Updated parameters and expanded simulation options for a model of the auditory periphery

Muhammad S. A. Zilany[a)]
Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia

Ian C. Bruce
Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, L8S 4K1, Canada

Laurel H. Carney
Departments of Biomedical Engineering and Neurobiology and Anatomy, University of Rochester, Rochester, New York 14642

A phenomenological model of the auditory periphery in cats was previously developed by Zilany and colleagues [J. Acoust. Soc. Am. **126**, 2390–2412 (2009)] to examine the detailed transformation of acoustic signals into the auditory-nerve representation. In this paper, a few issues arising from the responses of the previous version have been addressed. The parameters of the synapse model have been readjusted to better simulate reported physiological discharge rates at saturation for higher characteristic frequencies [Liberman, J. Acoust. Soc. Am. **63**, 442–455 (1978)]. This modification also corrects the responses of higher-characteristic frequency (CF) model fibers to low-frequency tones that were erroneously much higher than the responses of low-CF model fibers in the previous version. In addition, an analytical method has been implemented to compute the mean discharge rate and variance from the model's synapse output that takes into account the effects of absolute refractoriness. © 2014 Acoustical Society of America.
[http://dx.doi.org/10.1121/1.4837815]

I. INTRODUCTION

The auditory periphery is the sole conduit for acoustic information to reach the higher auditory centers. Thus a detailed understanding of the transformation of acoustic signals in the auditory periphery is necessary to determine how changes or deficits in certain underlying mechanisms in the healthy or impaired ear may affect perception. A phenomenological model of auditory-nerve responses in cats has been developed over the years (Carney, 1993; Zhang *et al*., 2001; Bruce *et al*., 2003, Zilany and Bruce, 2006, 2007; Zilany *et al*., 2009) to determine how the underlying sensory and neural mechanisms affect neural representations and more specifically to understand the staged transformations at different levels of the auditory periphery. The most recently published model (Zilany *et al*., 2009), which will be subsequently referred to as the "2009 version," incorporates most of the nonlinearities observed at the level of the auditory nerve (AN), such as nonlinear tuning, level-dependent phase, compression, suppression, shift in the best frequency as a function of level, adaptation, as well as some other nonlinearities seen at high sound levels. The 2009 version also includes power-law dynamics in the model of the inner-hair-cell (IHC)-AN synapse, which substantially improves the prediction of AN responses to a wide variety of complex sounds such as

amplitude-modulated (AM) stimuli, responses to forward-masking paradigms (Zilany *et al*., 2009), and also the long-term dynamics of AN responses (Zilany and Carney, 2010).

Although the 2009 version of the AN model captures features of responses in cat to a wide variety of simple (tone-like) and complex (speech and noise) acoustic stimuli, the model does not accurately simulate the data from Liberman (1978) describing the discharge rate at saturation as a function of characteristic frequency (CF) for higher CFs. The synapse model of the 2009 version has exponential adaptation followed by power-law adaptation. The synapse parameters, particularly the steady-state (Ass) and the spontaneous rate (Asp) of the three-store diffusion model (Westerman and Smith, 1988) were adjusted and set as a function of CF when the power-law adaptation was added to the model. Unfortunately, it has subsequently been realized that the model saturation rates at higher frequencies to CF tones are significantly higher than the rates described by Liberman (1978). Also the responses of higher-CF model fibers to low-frequency tones are erroneously much higher than the responses of low-CF model fibers to the same stimuli. This problem is evident from the response area to a 500-Hz tone signal [Fig. 1(A)].

To address the problem mentioned in the preceding text, some parameters of the IHC-AN synapse section of the model have been modified to limit the saturation rate of fibers as a function of CF to be more consistent with the data reported in Liberman (1978). Also an analytical method has

---

[a)]Author to whom correspondence should be addressed. Electronic mail: msazilany@gmail.com
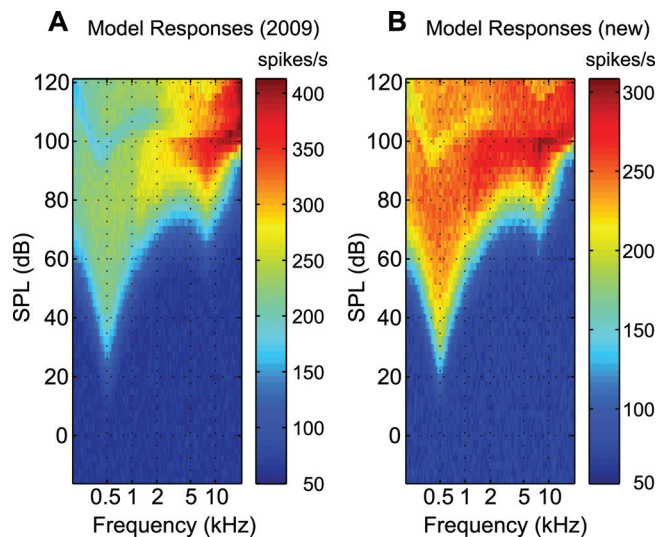
FIG. 1. Responses of a population of AN model fibers to a 500-Hz tone (50-ms duration with 2.5 ms on/off ramp) presented across a range of sound levels. Mean rate responses are shown for 100 AN fibers with CFs logarithmically spaced from 200 Hz to 20 kHz (along the *x* axis) and sound levels ranging from −15 to 120 dB SPL (along the *y* axis) in steps of 2.5 dB. (A) Responses from the Zilany *et al.* (2009). (B) Responses from the model presented here.

II. MODIFICATIONS OF THE MODEL AND RESULTS

A. Saturation rate as a function of CF

In the IHC-AN synapse section of the model (Zilany *et al.*, 2009), a three-store diffusion model is followed by a power-law adaptation. Westerman and Smith's (1988) three-store diffusion model produces an exponential adaptation with rapid and short-term time constants, and the output of this exponential adaptation stage is further adapted by the power-law dynamics. The power-law model has also two paths, namely, slow and fast. In general, the power-law affects (adapts) constant (dc) signals, which arise for the higher CFs because of the low-pass filter in the IHC, more than the respective fluctuating signals (which arise for the lower CFs). Therefore in the Zilany *et al.* (2009) model, the steady-state rate (*Ass*) parameter of the double-exponential adaptation was adjusted as a function of CF to achieve the desired responses to CF tones. However, this change resulted in higher saturation discharge rates (to CF tones) for model fibers with CFs above ∼2 kHz than reported in Liberman (1978). In addition, this adjustment created a problem of much higher saturation rates of model high-CF fibers to low-frequency tones than the saturation responses of low-frequency fibers to the same low-frequency stimulus. Simply readjusting the parameter *Ass* as a function of CF can solve the former problem, i.e., the model saturation rate can be

made to match AN data for high-CF fibers. However, the latter problem still persists; the response area plot in Fig. 1(A) shows that the rates of high-CF fibers in response to a low frequency tone (500 Hz) were much higher than the responses of low-CF fibers to the same tone. To solve this problem, it was necessary to readjust the parameters of the power-law model to achieve responses that agree with the cat AN data.

A general model of power-law adaptation (Drew and Abbott, 2006) is described as follows. Suppose, a stimulus $s(t)$ produces a response $r(t)$ that feeds back into an integrator $I(t)$, such that the adapted response is described as $r(t) = \max[0, s(t) - I(t)]$. The suppressive effects are accumulated with power-law memory, and are given by the following equation:

$$I(t) = \alpha \int_0^t \frac{r(t')}{t - t' + \beta} dt' = \alpha \; r(t) * f(t), \quad \text{where } f(t) = 1/(t + \beta).$$

(1)

Here α is a dimensionless constant that controls the amount of adaptation, β is a parameter with units of time, and $*$ indicates convolution (Drew and Abbott, 2006).

In this work, the amount of adaptation applied to higher CFs has been readjusted by changing the parameter α in the slow power-law path. With the combination of changing α from 5×10^{-6} to 2.5×10^{-6} and readjustment of *Ass* as a function of CF [$Ass = 800 \times (1 + CF/10^5)$ where CF is in units of Hz], the desired responses of the model have been achieved. The responses of the new model are shown in Fig. 1(B) along with the responses from the 2009 version of the model. The model saturation rate as a function of characteristic frequency is compared with the physiological data from cats (Liberman, 1978) in Fig. 2. The acoustic stimuli used were 50-ms tone bursts with a 2.5 ms rise-fall time, and

been adopted to compute the mean and variance of the discharge rate as a function of time from the model synapse output (before the discharge generator), accounting for the refractory effects of the AN. A new version of the model incorporating all of the above-mentioned changes is available on the following website (http://www.urmc.rochester.edu/labs/Carney-Lab/publications).
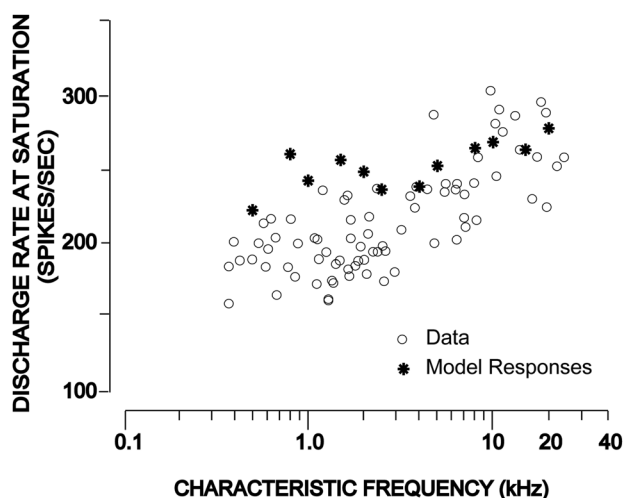

FIG. 2. Discharge rate at saturation as a function CF for HSR fibers. Physiological saturation rates (reprinted Fig. 17 from Liberman, 1978 with permission), shown by circles, were determined at each CF from the AN rate-level responses for sound levels up to and including 54 dB above threshold. Model saturation rate (shown by asterisks) at each CF was estimated from the simulated rate-level responses using the paradigm in Liberman (1978), except that the sound levels were extended up to 100 dB SPL to obtain a measurable saturation rate.

the rate of stimulus presentation was 10/s. Following the same paradigm, model rate-level responses were simulated for a range of CFs for sound levels ranging from 0 to 100 dB sound pressure level (SPL), and the discharge rate at saturation was estimated for each CF. Although the proposed model responses (shown by asterisks) are within a reasonable range of the published physiological data (circles), the discrepancy at lower CFs could lie in the fact that the saturation rate for the physiological data (Liberman, 1978) was computed from the rate-level responses for sound levels up to and including 54 dB above threshold. However, the model responses have been simulated at even higher levels to be able to estimate a measurable saturation rate. The new model has also been extensively tested against the published and recorded data from the AN, and the results are summarized in Table I. As expected, only the responses to forward-masking paradigm are affected by this change, although minimally; other responses are apparently unchanged.

In the 2009 version of the model, a fractional Gaussian noise (fGn) was added in the slow power-law path to model the distribution of spontaneous rates reported in Liberman (1978). A fractional (fractal) Gaussian noise is a generalization of the common white Gaussian noise that shows long range dependence. In this work, the option to use a pre-determined fixed seed instead of a variable seed for the fGn is made available; using the fixed seed removes the trial-to-trial variability in this aspect of the model response, which is useful to separate the independent effects of internal (physiological) and external (stimulus-driven) noise. The fixed seed chosen in the code gives a fairly average synapse response.

B. Analytical method to estimate mean rate and variance from the synapse output

The output of the model IHC section drives the IHC-AN synapse model, which provides an instantaneous synaptic release rate as output without taking into account the refractory effects (absolute and relative). In the final stage of the model, the discharge times are produced by a renewal process that includes refractory effects. However, to obtain a reliable post-stimulus time histogram (PSTH) based on the output of the discharge generator requires simulating responses to multiple repetitions of the same stimulus. To make the model computationally efficient, it is useful to employ a synapse output for which the effects of refractoriness can be directly estimated. Thus an analytical methodology has been applied here to calculate an approximate mean

TABLE I. Lists of model response properties that remained almost unchanged from the 2009 version after the modifications proposed in this paper.

No.	Response property
1	Distribution of spontaneous rates (Liberman, 1978)
2	Long-term recovery (Young and Sachs, 1973)
3	Responses to stimuli with amplitude increments/decrements (Smith et al., 1985)
4	Responses to SAM tones (Joris and Yin, 1992)
5	Responses to noise stimuli (Louage et al., 2004)

rate and variance as a function of time, including refractory effects from the model synapse output.

Vannucci and Teich (1978) derived a general expression for the dead-time (absolute refractory period) modified mean and variance for a counting process when the rate of input is an arbitrary function of time. Edwards and Wakefield (1990) also derived an analytical solution to compute the mean rate of nerve firing as a function of time [Eqs. (21) and (22) in Edwards and Wakefield, 1990], considering both the absolute and relative refractory effects. The latter method requires recursion and is computationally very expensive. However, the former method does not include relative refractory effects. In both cases, the assumption is that the rate of the input process is nearly constant with respect to the duration of the absolute refractory period. In the revised model presented here, the mean rate was computed from the model synapse output using the same recovery function as used in the discharge generator section of the 2009 version of the model (Zilany et al., 2009). The estimated rates are almost the same for both methods (Vannucci and Teich, 1978; Edwards and Wakefield, 1990) when the synapse output is fairly constant. Note that the estimate of mean rate deviates from the model PSTH when the synapse output changes rapidly because it violates the basic assumption made in both cases. This has been illustrated in Fig. 3. When the synapse output (solid line) is nearly constant or varies slowly, the estimated mean rate (dashed line, using Vannucci and Teich, 1978) closely matches the model PSTH (dotted line) from the output of the discharge generator [Fig. 3(A)]. However, when the synapse output varies abruptly, especially at the onset of the responses at higher levels [Fig. 3(B)], the estimated mean rate is substantially higher than the rate computed from the PSTH of the model AN discharge generator. However, the estimated mean rate still provides a better prediction of the PSTH than does the synapse output.

To make the calculation computationally less expensive, the following expression by Vannucci and Teich (1978) has been implemented in the code:

$$R(t) = \frac{Sout(t)}{1 + \tau Sout(t)}, \qquad (2)$$

where $R(t)$ is the mean discharge rate as a function of time, $Sout(t)$ is the model synapse output (spikes/s), and τ is the absolute refractory period (here fixed to 0.75 ms). The variance of the rate can be computed as [Eq. (12) in Vannucci and Teich, 1978]

$$\sigma^2(t) = \frac{Sout(t)}{[1 + \tau Sout(t)]^3}. \qquad (3)$$

III. CONCLUSIONS

A revised phenomenological model of the auditory periphery is presented here. The changes made to the 2009 version corrected the model saturation rates as a function of CF without adversely affecting other response properties. The forward-masking properties were affected to a small degree

J. Acoust. Soc. Am., Vol. 135, No. 1, January 2014

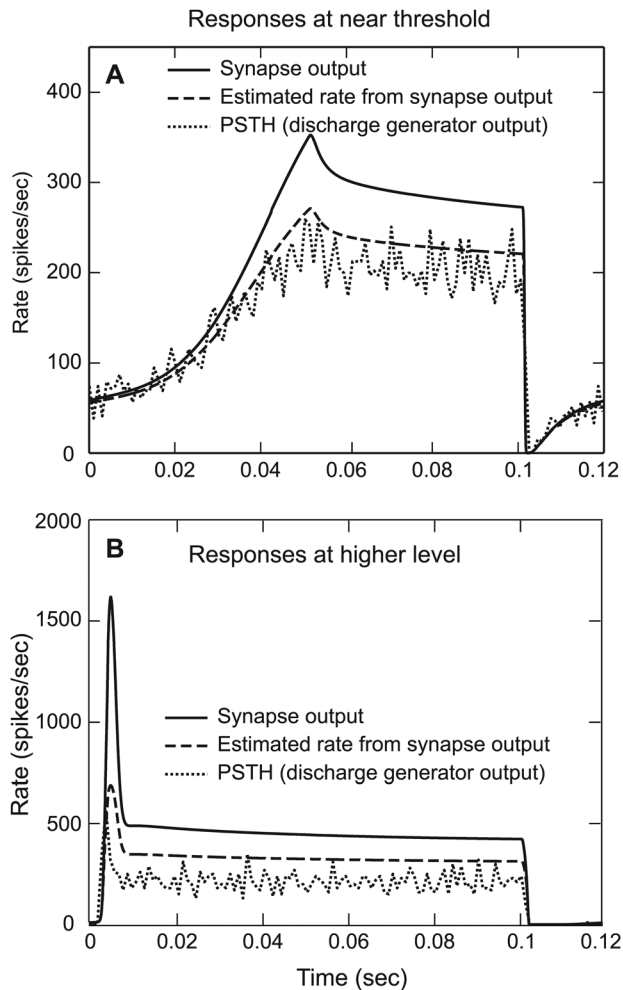Zilany et al.: An updated model of the auditory periphery   285

FIG. 3. Illustrations of the effect of changes in the stimulus on the model synapse output and discharge generator. The response of a tone at CF is shown for a model HSR fiber with CF equal to 10 kHz. The duration of the stimulus was 100 ms with a rise time of 50 ms (slowly varying). PSTH responses were simulated for 100 repetitions of the same stimulus, and the repetition time was 120 ms. (A) Responses at near threshold (8 dB SPL). (B) Responses at 40 dB SPL. The estimate of mean rate (dashed line) deviates from the PSTH output (dotted line) when the synapse output (solid line) varies abruptly especially at the onset of the model AN responses to higher level tones.

as could be expected after modifying synaptic adaptation. The effects of refractoriness have been incorporated into the model of the IHC-AN synapse. The model is now a better candidate to examine realistic neural-encoding hypotheses, especially those involving high CFs.

Bruce, I. C., Sachs, M. B., and Young, E. D. (**2003**). "An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses," J. Acoust. Soc. Am. **113**, 369–388.

Carney, L. H. (**1993**). "A model for the responses of low-frequency auditory-nerve fibers in cat," J. Acoust. Soc. Am. **93**(1), 401–417.

Drew, P. J., and Abbott, L. F. (**2006**). "Models and properties of power-law adaptation in neural systems," J. Neurophysiol. **96**, 826–833.

Edwards, B. W., and Wakefield, G. H. (**1990**). "On the statistics of binned neural point processes: The Bernoulli approximation and AR representation of the PST histogram," Biol. Cybern. **64**, 145–153.

Joris, P. X., and Yin, T. C. T. (**1992**). "Responses to amplitude-modulated tones in the auditory nerve of the cat," J. Acoust. Soc. Am. **91**, 215–232.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," J. Acoust. Soc. Am. **63**(2), 442–455.

Louage, D. H. G., Heijden, M. v. d., and Joris, P. X. (**2004**). "Temporal properties of responses to broadband noise in the auditory nerve," J. Neurophysiol. **91**, 2051–2065.

Smith, R. L., Brachman, M. L., and Frisina, R. D. (**1985**). "Sensitivity of auditory-nerve fibers to changes in intensity: A dichotomy between decrements and increments," J. Acoust. Soc. Am. **78**(4), 1310–1316.

Vannucci, G., and Teich, M. C. (**1978**). "Effects of rate variation on the counting statistics of dead-time-modified Poisson processes," Opt. Commun. **25**(2), 267–272.

Westerman, L. A., and Smith, R. L. (**1988**). "A diffusion model of the transient response of the cochlear inner hair cell synapse," J. Acoust. Soc. Am. **83**(6), 2266–2276.

Young, E. D., and Sachs, M. B. (**1973**). "Recovery from sound exposure in auditory-nerve fibers," J. Acoust. Soc. Am. **54**(6), 1535–1543.

Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (**2001**). "A phenomenological model for the responses of auditory-nerve fibers. I. Nonlinear tuning with compression and suppression," J. Acoust. Soc. Am. **109**, 648–670.

Zilany, M. S. A., and Bruce, I. C. (**2006**). "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," J. Acoust. Soc. Am. **120**, 1446–1466.

Zilany, M. S. A., and Bruce, I. C. (**2007**). "Representation of the vowel /ɛ/ in normal and impaired auditory nerve fibers: Model predictions of responses in cats," J. Acoust. Soc. Am. **122**(1), 402–417.

Zilany, M. S. A., Bruce, I. C., Nelson, P. C., and Carney, L. H. (**2009**). "A phenomeno-logical model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," J. Acoust. Soc. Am. **126**, 2390–2412.

Zilany, M. S. A., and Carney, L. H. (**2010**). "Power-law dynamics in an auditory-nerve model can account for neural adaptation to sound-level statistics," J. Neurosci. **30**(31), 10380–10390.