AI neurons to single tones, 2-tone complexes, narrow-excursion FM transitions, and linear and logarithmic FM sweeps, as recorded from the barbiturate-anesthetized cat and ferret.

Although the model shows a good agreement with the data analyzed here, there are some aspects of the responses of cortical neurons that the model does not account for. Specifically, AI neurons have longer time constants, as expressed for example in the relatively slow dynamics of STRFs measured using reverse correlation or moving ripple spectra (10–20 Hz; Depireux et al. 2001; Linden et al. 2003; Miller et al. 2002; Schnupp et al. 2001), or as expressed in the dependency of AI responses on the stimulus history at a time scale of about ± 1 s (Brosch and Schreiner 1997; Calford and Semple 1995; Heil et al. 1992a; Malone et al. 2002; Nelken and Versnel 2000; Ulanovsky et al. 2003). These phenomena cannot be reproduced in the model because the model integrates acoustic input over a time window of ≤ 10 ms. Furthermore, to keep the model as simple as possible the model components are I&F neurons that lack mechanisms such as synaptic depression and facilitation. As a result, the responses of the model neurons depend on a relatively short stimulus history of about 10 ms. The success of the model as it stands suggest that multiple time constants coexist in cortical neurons. Whether including mechanisms with longer time constants, such as synaptic depression, to the model will adequately address this problem remains an open question (Ulanovsky et al. 2003).

Part of the motivation for our study stemmed from the assumption that transient information may be used by the auditory system as important cues for the task of primary segmentation of complex auditory scenes (Fishbach et al. 2001). Some psychoacoustical studies demonstrate that auditory perception is sensitive to the gradient of amplitude transitions and that a larger gradient enables easier separation of auditory components (Bregman et al. 1994; Turner et al. 1994; also see reanalysis of these studies in Fishbach et al. 2001). We presented in our previous study a neural model for the detection of auditory temporal transitions (edges) and demonstrated that the model is able to reproduce and predict various physiological and psychoacoustical responses to amplitude transitions.

The application of a rotatable receptive field to a 2D spectrotemporal auditory space, as presented in this study, is a natural generalization of the previously suggested temporal model. The 2D model studied here preserves the temporal properties of the 1D model and therefore it suggests a common unifying mechanism for a vast physiological and some psychoacoustical responses to acoustic transitions.

Although the model is abstract, all of its components are physiologically plausible. Even though the existence of a neu-

ral frequency-delay layer is not well substantiated, there is evidence that validates the use of such an auditory mechanism. Hattori and Suga (1997) report that the latency (ranging from 4 to 12 ms) of neurons in the IC of unanesthetized mustached bats is topographically organized orthogonally to the tonotopic organization of the IC, forming a frequency versus latency map. A similar organization of onset latencies in the cat IC was reported by Schreiner and Langner (1988b). Both the range of values and the topographic organization of the latency in the bat and in the cat IC are consistent with the model's frequency-delay layer. However, more research is needed to establish a direct link between these findings and the proposed model.

Relationships to STRF models

During recent years, a number of methods have been proposed for constructing spectrotemporal receptive fields of auditory neurons, using synthetic (e.g., deCharms et al. 1998; Depireux et al. 2001; Miller et al. 2002) and natural (Sen et al. 2001; Theunissen et al. 2000) sounds. The STRF obtained using these techniques is the best linear model that approximates the responses of a neuron to the stimulus ensemble used for the estimation process. It is important to note that the STRF estimated using these techniques is very different from the STRF used in our model. The STRF in our model is being applied to the model's spectrotemporal representation of the acoustic stimulus. This representation is spanned by the frequency-delay neurons that have adjustable bandwidth and their output undergoes 2 compressive nonlinearities (*Eqs. 2* and *3*). Because of these nonlinearities, an STRF of a synthetic neuron simulated by the model, measured using the above-mentioned techniques, will not necessarily coincide with the model's STRF as represented by the weights of the frequency-delay layer (Fig. 1*D* and *Eq. 4*). Moreover, the STRFs used in our study constitute a restricted subset of all possible STRFs, and can be characterized by only 2 free parameters (*Eq. 4*). Although this probably constrains the ability of the model to fully

describe the properties of a neuron, it allows a more robust characterization of part of its spectrotemporal properties based on a limited set of stimuli.

These differences between the model's STRF and a linearly estimated STRF have an important implication to the ability of the model to replicate experimental data. For example, the model is able to demonstrate sensitivity to very fast frequency transitions ($>$150 oct/s), which is common to many AI neurons (Kowalski et al. 1995; Nelken and Versnel 2000; and see Fig. 6*C*). The ability of the model to respond optimally to frequency transitions of such a high velocity is partly attributed to the relatively short time constants (3–9.8 ms; see the APPENDIX) that are used to span the frequency-delay layer; however, the fast dynamics of the model are further boosted by the nonlinearities the delay layer undergoes. This property of the model is illustrated in Fig. 12*A*, which plots the average velocity sensitivity of the model (*V-log*) as a function of the filter bandwidth (*T*) and saturation scaling (*D*) parameters of the model's delay layer. For this plot we use the data of the 720 model neurons in the canonical set we have constructed using representative values for each of the 4 model parameters (see Figs. 7 and 8). Therefore each data point in Fig. 12*A* is the averaged *V-log* of the same 36 configurations of the model's STRF (6 values for α by 6 values for σ_f). Despite the fact that the STRF configurations at each data point are the same, the model's velocity sensitivity changes systematically as a function of the delay-layer filter bandwidth (*T*) and saturation scaling (*D*) parameters. Note that the average velocity sensitivity increases as the saturation transformation becomes steeper (smaller values of *D*), which implies that in the absence of this nonlinear operator the average velocity sensitivity of the model would be lower.

Typically, the STRFs estimated for AI neurons show much more sluggish characteristics than what is needed to produce sensitivity to high-velocity transitions (e.g., deCharms et al.
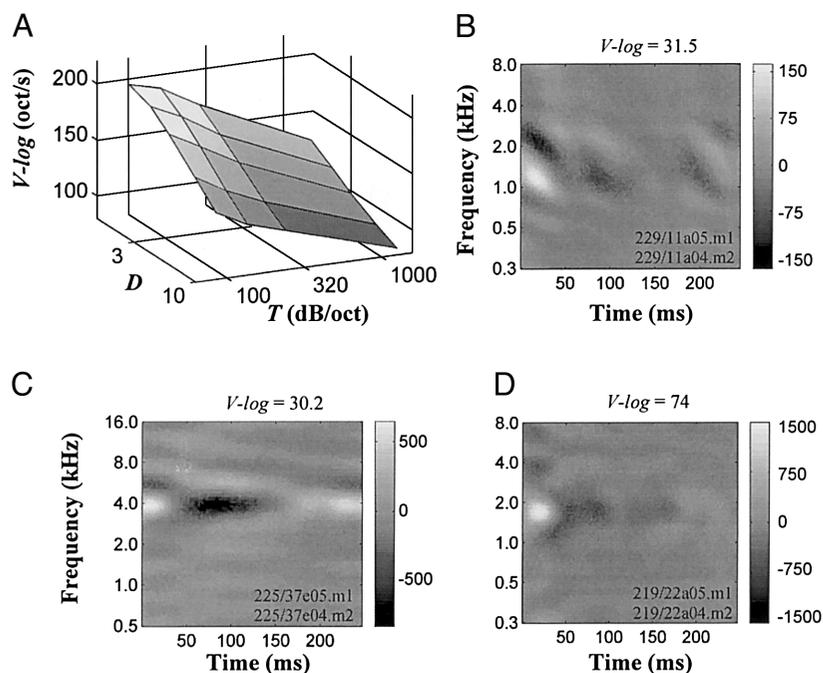
FIG. 12. *A*: average velocity sensitivity of model (*V-log*) as a function of filter bandwidth (*T*) and saturation scaling (*D*) parameters of model's delay layer. Data were calculated from canonical set of 720 artificial neurons we constructed using representative values for each of 4 model parameters (see Fig. 7). *B–D*: replot of 3 STRFs estimated by Depireux et al. (2001) from ferret AI. Estimated velocity sensitivity (*V-log*) as estimated by simulating responses of each STRF to stimuli used by Nelken and Versnel (2000) appears above each plot. These values are much lower than typical values estimated for AI neurons. See DISCUSSION for more details.

1998; Depireux et al. 2001; Miller et al. 2001). To verify this qualitative observation, we used 3 of the STRFs displayed in Depireux et al. (2001) and calculated their response to the stimuli used by Nelken and Versnel (2000). The STRFs and their corresponding velocity sensitivity (*V-log*) are shown in Fig. 12, *B–D*. The velocity sensitivity for these STRFs is indeed much lower than the typical values estimated for AI neurons. A similar observation was made by Bar-Yosef et al. (2002), who showed that AI neurons in the cat are sensitive to fluctuations in stimulus envelopes on a time scale that is much faster than predicted by most STRF studies.

The model does more than fitting the data

The fact that the model is able to fit the responses of AI neurons to a variety of stimuli suggests that the model may capture some fundamental aspect of the computation that is being performed by neurons in the primary auditory cortex. Moreover, we show that a correlation-free combination of the model parameters reproduces the correlations between the experimental response indices, as reported in several studies. Thus these correlations are being reproduced by the intrinsic properties of the model and not by a particular choice of parameters. These findings support the claim that the model parameters better characterize the recorded neurons than the response indices that are extracted from the responses to a single type of stimulus. In this respect, the model performs a nonlinear transformation of the neural characteristics from a nonorthogonal experimentally measured coordinate system to a much more orthogonal coordinate system of the underlying computation.

The concept of a model-based characterization of the recorded neurons is further bolstered by the clustered topographical organization of the model parameters along the cortical surface. In 2 out of the 3 experiments we have analyzed, 7 out of the 8 possible maps for the 4 model parameters show a significant topographical clustering. In the other experiment (*162*), only the pseudo-rotation parameter was found to be significantly clustered. Interestingly, the quality of the model fit to the responses recorded from that experiment was significantly lower than that for the other 2 experiments. The relation between the quality of the fit and the quality of the topographical organization of the model parameters is not trivial, given that the experimental response indices showed equally weak topographical clustering in all 3 experiments. Also, as we already noted, the model parameters were adjusted separately to fit the responses of each multiunit cluster, without any information regarding its topographical location, nor the properties of its neighbors.

Primary sensory areas in the cortex of all mammals investigated to date demonstrated topographically ordered organization (Kaas 1997). Across the cortical surface, neurons that share a topographical proximity tend to have similar values of some functional properties. In many cases the sensory cortical maps relate to some dimension of the receptor organ. However, there are examples for cortical maps of computational properties (e.g., Bonhoeffer and Grinvald 1991; Kaas et al. 1979). This in mind, the fact that the model parameters reveal topographical organization that is substantially better than the topographical organization of the indices extracted from the responses the model fits strongly

support the claim that the model parameters more fully and more simply characterize the recorded neurons than the experimental response indices. Thus we suggest that the model reflects some essential aspects of the computation performed by neurons in primary auditory cortex.

APPENDIX: MATHEMATICAL EQUATIONS AND PARAMETERS OF THE MODEL

Neural representation

The neural representation consists of a bank of 81 band-pass filters. The output of the *i*th filter is given by

$$N_i(t) = \frac{1}{\tau_m^2} \int_0^t \left[\sum_{k=1}^{n} E_k(x) - T \left| \log_2\left(\frac{F_k(x)}{f_i}\right) \right| \right](t - x)e^{-(t-x)/\tau_m} \, dx$$

$$f_i = MF \times 2^{0.05(i-40)} \quad i = 1, 2, \ldots, 81 \quad (A1)$$

where E_k and F_k are the instantaneous amplitude level (in dB SPL) and the frequency of each of the *n* input sinusoidals. *MF* is the middle frequency of the model, τ_m is set to 1 ms, and *T* is one of the 4 adjustable parameters of the model.

Delay layer

The operation of each unit $U_{i,j}(t)$ of the delay layer on its input $N_i(t)$ is given by

$$U_{i,j}(t) = \frac{1}{\tau_j^2} \int_0^t N_i(x)(t - x)e^{-(t-x)/\tau_j} \, dx \quad (A2)$$

In our simulations we used 69 τ_j values equally spaced between 3 and 9.8 ms. The output of the units is saturated using the following sigmoidal transformation

$$\tilde{U}_{i,j}(t) = F_{max}\left\{ \frac{2}{[1 + e^{-U_{i,j}(t)/D}] - 1} \right\} \quad (A3)$$

where F_{max} is set to 225 spines/s and *D* is one of the adjustable parameters of the model.

Receptive field

The delay layer units are connected to a single neuron. The neuron's input $I(t)$ is given by

$$I(t) = \sum_{i=1}^{81} \sum_{j=1}^{69} R(i, j)\tilde{U}_{i,j}(t) \quad (A4)$$

where

$$R(i, j) = R_f(f_i)R_\tau(\tau_j, f_i)$$

$$R_f(f_i) = \frac{1}{\sigma_f} e^{-(F/\sigma_f)^2} \quad \text{where} \quad F = \log_2(f_i/MF)$$

$$R_\tau(\tau_j, f_i) = \frac{2}{\sigma_\tau^3} te^{-(t/\sigma_\tau)^2} \quad \text{where} \quad t = \tau_j - (\mu + \alpha F) \quad (A5)$$

and where μ_τ is the mean delay of the receptive field, set to 6.4 ms; σ_τ is the width of the delay time window, set to 0.2 ms; α, the pseudo-rotation parameter, and σ_f, the frequency integration bandwidth, are 2 of the 4 adjustable parameters of the model.

Edge-detection neuron

The neuron is modeled as a simple leaky integrator with a voltage threshold (T), an absolute refractory period ($\delta^{abs} = 1$ ms), and a refractoriness function

$$\varphi(t) = -Te^{-(t-\delta^{abs})/\lambda}S(t - \delta^{abs}) - KS(t)S(\delta^{abs} - t) \quad (A6)$$

with a constant $K \to \infty$, $\lambda = 1.5$ ms, and where $S(t)$ is the positive step function (Gerstner 1999a).

The membrane potential $M(t)$ of the neuron is given by

$$M(t) = \frac{1}{\tau_{ed}^2}\int_0^t [I(x) + \xi(x)](t - x)e^{-(t-x)/\tau_{ed}}dx + \sum_{s_i \in \{s_1, \ldots, s_n\}} \varphi(t - s_i) \quad (A7)$$

where $\{s_1, \ldots, s_n\}$ represents the set of firing times of the neuron, the membrane time constant τ_{ed} is set to 3 ms, and $\xi(x)$ is a random Gaussian noise with a zero mean and SD $\sigma = 0.2T$.

DISCLOSURES

REFERENCES

Agmon-Snir H and Segev I. Signal delay and input synchronization in passive dendritic structure. *J Neurophysiol* 70: 2066–2085, 1993.

Barth CD and Burkard R. Effects of noise bursts rise time and level on the human brainstem auditory evoked response. *Audiology* 32: 225–233, 1993.

Bar-Yosef O, Rotman Y, and Nelken I. Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. *Neuroscience* 22: 8619–8632, 2002.

Bonhoeffer T and Grinvald A. Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature* 353: 429–431, 1991.

Bregman AS, Ahad PA, Kim J, and Melnerich L. Resetting the pitch-analysis system. 1. Effects of rise times of tones in noise backgrounds or of harmonics in a complex tone. *Percept Psychophys* 56: 155–162, 1994.

Brosch M and Schreiner CE. Time course of forward masking tuning curves in cat primary auditory cortex. *J Neurophysiol* 77: 923–943, 1997.

Calford MB and Semple MN. Monaural inhibition in cat auditory cortex. *J Neurophysiol* 73: 1876–1891, 1995.

deCharms RC, Blake DT, and Merzenich MM. Optimizing sound features for cortical neurons. *Science* 280: 1439–1443, 1998.

Depireux DA, Simon JZ, Klein DJ, and Shamma SA. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol* 85: 1220–1234, 2001.

Eggermont JJ. Differential effects of age on click-rate and amplitude-modulation-frequency coding in cat primary auditory cortex. *Hear Res* 65: 175–192, 1993.

Fishbach A, Nelken I, and Yeshurun Y. Auditory edge detection: a neural model for physiological and psychoacoustical responses to amplitude transients. *J Neurophysiol* 85: 2303–2323, 2001.

Gerstner W. Spiking neurons. In: *Pulsed Neural Networks*, edited by Maass W and Bishop CM. Cambridge, MA: MIT Press, 1999a.

Gerstner W. Population of spiking neurons. In: *Pulsed Neural Networks*, edited by Maass W and Bishop CM. Cambridge, MA: MIT Press, 1999b.

Hattori T and Suga N. The inferior colliculus of the mustached bat has the frequency-vs-latency coordinates. *J Comp Physiol A Sens Neural Behav Physiol* 180: 271–284, 1997.

Heil P. Auditory onset responses revisited. I. First-spike timing. *J Neurophysiol* 77: 2616–2641, 1997a.

Heil P. Auditory onset responses revisited. II. Response strength. *J Neurophysiol* 77: 2642–2660, 1997b.

Heil P and Irvine DRF. On determinants of first-spike latency in auditory cortex. *Neuroreport* 7: 3073–3076, 1996.

Heil P and Irvine DRF. First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J Neurophysiol* 78: 2438–2454, 1997.

Heil P and Irvine DRF. Functional specialization in auditory cortex: response to frequency-modulated stimuli in the cat's posterior auditory field. *J Neurophysiol* 79: 3041–3059, 1998a.

Heil P and Irvine DRF. The posterior field P of cat auditory cortex: coding of envelope transients. *Cereb Cortex* 8: 125–141, 1998b.

Heil P, Langner G, and Scheich H. Processing of frequency-modulated stimuli in the chick auditory cortex analogue: evidence for topographic representations and possible mechanisms of rate and directional sensitivity. *J Comp Physiol A Sens Neural Behav Physiol* 171: 583–600, 1992a.

Heil P, Rajan R, and Irvine DRF. Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. I. Effects of variation of stimulus parameters. *Hear Res* 63: 108–134, 1992b.

Heil P, Rajan R, and Irvine DRF. Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. II. Organization of response properties along the "isofrequency" dimension. *Hear Res* 63: 136–156, 1992c.

Kaas JH. Topographic maps are fundamental to sensory processing. *Brain Res Bull* 44: 107–112, 1997.

Kaas JH, Nelson RJ, Sur M, Lin CS, and Merzenich MM. Multiple representations of the body within the primary somatosensory cortex of primate. *Science* 204: 521–523, 1979.

Kitzes LM, Gibson MM, Rose JE, and Hind JE. Initial discharge latency and threshold considerations for some neurons in cochlear nucleus complex of the cat. *J Neurophysiol* 41: 1165–1182, 1978.

Kowalski N, Versnel H, and Shamma AS. Comparison of responses in the anterior and primary auditory fields of the ferret cortex. *J Neurophysiol* 73: 1513–1523, 1995.

Linden JF, Liu RC, Sahani M, Schreiner CE, and Merzenich MM. Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J Neurophysiol* 90: 2660–2675, 2003.

Malone BJ, Scott BH, and Semple MN. Context-dependent adaptive coding of interaural phase disparity in the auditory cortex of awake macaques. *Neuroscience* 22: 4625–4638, 2002.

Mendelson JR, Schreiner CE, Sutter ML, and Grasse KL. Functional topography of cat primary auditory cortex: responses to frequency modulated sweeps. *Exp Brain Res* 94: 65–87, 1993.

Miller LM, Escabi MA, Read HL, and Schreiner CE. Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. *J Neurophysiol* 87: 516–527, 2002.

Nelken I and Versnel H. Responses to linear and logarithmic frequency-modulated sweeps in ferret primary auditory cortex. *Eur J Neurosci* 12: 549–562, 2000.

Phillips DP. Effect of tone-pulse rise time on rate-level functions of cat auditory cortex neurons: excitatory and inhibitory processes shaping responses to tone onset. *J Neurophysiol* 59: 1524–1539, 1988.

Phillips DP. Factors shaping the response latencies of neurons in the cat's auditory cortex. *Behav Brain Res* 93: 33–41, 1998.

Phillips DP and Burkard R. Response magnitude and timing of auditory response initiation in the inferior colliculus of the awake chinchilla. *J Acoust Soc Am* 105: 2731–2737, 1999.

Phillips DP, Mendelson JR, Cynader MS, and Douglas RM. Responses of single neurons in cat auditory cortex to time-varying stimuli: frequency-modulated tones of narrow excursion. *Exp Brain Res* 58: 443–454, 1985.

Phillips DP, Semple MN, and Kitzes LM. Factors shaping the tone level sensitivity of single neurons in posterior field of cat auditory cortex. *J Neurophysiol* 73: 674–686, 1995.

Ramachandran R, Davis KA, and May BJ. Single-unit responses in the inferior colliculus of decerebrate cats. I. Classification based on frequency response maps. *J Neurophysiol* 82: 152–163, 1999.

Rees A and Møller AR. Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hear Res* 10: 301–330, 1983.

Rotman Y, Bar-Yosef O, and Nelken I. Relating cluster and population responses to natural sounds and tonal stimuli in cat primary auditory cortex. *Hear Res* 152: 110–127, 2001.

Schnupp JW, Mrsic-Flogel TD, and King AJ. Linear processing of spatial cues in primary auditory cortex. *Nature* 414: 200–204, 2001.

Schreiner CE and Langner G. Central coding of temporal patterns. In: *Auditory Function, Neurobiological Bases of Hearing*, edited by Edelman G, Gall W, and Cowan W. New York: Wiley, 1988a.

Schreiner CE and Langner G. Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J Neurophysiol* 60: 1823–1840, 1988b.

Schreiner CE and Mendelson JR. Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J Neurophysiol* 64: 1442–1459, 1990.

Sen K, Theunissen FE, and Doupe AJ. Feature analysis of natural sounds in the songbird auditory forebrain. *J Neurophysiol* 86: 1445–1458, 2001.

Shamma SA, Fleshman JW, Wiser PR, and Versnel H. Organization response areas in ferret primary auditory cortex. *J Neurophysiol* 69: 367–383, 1993.

Suga N. Responses of inferior collicular neurons of bats to tone bursts with different rise times. *J Physiol* 217: 159–177, 1971.

Theunissen FE, Sen K, and Doupe AJ. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *Neuroscience* 20: 2315–2331, 2000.

Tian B and Rauschecker JP. Processing of frequency-modulated sounds in the cat's anterior auditory field. *J Neurophysiol* 71: 1959–1975, 1994.

Tian B and Rauschecker JP. Processing of frequency-modulated sounds in the cat's posterior auditory field. *J Neurophysiol* 79: 2629–2642, 1998.

Turner CW, Relkin EM, and Doucet J. Psychophysical and physiological forward masking studies: probe duration and rise-time effects. *J Acoust Soc Am* 96: 795–800, 1994.

Ulanovsky N, Las L, and Nelken I. Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6: 391–398, 2003.

Young ED and Voigt HF. Response properties of type II and type III units in dorsal cochlear nucleus. *Hear Res* 6: 153–169, 1982.