Modeling signal propagation in the human cochlea

Stephen T. Neely[a] and Daniel M. Rasetshwane
Boys Town National Research Hospital, 555 North 30th Street, Omaha, Nebraska 68131, USA

(Received 12 February 2017; revised 28 September 2017; accepted 29 September 2017; published online 19 October 2017)

The level-dependent component of the latency of human auditory brainstem responses (ABR) to tonebursts decreases by about 38% for every 20-dB increase in stimulus level over a wide range of both frequency and level [Neely, Norton, Gorga, and Jesteadt (1998). J. Acoust. Soc. Am. **31**, 87–97]. This level-dependence has now been simulated in an active, nonlinear, transmission-line model of cochlear mechanics combined with an adaptation stage. The micromechanics in this model are similar to previous models except that a dual role is proposed for the tectorial membrane (TM): (1) passive sharpening the tuning of sensory-cell inputs (relative to basilar-membrane vibrations) and (2) providing an optimal phase shift (relative to basilar-membrane vibrations) of outer-hair-cell feedback forces, so that amplification is restricted to a limited range of frequencies. The adaptation stage, which represents synaptic adaptation of neural signals, contributes to the latency level-dependence more at low frequencies than at high frequencies. Compression in this model spans the range of audible sound levels with a compression ratio of about 2:1. With further development, the proposed model of cochlear micromechanics could be useful both (1) as a front-end to functional models of the auditory system and (2) as a foundation for understanding the physiological basis of cochlear amplification. © *2017 Acoustical Society of America*. https://doi.org/10.1121/1.5007719

[CAS] Pages: 2155–2167

I. INTRODUCTION

The primary sensory organ for our sense of hearing is the cochlea, which is the part of the inner ear that converts vibrations caused by sounds into the neural signals that are conveyed by the auditory nerve (AN) to the brain. Computational modeling is one of many tools that researchers use to answer questions about cochlear function (e.g., Allen and Neely, 1992). One question is how the cochlea delivers to each AN fiber only a specific range of frequencies. Another question is how the cochlea compresses the large dynamic range of audible sounds into the much smaller dynamic range of neural signals. The answers to these questions lie within the realm of cochlear micromechanics.

There are two types of sensory cells in the cochlea that detect vibrations: inner hair cells (IHCs), which are mainly responsible for converting cochlear vibrations into neural signals, and outer hair cells (OHCs), which have a different functional role that utilizes their shape-shifting ability (called motility) to influence cochlear vibrations (Brownell *et al.*, 1985). Although it is widely accepted that OHC motility amplifies cochlear vibrations (e.g., Dallos and Evans, 1995), evidence supporting this action is inconclusive and some details of OHC function remain controversial (e.g., van der Heijden and Versteegh, 2015). A goal of modeling cochlear micromechanics is to contribute toward resolving the lingering uncertainty surrounding the role of OHCs underlying these processes (Allen and Neely, 1992).

The existence of cochlear compression implies the presence of nonlinear elements in cochlear mechanics. An important consequence of the fact that cochlear mechanics is nonlinear is that the latency of AN responses to transient sounds is level dependent. Specifically, the *latency* of AN transient responses decreases as sound level increases. Evidence suggests that the level-dependent component of the latency of human auditory brainstem responses (ABR) to toneburst stimuli decreases by about 38% for every 20 dB increase in stimulus level (Neely *et al.*, 1988; Rasetshwane *et al.*, 2013). This rate of latency decrease remains approximately constant across a wide range of toneburst frequencies and intensities.

The level dependence of the latency of cochlear transients has previously been partially simulated in models of cochlear mechanics by applying a saturating nonlinearity to the feedback force associated with OHCs (e.g., Neely, 1989). However, modeling the full frequency and level range of this level dependence has been elusive (e.g., Neely, 1989; Rønne *et al.*, 2012; Verhulst *et al.*, 2016).

The present study proposes a new way to model OHC feedback forces that operates over a wider range of stimulus levels and stimulus frequencies than has been achieved in previous models. An important feature of this model is that it maintains numerical stability for wideband stimuli in an active, nonlinear time–domain implementation. The key to the success of this approach is a new interpretation of the OHC feedback force in which the tectorial membrane (TM) plays an essential role in controlling the activation of cochlear amplification.

II. METHODS

The model equations and mathematical notation presented in this are essentially the same as Neely and Kim (1986) and Neely (1990). Differences in the present model formulation include (1) the physiological interpretation of

[a]Electronic mail: Stephen.Neely@boystown.org

micromechanics, (2) details of the OHC nonlinearity, and (3) the inclusion of an adaptation stage. A frequency–domain formulation of the equations is useful for understanding how the model input to the hair cells differs from basilar membrane (BM) vibrations; however, a time–domain formulation is required to properly simulate the level dependence of OHC feedback forces.

A. Fluid mechanics

A one-dimensional (or transmission-line) approximation of cochlear–fluid mechanics (e.g., Zweig *et al.*, 1976) can be written as

$$\frac{d^2}{dx^2}p_d(x) = \frac{2\rho}{H}\ddot{\xi}_p(x), \tag{1}$$

where p_d is the difference in pressure across the cochlear partition (CP), ξ_p is the average displacement across the width of the CP, H is the height of the scale, and ρ is the density of the fluid. (The dots above ξ_p indicate a second derivative with respect to time.) A finite-difference approximation of the second-order spatial derivative (with respect to x) in Eq. (1) yields the following equation:

$$2p_d(x) - p_d(x-\Delta) - p_d(x+\Delta) = -m_f\ddot{\xi}_p(x), \tag{2}$$

where $\Delta = L/(N-1)$ is the small distance along the CP over which mechanical properties of the cochlea are lumped as a single section, L is the length of the CP, N is the number of discrete location that represent this length in the model, and $m_f = 2\rho\Delta^2/H$ is the effective mass (per unit area) of the scala fluid. In the cochlea, scala height decreases from base to apex, which has been shown to have an important influence on cochlear input impedance (e.g., Puria and Allen, 1991; Shera and Zweig, 1991). In the present model, scala height decreases exponentially such that the height at the apex is about one-fourth the height at the base.

B. Cochlear micromechanics

The macromechnical equations that govern the relation between pressure difference p_d and CP displacement ξ_p are described by Eqs. (1) and (2). The variables p_d and ξ_p are also related by cochlear micromechanics.

In the model, each lumped section of the micromechanics contains two mechanical degrees-of-freedom (DOF), as shown in Fig. 1. The mass m_1 represents the CP, which comprises the portion of the Organ of Corti (OC) between (and including) the BM and reticula lamina (RL). The mass m_2 represents the TM, which is attached to the bony spiral limbus and is also elastically coupled to the RL (through the stiffness of the OHC hair bundles and the viscosity of the subtectorial fluid).

In the frequency domain, the micromechanical relation between p_d and ξ_p can be expressed in terms of a partition impedance (Z_p) and a Fourier transform of $\dot{\xi}_p$ (denoted as $\dot{\Xi}_p$):

$$P_d(x) = Z_p(x)\,\dot{\Xi}_p(x), \tag{3}$$

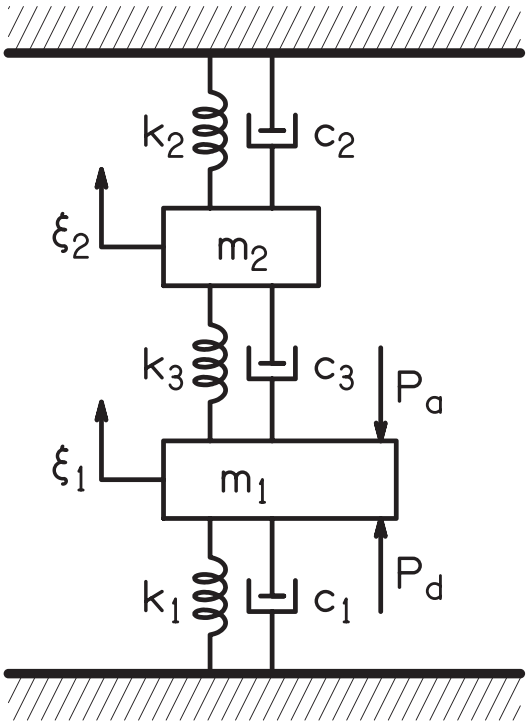

FIG. 1. Lumped-element model of cochlear micromechanics with two DOF. The mass m_1 represents the portion of the OC between the BM and reticular lamina, which is attached to rigid bone by stiffness and damping elements k_1 and c_1. A difference in fluid pressure (P_d) across the OC creates a force that acts on the OC mass. Feedback force due to OHC motility creates a second force that also acts on the OC mass and is represented in this figure as P_a. The mass m_2 represents the TM, which is attached to rigid bone by k_2 and c_2. The two masses are coupled to each other by k_3 and c_3. Displacements of m_1 and m_2 are denoted by ξ_1 and ξ_2, respectively. In the model, the relative (or shear) displacement between m_1 and m_2 (defined as $\xi_c = \xi_1 - \xi_2$) becomes the input to outer hair cells. The relative velocity $\dot{\xi}_c$ drives inner hair cells.

$$bZ_p = Z_1 + (Z_3 - \gamma Z_4)H_c, \tag{4}$$

$$H_c = \frac{Z_2}{Z_2 + Z_3}, \tag{5}$$

where P_d is the Fourier transform of p_d, $Z_1 = m_1 s + c_1 + k_1/s$, $Z_2 = m_2 s + c_2 + k_2/s$, $Z_3 = c_3 + k_3/s$, and $Z_4 = c_4 + k_4/s$ are impedance functions, which are specified in the frequency domain, and $s = i\omega$.[1]

To compensate for the fact that local displacement of the CP varies across its width, a scale factor b is introduced in Eq. (4) that relates the average displacement of CP (across its width) to the point displacement of m_1 by the equation $\xi_p = b\xi_1$. In effect, the scale factor b resolves the need to maintain conservation of volume in the fluid mechanics with the need to report point displacements of the micromechanics for comparison with experimental measurements. For simplicity, b is a constant (independent of x) in the present model.

The term Z_4 in Eq. (4) implements an OHC feedback force that is controlled by OHC hair-bundle deflection and acts on the CP. This feedback force provides the basis for amplification of CP vibrations. The parameter γ, which is a constant in the frequency–domain model and varies with hair-bundle displacement in the time–domain model,

controls the magnitude of the OHC feedback force and, thereby, the amplifier gain. The time dependence of γ is what causes the time–domain model to become level dependent or, equivalently, to be nonlinear. The model is *linear* when γ is constant and *nonlinear* when γ varies with time. The parameter γ also determines when the model is active. The model is *active* when $\gamma > 0$ and *passive* when $\gamma = 0$. The model is most sensitive when $\gamma = 1$.

The signal delivered to both IHCs and OHCs is the deflection of their hair bundles. Because the OHC hair bundle is embedded in the TM, deflection of the OHC hair bundle is equal to the relative displacement $\xi_c = \xi_1 - \xi_2$ between the TM and the CP. In Eq. (5), H_c is defined as a transfer function between CP displacement and OHC hair-bundle deflection. So, in the frequency domain, $\dot{\Xi}_c = H_c \dot{\Xi}_1$.

The role of H_c in sharpening the tuning of cochlear responses as a passive filter has been suggested previously by Zwislocki and Kletsky (1979) and Allen (1980). In the present model, passive sharpening due to H_c is mainly limited to the attenuation of low-frequency responses. However, in addition to this passive sharpening, H_c also contributes to active sharpening by limiting the frequency range over which OHC feedback forces can counteract intrinsic damping. This point is discussed further below.

The magnitude and phase of $H_c(x)$ are shown in Fig. 2 at four frequencies (1, 2, 4, and 8 kHz). Note that the "high-pass filter" shape of the transfer-function magnitude (in the upper panel of Fig. 2) provides an improvement in frequency selectivity of about 12 dB between its maximum and minimum.[2] Also, note that (at each frequency) the phase of H_c increases to more than one-eighth cycle. As we shall see below, one consequence of this phase shift is that the OHC feedback force is shifted toward *negative* damping.

The IHC hair bundle is not embedded in the TM, so it is displaced by viscous drag due to motion of the sub-tectorial fluid that surrounds it. Considering this, we follow Verhulst *et al.* (2016) in modeling the IHC hair-bundle deflection as being proportional to the velocity $\dot{\xi}_c$.

The real part of $Z_p(x)$ is shown in Fig. 3 at the same four frequencies as in Fig. 2. The solid and dashed lines in Fig. 3 show the impedance when $\gamma = 1$ and $\gamma = 0$, respectively. At each frequency, the real part of $Z_p(x)$, which represents the damping effect of the CP on motion of the cochlear fluid, becomes negative across a limited distance (when $\gamma = 1$). Note that the damping becomes most negative near the location where the phase of $H_c(x)$ equals one-eighth cycle, which is indicated by the filled circle on each curve. The presence of negative damping means that the CP introduces energy into the cochlear fluid. The source of this energy is OHC motility. The open circles indicate locations of maximum response.

A time–domain implementation is required to include nonlinear elements in the model. In our time–domain model, the only time-varying parameter is γ, which was introduced in Eq. (4) as a multiplier of the OHC feedback force. In other words, the only nonlinear element in the model represents OHC mechanoelectric transduction. The OHC force is greatest when $\gamma = 1$, which occurs when signal levels are very low. The OHC force is absent when $\gamma = 0$, which is the passive condition of the model that occurs (1) when signal levels are extremely high or (2) when OHC function is impaired. The present time–domain model uses the following equation to gradually decrease γ as the magnitude of ξ_c increases:

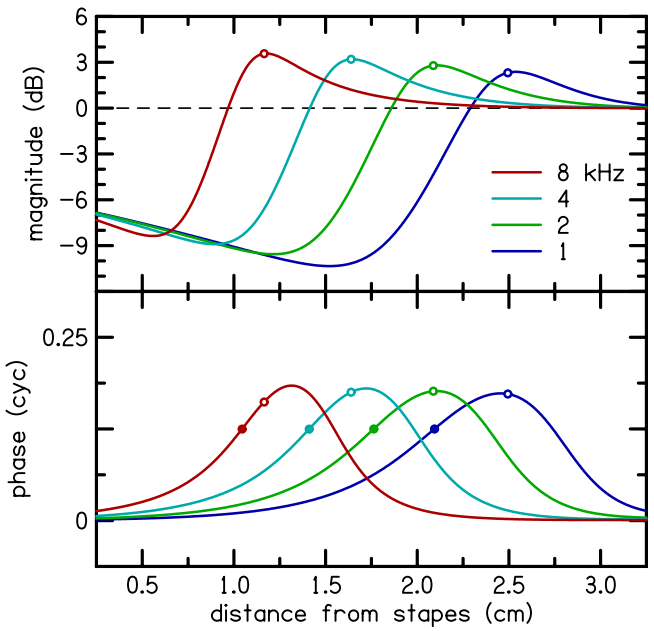

FIG. 2. (Color online) Transfer function between basilar-membrane (or reticular lamina) displacement and the shearing displacement between reticular lamina and the TM $H_c(x)$ at four frequencies, 1, 2, 4, and 8 kHz. In the model, this shearing displacement becomes the input to both IHCs and OHCs. These results were computed from a linearized version of the active, time–domain cochlear model. The filled circles indicate the locations where the phase equals one-eighth cycle, which is equal to 45° or $\pi/4$ radians. The open circles indicate the location of maximum OHC displacement (ξ_c).

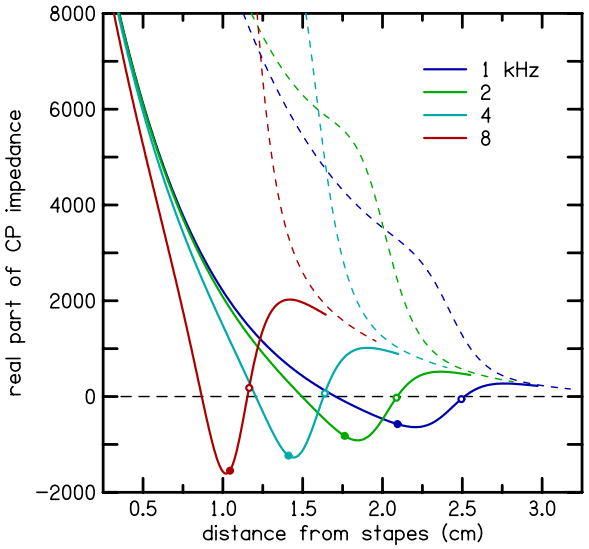

FIG. 3. (Color online) Real part of the impedance of the CP $Z_p(x)$ at four frequencies, 1, 2, 4, and 8 kHz. The solid lines represent the active condition of the model with $\gamma = 1$. The dashed lines represent the passive version of the model with $\gamma = 0$. These results were computed from a linearized version of the active, time–domain cochlear model. The filled circles indicate the locations where the transfer-function (H_c) phase equals one-eighth cycle (see Fig. 2). The open circles indicate the locations of maximum OHC displacement (ξ_c).

$$\gamma = \begin{cases} 1, & \xi_c \le d_m, \\ \dfrac{1}{1 + a_\gamma \cdot \ln|\xi_c/d_m|}, & \xi_c > d_m. \end{cases} \quad (6)$$

In the model, γ varies with time and location, but only because of its dependence on ξ_c. The parameters a_γ and d_m in Eq. (6) are (for simplicity) constant across time and location.[3]

C. Cochlear model parameters

Parameter values for the present model are listed in Table I. Parameters associated with fluid mechanics were selected to represent human cochlear anatomy and have the same values as in previous versions of the model (e.g., Neely, 1990). Other parameters (associated with cochlear micromechanics) were selected for consistency with (1) the human frequency-to-place map suggested by Greenwood (1990) and (2) the level-dependence of ABR Wave-V latency reported first by Neely *et al.* (1988) and replicated by Rasetshwane *et al.* (2013). Model results that demonstrate this consistency are presented below.

D. Stimulus delivery and model response

In the model, stimuli were delivered to the cochlea though simple earphone and ear-canal/coupler sections, which were connected (by an eardrum) to a middle-ear section. These sections are the same in the present model as described by Liu and Neely (2010).

For the purposes of this paper, the primary measure of the model's response was a whole-nerve response (WNR), which was defined as the sum of the outputs of all IHCs at each time step. To demonstrate how synapses may influence the latency of cochlear transient responses, each IHC output was computed in two different ways: (1) as a voltage representing an IHC receptor potential due to hair-bundle deflection or (2) as an instantaneous neural spike rate obtained by passing the IHC voltage through an adaptation stage.

The adaptation stage is intended to represent the combined effect of all of the IHC synapses of the subset of AN fibers attached to a single IHC. The adaption stage enhances the WNR onset by reducing steady response that occurs after the onset. This sensitivity reduction persists for a few milliseconds after the stimulus ends and gradually recovers.

TABLE I. Model parameter values (cgs) units. The spatial dependence of parameter values is expressed in terms of a longitudinal place variable $\check{x} = x \cdot [1 + (0.234x)^6]$, which increases parameter gradients near the apex to better fit to the desired frequency-place map.

$k_1(x) = 2.394 \times 10^8 e^{-2.666\check{x}}$ (dyn cm^{-3})	$L = 3.5$ (cm)
$c_1(x) = 871.4e^{-1.394\check{x}}$ (dyn s cm^{-3})	$H = 0.1e^{-0.4x}$ (cm)
$m_1(x) = 7.417 \times 10^{-3} e^{-0.063\check{x}}$ (gm cm^{-2})	$W = 0.1$ (cm)
$k_2(x) = 3.036 \times 10^8 e^{-3.476\check{x}}$ (dyn cm^{-3})	$\rho = 1$ (gm cm^{-3})
$c_2(x) = 1979e^{-1.247\check{x}}$ (dyn s cm^{-3})	$b = 0.05$
$m_2(x) = 3.417 \times 10^{-2} e^{-0.083\check{x}}$ (gm cm^{-2})	$N = 3501$
$k_3(x) = 3.151 \times 10^8 e^{-2.909\check{x}}$ (dyn cm^{-3})	$a_\gamma = 0.05$
$c_3(x) = 1.049e^{-0.103\check{x}}$ (dyn s cm^{-3})	$d_m = 6 \times 10^{-9}$ (cm)
$k_4(x) = 4.045 \times 10^8\, e^{-2.795}$ (dyn cm^{-3})	$t_{ihc} = 0.2$ (ms)
$c_4(x) = 0$	

IHC voltage was computed by taking a quantity that was proportional to the shear velocity $\dot{\xi}_c$, which was described above in the context of cochlear micromechanics, then subjecting it to (1) half-wave rectification and (2) a first-order low-pass filter (with a time-constant of $t_{ihc} = 0.2$ ms). The rectification is attributed to mechanoelectric transduction at the IHC hair bundle and the low-pass filter is attributed to the combination of IHC membrane capacitance and conductance. For comparison, the time-constant of the IHC membrane is reported to be 0.26 ms and constant throughout the cochlea (Johnson *et al.*, 2011).

Our adaptation-stage formulation is essentially a nonlinear single-reservoir diffusion model. The time-varying inner-hair-cell membrane potential $v_{ihc}(t)$ is the input to the model. The model output $r_{ns}(t)$ is similar to a neural spike rate. For convenience, the synapse–model equations are expressed in terms of the dimensionless quantities $v_n = v_{ihc}(t)/v_{norm}$ and $r_n = r_{ns}(t)/r_{norm}$,

$$\dot{\sigma}_1 = -\rho_d r_n + \rho_r(1 - \sigma_1), \quad (7)$$

$$\sigma_2 = \begin{cases} 0, & \sigma_1 \le 0, \\ \sigma_1, & 0 < \sigma_1 < 1, \\ 1, & \sigma_1 \ge 1, \end{cases} \quad (8)$$

$$r_n = \sigma_2^\alpha v_n. \quad (9)$$

In these equations, σ_1 and σ_2 are similar to diffusion permeabilities, the dot above σ_1 denotes a time derivative, ρ_d is the reservoir depletion rate, ρ_r is the reservoir recovery rate, and α is an exponent that has the effect of increasing the amount of forward masking produced by the model. For the results presented in this paper, the model parameters were $\rho_d = 1.906$/s, $\rho_r = 17.52$/s, and $\alpha = 10.78$. The model parameters (including v_{norm}) were selected to achieve a favorable compromise between the onset enhancement needed for latency comparisons and the amount of masking needed for forward-masking comparisons.

III. RESULTS

Features of the most sensitive state of the model, which represents its behavior for the lowest audible stimulus levels, are best demonstrated by impulse responses from a linear version of the time–domain model with $\gamma = 1$. A nonlinear version of the model is required to demonstrate the level dependence of the latency of cochlear responses to toneburst stimuli. Model responses to longer-duration tones provide additional context for consideration of (1) tone detection, (2) tone-on-tone forward masking, and (3) power flow as a function of spatial location.

A. Power flow

Figure 4 shows the cochlear–partition displacement (ξ_1) magnitude (upper panel) and power flow[4] (lower panel) as functions of place for a 1 kHz stimulus at two different levels. At 0 dB sound pressure level (SPL), power flow is

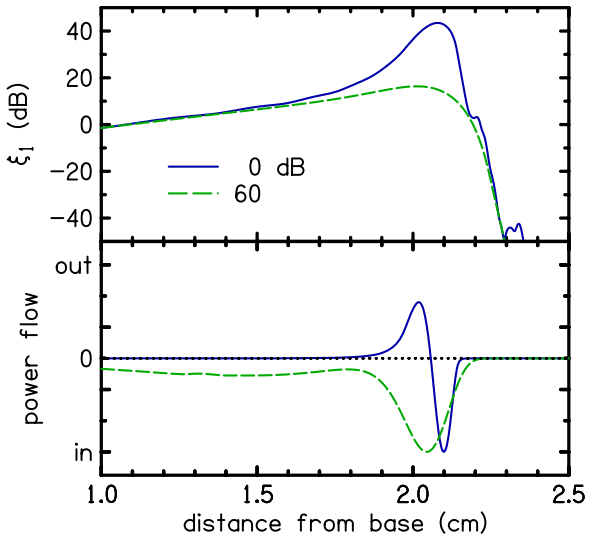

FIG. 4. (Color online) Cochlear-partition displacement (ξ_1) and power flow at 1 kHz as a function of place at two stimulus levels (0, 60 dB SPL). The dB units for displacement magnitude (upper panel) are referenced to 1 pm and 1 nm for 0 and 60 dB SPL, respectively, which causes the two curves to have approximately the same dB values between $x = 1.0$ and $x = 1.5$. The two power-flow curves (lower panel) are normalized to have the same maximum *in-flow* to facilitate visual determination of the regions where power flow is positive or negative, which indicates that power is being produced or absorbed by the CP, respectively.

positive at some locations. Positive power flow means that power is flowing *out* of the CP, which is evidence of a net power gain.[5] At 60 dB SPL, power flow is negative everywhere. Negative power flow means that power is flowing into the CP, which indicates the absence of any power gain.

Displacement gain (defined as CP displacement amplitude at the characteristic place when $\gamma = 1$ compared to the amplitude when $\gamma = 0$) is typically larger in the model than power gain. Figure 5 shows how displacement gain and power gain both decrease as stimulus level increases. Note that the slope of the displacement gain is about −0.5, which is equivalent to a compression ratio of 2:1.

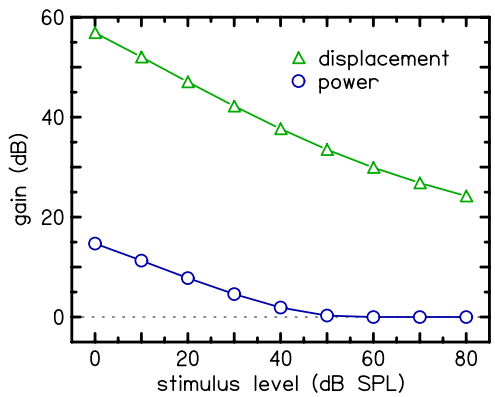

FIG. 5. (Color online) Cochlear-amplifier gain as a function of stimulus level for a 1-kHz tone. *Displacement* gain (triangles) is defined as the amplitude of displacement of the CP at the 1-kHz characteristic place ($x = 2.051$ cm) when $\gamma = 1$ relative to the amplitude at that place when $\gamma = 0$. *Power* gain (circles) is defined as the total power absorbed by the CP at 1 kHz relative to the power that enters the cochlea through the stapes.

B. Impulse responses

Isodisplacement tuning curves from the linearized time–domain model (with $\gamma = 1$) are shown in Fig. 6. The input stimulus to the model was a wide-band click. The model response was computed for 160 ms beyond the onset of the click. At five places (20, 40, 60, 80, and 90% of the distance from the base), CP displacement (ξ_p) and OHC displacements (ξ_c) were Fourier transformed (and then divided by the Fourier transform of the pressure at the eardrum) to obtain the curves shown in Fig. 6. The isodisplacement curves (upper panel), indicate the pressure needed at the eardrum to elicit a displacement of $\xi_c = 1$ nm (rms) in the model. The group delay of ξ_c (relative to eardrum pressure at the same five places; lower panel) is defined as *minus* the slope of the phase with respect to frequency.

The isodisplacement tuning curves (upper panel of Fig. 6) show an increase in sharpness of tuning as characteristic frequency (CF) increases. A useful measure of the sharpness of tuning is Q_{erb}, which is defined as its CF divided by the "equivalent rectangular bandwidth" of the tuning curve. Direct measurement of tuning curves is not possible for human cochlear responses; however, Q_{erb} may be inferred from human otoacoustic emission (OAE) latency (e.g., Shera *et al.*, 2002). Table II compares Q_{erb} of the tuning curves in Fig. 6 with estimates derived from stimulus-frequency OAE latency measurements. Compared to the OAE estimate, Q_{erb} in the active model is 37% less at 0.11 kHz and 41% greater at 7.8 kHz. In the passive model, Q_{erb} is about 1 at all frequencies.

Tonotopicity is a key feature of auditory-system organization, so it is important that the model reproduces the frequency-place map of the human cochlea. In Fig. 7, the

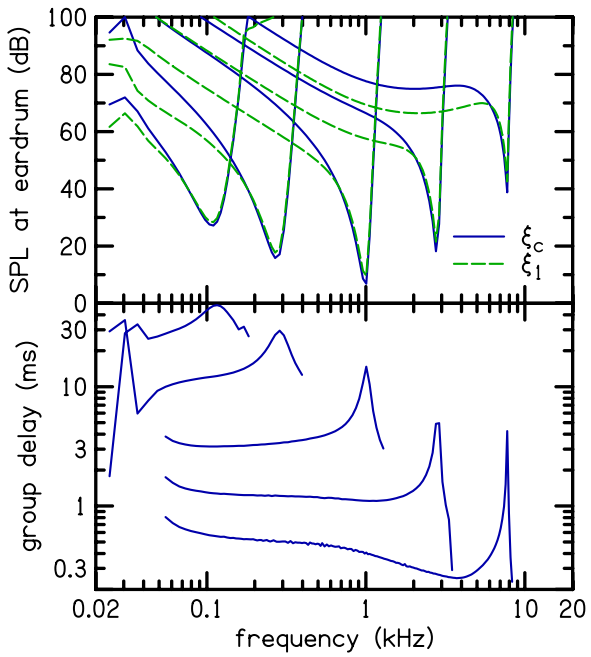

FIG. 6. (Color online) Isodisplacement tuning curves representing ξ_c and ξ_1 at five locations. The locations are at 20, 40, 60, 80, and 90% of the distance from the stapes (to the apex). These results were computed from a linear version of the time–domain cochlear model with $\gamma = 1$.

TABLE II. Sharpness of tuning (Q_{erb}) in the model compared with estimates derived from OAE latency measurements. For the model, Q_{erb} was calculated (as the best frequency divided by an equivalent-rectangular bandwidth) from the ξ_c tuning curves shown in Fig. 6 for both active and passive conditions of the model. The OAE estimates were calculated from a power law fit $Q_{erb} = \beta f^{\alpha}$ to SFOAE-delay data that was suggested by Shera *et al.* (2002) with $\alpha = 0.30$ and $\beta = 12.7$.

	0.11 kHz	0.27 kHz	1.0 kHz	2.8 kHz	7.8 kHz
Passive ξ_c	0.9	1.0	1.2	1.1	0.6
Active ξ_c	4.0	5.9	11.0	17.1	32.3
OAE estimate	6.4	8.4	12.2	16.8	22.9

model frequency–place map is compared to maps derived by Greenwood (1990) for the human cochlea and the cat cochlea. The good agreement between the present model and the human map required special attention to the slope of the model map at low frequencies (see Fig. 7). The map–slope decrease was achieved by increasing the rate-of-change (with respect to distance from the stapes) of all micromechanical model parameters near the cochlear apex (see Table I caption). This feature of the frequency-place map has the important benefit of reducing reflections from the helicotrema (Puria and Allen, 1991; Shera and Zweig, 1991), which would otherwise create standing waves at low frequencies.

C. Toneburst responses

The set of toneburst stimuli that were selected to demonstrate the latency of cochlear transient responses consists of Blackman-windowed tones at four frequencies (0.5, 1, 2, and 4 kHz) at four levels each (20, 40, 60, and 80 dB peSPL[6]). The duration of each toneburst is inversely proportional to the square root of its frequency (5.66, 4, 2.83, and 2 ms). Model responses to each of the tonebursts were computed for 40 ms beyond the onset of the stimulus. Examples of the time course of the WNR are shown in Fig. 8 for the 1-kHz toneburst at three levels (40, 60, and 80 dB peSPL). The middle and lower panels show WNR waveforms with and without inclusion of the adaptation stage, respectively. The latency of the WNR is measured from the onset of the stimulus to the peak (i.e., maximum value) of the WNR. In Fig. 8, WNR peaks are indicated by dots. Note that enhancement of the WNR onset when the adaptation stage is included causes more rapid decrease in the latency of the peak as level increases.

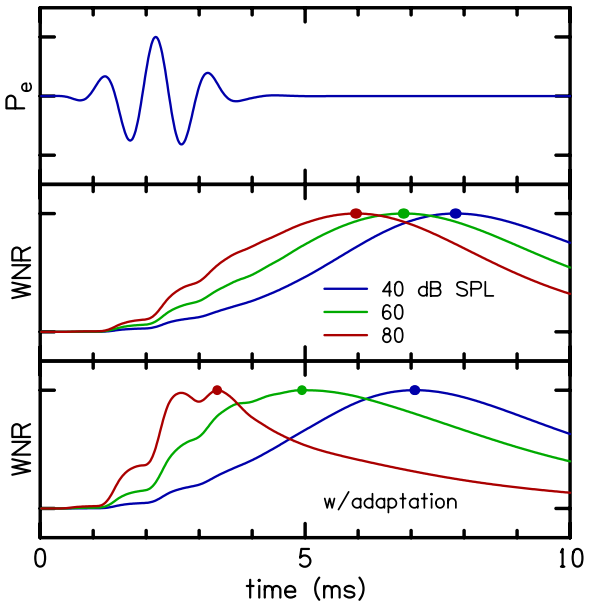

FIG. 8. (Color online) Pressure at eardrum (P_e) and WNR for a 1-kHz toneburst. Each curve is normalized to its maximum value. WNR is shown at three stimulus levels (40, 60, and 80 dB peSPL). The dots at the maximum value of each WNR curve indicate its latency, which decreases as stimulus level increases.

In Fig. 9(a), the WNR latency for all 16 toneburst stimuli (for the model without synapses) is compared to lines fitted to measurements of human Wave-V ABR latencies after subtraction of 5 ms attributed to neural propagation delay (Neely *et al.*, 1988). The dependence of the WNR latency on level and frequency is similar to the ABR data at the highest frequency, but the WNR latency is *less* dependent on level at the lowest frequency.

Figure 9(b) shows WNR latency for the same tonebursts when the adaptation stage was included. The addition of the

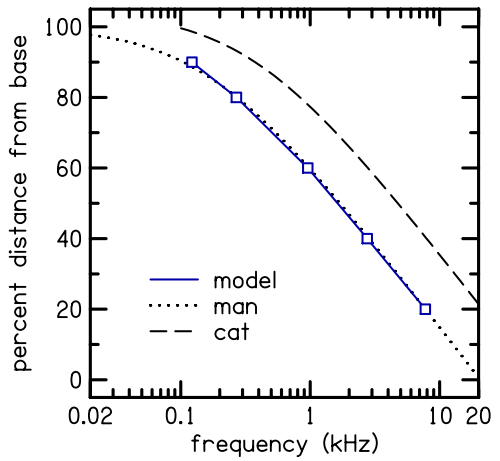

FIG. 7. (Color online) Frequency-place map showing the CF at five locations. The locations are the same as for the tuning curves in Fig. 4. These frequency-place results were computed from a linear version of the time–domain cochlear model with $\gamma = 1$. The man and cat maps shown for comparison are based on equations suggested by Greenwood (1990).

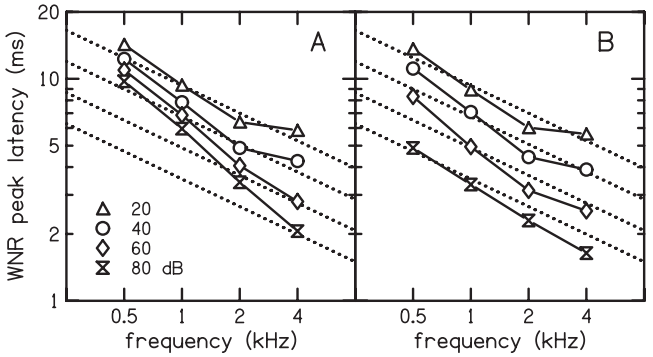

FIG. 9. WNR latency for tone bursts. The symbols indicate model latency at four levels (20, 40, 60, 80 dB peSPL) and four frequencies (0.5, 1, 2, 4 kHz). Panel A shows model results without any adaptation. Panel B shows model results with adaptation applied to each IHC. The dotted lines are fits to the level-dependent component of human ABR Wave-V latency (Neely *et al.*, 1988).

adaptation stage causes the latency to become more level dependent at the lowest frequency, but has little effect at the highest frequency. These results demonstrate that including the adaptation stage in the model improves the agreement between WNR latency and the ABR latency.

D. Tone detection and forward masking

Onset emphasis due to adaptation is illustrated in Fig. 10, which superimposes WNR for 1-kHz tones at three stimulus levels (80, 60, and 40 dB SPL). The duration of these tones is 100 ms, which includes 10 ms cosine-shaped ramps at the beginning and end. To facilitate comparison, WNR at all stimulus levels has been normalized to the peak value of WNR at 80 dB in the time range from 80 to 90 ms. Note that the WNR peak-to-steady ratio is 5.9 at 80 dB and decreases to 3.0 at 40 dB SPL.

Reduction in sensitivity of the adaption model due to the 100 ms tone is illustrated in Fig. 10 by showing a corresponding reduction in WNR when a second tone, which we call the *probe*, was presented with a delay of 20 ms after the end of the first tone, which we call the *masker*. The probe tone had a duration of 20 ms (with 10 ms ramps) and a level of 20 dB SPL. Figure 10 shows that WNR produced by the probe decreases in amplitude as masker level increases. The decrease in WNR amplitude is interpreted as an increase in forward masking due to reduced sensitivity of the adaptation stage.

The adaptation stage produces onset emphasis because steady signals experience decreased throughput. In other words, the adaptation stage adapts to its input signal level over a short period of time. After the input signal ends, the adaptation stage recovers from its reduced-sensitivity state over a short period of time. The adaptation model is intended to mimic a reduced sensitivity observed in IHC synapses that is known as short-term adaptation. When observed psychophysically, reduced sensitivity in the presence of a preceding tone is known as forward masking. Although short-term adaptation may play a role in forward masking, there are also differences between these two phenomena (e.g., Harris and Dallos, 1979), so these two phenomena are apparently not exactly equivalent. However, our adaptation stage is intended to combine simplified versions of both phenomena into a single stage of signal-processing.

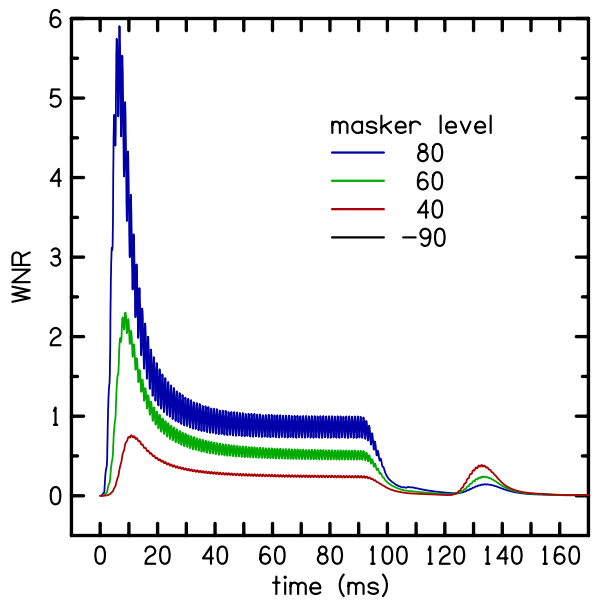

FIG. 10. (Color online) Time-course of WNR showing adaption due to IHC synapses. The three superimposed curves each show the WNR for two consecutive short-duration tones. The first tone had a duration of 100 ms (with 10-ms ramps) and three different levels (80, 60, and 40 dB peSPL). The second tone started 20 ms after the end of the first tone, had a duration of 20 ms, and had the same level (40 dB peSPL) for all curves. WNR has been normalized to have a steady value (defined as its local maximum in the 80 to 90 ms time interval) of 1 when the masker level is 80 dB.

Table III lists the signal levels of tones at several frequencies that produce the same WNR-peak value as a 1-kHz tone at 9 dB SPL. These signal levels are compared to *reference equivalent threshold* sound-pressure levels (RETSPL) for ear-canal sound pressure at threshold (ANSI, 2004). The model matches the RETSPL at 1 kHz (by design), but deviates by more than 40 dB at 8 kHz. This result suggests that further investigation is needed to determine whether the WNR is a reasonable proxy for behavioral threshold. However, in this paper, we restrict our consideration of masking to 1 kHz.

The set of stimuli that was selected to quantify forward masking in the model consists of 1-kHz maskers (with 250 ms duration) at three levels (40, 60, and 80 dB SPL) followed by a 1-kHz signal (with 20 ms duration) at one of three delays (10, 20, or 40 ms). The signal level was varied to determine the threshold level that produced the WNR-peak value as a signal at 19 dB SPL without a preceding tone (Jesteadt *et al.*, 1982). The *amount of masking* was defined as the dB difference between the masked and unmasked threshold levels. Figure 11 shows how the amount of masking produced by the model synapse (circles and solid lines) varies across the 9 stimulus conditions. For comparison, power-law fits to psychophysical measurements in human subjects (Jesteadt *et al.*, 1982) are superimposed (dashed lines). The agreement between the model and measurements is best for the shortest delay (10 ms). Masking in the model exceeds psychophysical masking by less than 3 dB at the longer latencies (20 and 40 ms).

IV. DISCUSSION

A. Modeling cochlear micromechanics

The modeling approach described in this paper consolidates the two opposing views of cochlear micromechanics expressed by Allen and Neely (1992). One of these viewpoints emphasized the passive role of the TM in sharpening the tuning of signals delivered to the IHC and argued against

TABLE III. Tone-detection threshold levels (dB SPL). The threshold criterion for the model was specified as having a peak-WNR value that matches a 1-kHz tone at 9 dB SPL. Reference equivalent thresholds for ear-canal sound levels (RETSPL, ANSI) are also listed for comparison.

	0.5 kHz	1 kHz	2 kHz	4 kHz	8 kHz
Model WNR	17	9	11	34	56
RETSPL	12	9	15	13	14

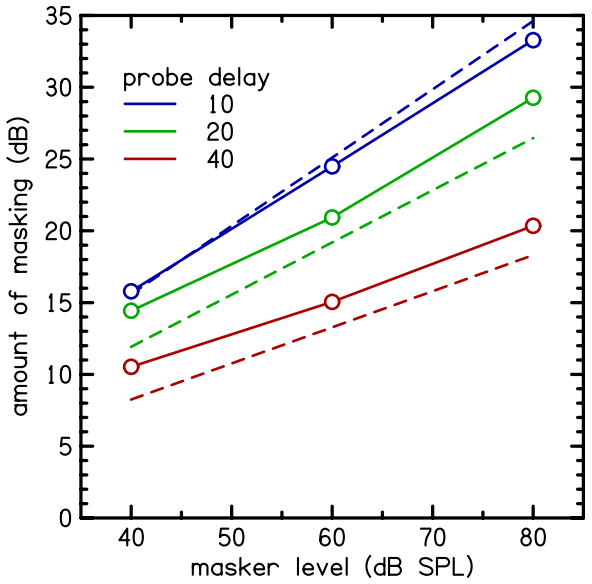

FIG. 11. (Color online) Amount of masking observed as a function of masker level for maskers for three probe delays (10, 20, 40 ms). The masker and probe had the same frequency (1-kHz) and their durations were 250 and 20 ms, respectively. The method for calculating the amount of masking is described in the text. For comparison, the dashed lines are fits to psychophysical data (Jesteadt *et al.*, 1982).

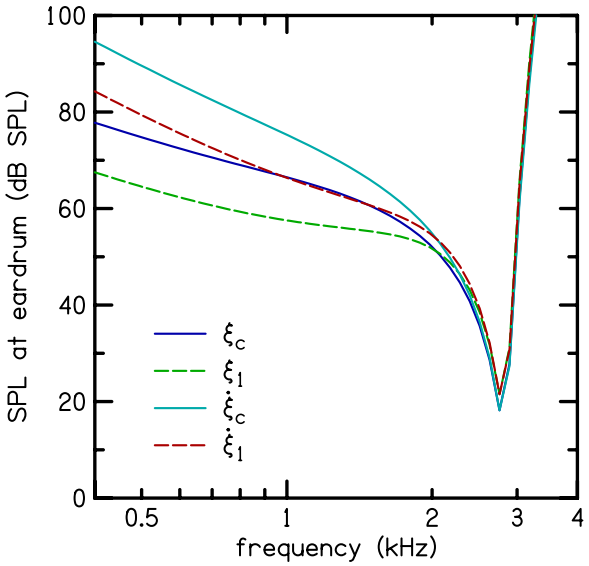

FIG. 12. (Color online) Isodisplacement and isovelocity tuning curves representing ξ_c, ξ_1, $\dot{\xi}_c$, and $\dot{\xi}_1$ at 80% of the distance from the stapes (to the apex). The isodisplacement curves are the same as shown in Fig. 6. IHC excitation in the model is proportional to $\dot{\xi}_c$. The isovelocity criterion was selected to place the tips of the ξ_1 and $\dot{\xi}_1$ tuning curves at the same SPL.

any cochlear amplification that required active elements. The other viewpoint argued the need for active elements (in the form of negative damping) to increase cochlear sensitivity and thereby sharpen tuning. The present model incorporates both the *passive* sharpening due to the TM and the *active* sharpening due to OHC feedback forces. Besides having the conceptual benefit of consolidating opposing viewpoints, the formulation of the nonlinearity in the present model achieves the dual objectives of (1) implementing the requisite amount of dynamic-range compression and (2) being consistent with level-dependence observed in ABR latencies.

The present model results differ from the model results of Allen (1980) in which the effect of H_c was to make ξ_c tuning much sharper at the tip compared to ξ_p tuning. In the present model, the passive sharpening of ξ_c tuning due to the H_c is observable, but it is much less than the active sharpening due to OHC feedback forces. Comparison of the isodisplacement tuning curves (upper panel of Fig. 6) reveals that enhancement of tuning due the transfer function H_c is mainly observed in elevation of the tails of the tuning curves, while only a slight sharpening is seen in the tips of the tuning curves. This feature of the model is reminiscent of similar experimental observation of Narayan *et al.* (1998) based on comparison of AN and BM tuning. However, in the present model, TM–RL shear velocity $\dot{\xi}_c$ was selected as the input to the IHC instead of TM–RL shear displacement ξ_c. Differences between velocity and displacement tuning are illustrated in Fig. 12, which shows tuning curves for representing ξ_c, ξ_1, $\dot{\xi}_c$, and $\dot{\xi}_1$. Comparison of Fig. 12 with Fig. 1(a) of Narayan *et al.* suggests that ξ_1 would have been a better choice for input to the IHC because it lies mostly between BM displacement ξ_1 and BM velocity $\dot{\xi}_1$ similar to the neural tuning curve of Narayan *et al.*

Best results were obtained with the damping coefficient c_4 set to zero and c_3 set to a relatively small value. Setting c_3 so small is probably unrealistic because the motion of the TM relative to the RL is thought to be a major source of damping in cochlear micromechanics. The need to set c_3 to zero in the present model is attributed to the lack of an independent motion of the RL relative to the BM. The multi-chamber model described in the Discussion section could provide a means of testing this hypothesis.

TM damping in the present model is larger than expected from previous work. For example, the TM model suggested by Allen (1980) produces an abrupt positive phase shift of nearly 180° in the transfer function that resembles a phase shift observed in neural data by Kim *et al.* (1979). Furthermore, a phase shift that was observed experimentally by Dong and Olson (2013) between BM displacement and OHC voltage (that might be caused by a TM resonance) is more sharply tuned than the TM resonance in the present model. In both cases, the present model appears to have some features that are qualitatively similar to the experimental data. However, the phase shift of the transfer function in the present model is less abrupt than observed in the experimental data. This dissimilarity can be attributed to the limitations imposed by representing the micromechanics with only two DOFs. Better agreement would be expected with the multi-chamber model described below.

Although the representation of cochlear micromechanics in the present model differs in many details from the recent model of Lamb and Chadwick (2014), the role of the TM is similar. Specifically, we agree with their conclusion that the phase of shear vibrations within CP is crucial to activation of the cochlear amplifier.

The present model is consistent with the recent model reported by Meaud and Grosh (2011) in favoring OHC somatic motility over hair bundle motility as the predominate

source of feedback force that drives the cochlear amplifier. Likewise, the present model is consistent with the suggestion of Meaud and Grosh (2014) for the role of the TM in two-tone suppression because the phenomena of two-tone suppression is generally regarded as a reduction in cochlear amplifier gain, which is influenced by mechanical properties of the TM. Variations in mechanical properties of the TM that have been observed in mutant mice could be simulated in the present model and the effect of these TM variations on two-tone suppression could be explored.

The finding that achieving agreement with the level-dependence observed in ABR latency requires the inclusion of an adaptation stage is not entirely consistent with the suggestion by Neely *et al*. (1988) that the level dependence of ABR Wave-V latency can be attributed entirely to cochlear mechanics. As an alternative to the previous claim, the present model suggests that the level-dependence is mostly due to cochlear mechanics at high frequencies and only partially due to cochlear mechanics at low frequencies. In other words, onset enhancement due to the adaptation stage makes a substantial contribution to the level dependence of WNR latency at low frequencies, but not so much at high frequencies.

Some aspects of the present model are less favorable compared to other models of cochlear mechanics. For example, although the present model can distinguish OHC feedback force as being driven by OHC hair-bundle deflection better than the functional model of Verhulst *et al*. (2012) and may be computationally more efficient, it does not preserve intensity invariance of zero crossings as well as their model. Resolving this issue in the present model deserves further attention.

The present model may have the advantage of computational efficiency over more detailed finite-element models (e.g., Meaud and Grosh, 2010, 2011), but it does not do as well as finite-element models in distinguishing separate motions of micromechanical structures (such as the TM and the RL). Nevertheless, the simplification of OHC feedback forces in the present model may be useful in efforts to make finite-element models more functionally realistic.

The cochlear-response delays in the present model are longer than in previous active cochlear models. Previous attempts to obtain such long delays (e.g., Neely, 1990) have encountered computational instability due to excessive power being delivered to the cochlear fluid at high frequencies. The present formulation of cochlear micromechanics alleviates the long-standing issue of computational instability in cochlear models with negative damping, so provides a useful foundation for future modeling research.

B. Cochlear amplification

The benefit of providing signal amplification in the cochlea prior to detection was first suggested by Gold (1948), but this suggestion was largely ignored until the discovery of OAEs (Kemp, 1978). Shortly after the discovery of OAEs, the term *cochlear amplifier* was coined to describe the high sensitivity and sharp tuning that was observed experimentally in viable cochleae, but was not present *post mortem* (Davis, 1983). The first computer models of the cochlear amplifier (e.g., Neely and Kim, 1983) used negative-damping elements as a mechanism to sharpen tuning and increase sensitivity similar to what was being observed experimentally. The potential of negative-damping elements to produce more signal power than what entered the cochlea through the stapes was recognized; however, it remains an open question, due to lack of empirical evidence, whether the signal power delivered to IHCs ever exceeds the signal power that enters the cochlea from the middle ear.

The presence of negative damping at low levels in the present model predicts that power gain would be observable experimentally in measurements of BM motion. However, recent experimental observations (e.g., van der Heijden and Versteegh, 2015) are incompatible with even as little as 6 dB of power gain in the BM traveling wave. As an alternative to amplification, van der Heijden and Versteegh suggest that OHCs could cause damping to increase at high stimulus levels through a mechanism that they call *braking*. A different alternative, that may be preferable, is that a more detailed model of cochlear micromechanics could confine longitudinal power flow (due to OHC feedback forces) to structures internal to the OC in a manner that would not make this power flow evident in BM motion. In any case, the failure by van der Heijden and Versteegh to observe power gain in BM motion is useful for restricting the scope of acceptable models of cochlear micromechanics. However, their observations do not appear to be conclusive evidence against the existence of cochlear amplification.

To illustrate the influence of OHC feedback forces on cochlear tuning in the present model, the tuning curves shown in Fig. 6 are replotted in Fig. 13 for a passive version of the model that lacks any OHC feedback force. All tuning curves in Fig. 13 are broader and less sensitive. The group-delay curves are relatively constant across frequency and have a smaller maximum. The passive state of the model simulates cochlear responses in the presence of OHC dysfunction. Cochlear tuning approaches passive tuning when stimulus levels are so high that OHC feedback forces become negligible because their amplitude is limited by saturation of OHC transduction.

C. Modeling auditory-system function

Our interest in exploring the limits of the present model regarding forward masking stems from the importance of masking in determining the loudness of sounds.

Prior to implementing the adaptation stage described by Eqs. (7)–(9), we considered the use of other synapse models in combination with the present cochlear model. Specifically, the models of Meddis *et al*. (1990) and Verhulst *et al*. (2016) were evaluated. These two synapse models are similar to each to other in many ways. Both use diffusion from multiple reservoirs to mimic responses of single AN fibers. Both produce sufficient onset enhancement to improve the agreement between model WNR and ABR latencies. However, neither of these synapse models produces more than 3 dB of forward masking for the stimulus conditions shown in Fig. 11. The inability of a single synapse to account for psychophysical

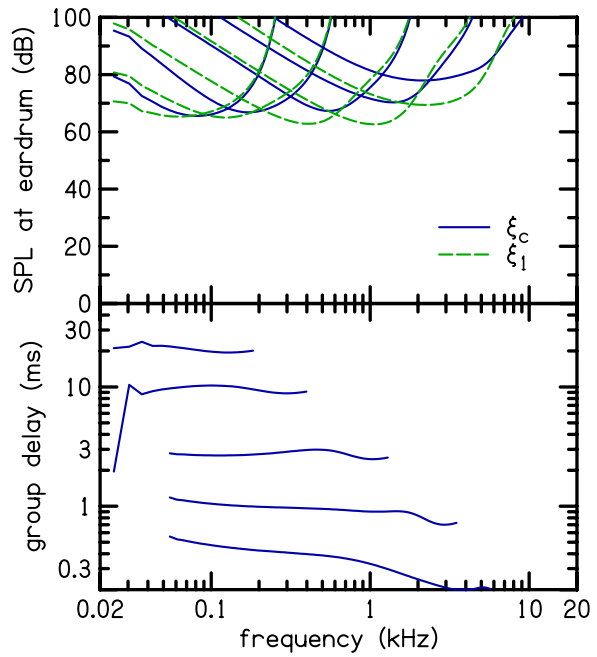

FIG. 13. (Color online) Isodisplacement tuning curves representing ξ_c and ξ_1 at five locations for a passive version of the model. The locations are at 20, 40, 60, 80, and 90% of the distance from the stapes (to the apex). These results were computed from a linear version of the time–domain cochlear model with $\gamma = 0$. Compare with the active version of the model with $\gamma = 1$ shown in Fig. 6.

forward masking is not unexpected. Experimental observations of neural responses (Relkin and Turner, 1988) and previous modeling efforts (Meddis and O'Mard, 2005) also support this conclusion.

Efforts have been made to extend single-synapse models to provide better simulation of psychophysical masking by including complex neural circuits that contain multiple synapses with either coincidence detectors (e.g., Meddis and O'Mard, 2005) or inhibitory connections (e.g., Verhulst *et al.*, 2016). However, if the objective is only to simulate the *function* of forward masking (and not necessarily its associated neural physiology), then the synapse model described by Eqs. (7)–(9) offers a simpler alternative. The key to its effectiveness in producing a greater amount of forward masking (compared to other synapse models) is the parameter α, which raises a variable representing diffusion permeability to the 10th power. Perhaps it is this mechanism that allows a single-reservoir model to achieve amounts of masking that would otherwise require multiple reservoirs. It should be noted, however, raising the diffusion permeability to any power is incompatible with linear diffusion and there is no experimental evidence to support a power of 10.

The peak-to-steady ratio was specified indirectly in our adaptation stage by requiring agreement with psychophysical forward masking. In other words, the same inherent balance between depletion and recovery in the synapse that produces forward masking is also responsible for determining the peak-to-steady ratio. Besides its function in increasing the level dependence of cochlear transient-response latencies at low frequencies, the peak-to-steady ratio may also play a role in the phenomenon known as *overshoot* (e.g., Elliott, 1965; Bacon, 1990), in which a short-duration tone presented simultaneously with a longer-duration masker is easier to detect when its onset is delayed relative to the onset of the masker.

In general, models of cochlear compression agree on exhibiting linear behavior at very low levels and compressive behavior at very high levels. However, there is disagreement among models on the sound level at which the transition from linear to compressive behavior, sometimes called the *compression knee-point* (CK), occurs. In the present model, low-level linear behavior is indicated in Eq. (5) by γ being constant for $\xi_c \leq d_m$. Some functional models of the auditory system place the CK at about 40 dB SPL (e.g., Meddis *et al.*, 2001). Some evidence, such as measurements of BM motion (e.g., Rhode, 1971; Ruggero and Rich, 1991) and stimulus-frequency OAEs (Shera and Zweig, 1993), suggests that the CK occurs at or above 20 dB SPL. Other evidence, such as measurements of ABR latency (Gorga *et al.*, 1988), categorical loudness scaling (Al-Salim *et al.*, 2010), distortion-product OAEs (Neely *et al.*, 2003), and BM measurements (Rhode and Recio, 2000) suggest that the CK occurs at or below 20 dB SPL. Knowing the sound level at which the CK occurs has clinical relevance when fitting hearing aids that feature *wide-dynamic range compression*, which has the objective of restoring normal cochlear compression. The fact that the displacement-gain curve in Fig. 4 has a negative slope at 0 dB SPL indicates that the CK in the present model occurs at or below 0 dB SPL. The weight of evidence from measurements and models appears to favor the view that the compressive range of human cochlear mechanics extends all the way down to the threshold of hearing.

D. OAE

Although the focus of this model development was on simulating ABR toneburst latencies, it is of interest to know how these latencies compare with toneburst otoacoustic emission (TBOAE) latencies. Because the model described in this paper includes ear-canal and middle-ear sections between the sound source and cochlea, TBOAEs may be simulated in the modeled pressure at the eardrum (P_e) following the toneburst stimulus. Unfortunately, this model did not produce TBOAE waveforms with well-defined latencies at all levels and frequencies. In Fig. 14, TBOAE waveforms are shown for responses from tonebursts at four frequencies (0.5, 1, 2, 4 kHz) at the same level (20 dB SPL). The model is the same as for all of the previous model results, except that $\pm 0.01\%$ random variation was added to the partition stiffness along the entire length of the cochlea to simulate roughness in the partition impedance. This roughness increased the OAE amplitude at 4 kHz, but had little effect at lower frequencies. The presence of multiple local maxima in the TBOAE waveforms at all frequencies suggests interference among contributions to the TBOAE that originate from different locations. At higher stimulus levels, the modeled TBOAE waveforms were irregular and did not have well defined peaks, so are not included in the results presented here.

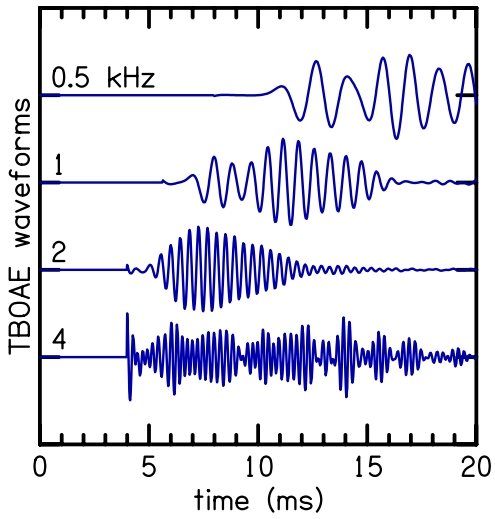

FIG. 14. (Color online) TBOAE waveforms modeled as eardrum pressure following the stimulus. The responses to tonebursts are shown at four frequencies (0.5, 1, 2, 4 kHz) at the same level (20 dB SPL) and with the same durations used for the WNR data shown in Fig. 7. In this figure, the stimulus portion of the eardrum pressure has been set to zero and the pressure following the stimulus has been scaled to the same maximum value in each case.

The TBOAE latencies shown in Fig. 15 are nearly equal to the corresponding WNR latencies at the lower frequencies. This means that the modeled OAE did not encounter any additional traveling-wave delay at the lower frequencies. At the highest stimulus frequency (4 kHz), the TBOAE latency is more than twice the WNR latency, which suggests the presence of an additional travelling-wave delay in the propagation of the return wave toward the stapes. The ratio of OAE latency to WNR latency is 2.3 at 4 kHz and has a mean value of 1.0 at the three lowest frequencies. Comparable latency ratios based on OAE and ABR measurements were 1.33 for low-frequency (0.5–1.4 kHz) tonebursts and 2.33 for high-frequency (2–8 kHz) tonebursts (Rasetshwane *et al.*, 2013).

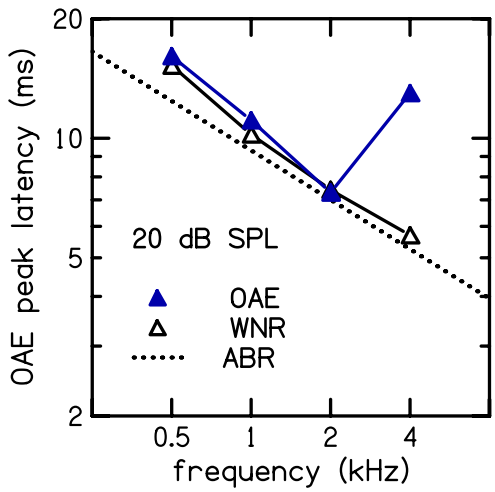

FIG. 15. (Color online) TBOAE latency as a function of frequency. TBOAE latencies were defined as the peak of the TBOAE waveforms shown in Fig. 14. For comparison, the WNR and ABR latencies at the same stimulus level (20 dB SPL) were copied from Fig. 9. Note that TBOAE latency is nearly equal to WNR latency at the lower frequencies (0.5, 1, 2 kHz). At the highest frequency (4 kHz), TBOAE latency is more than twice WNR latency.

These comparisons between TBOAE and WNR latencies should be regarded as preliminary. Further development and exploration of the model are needed before any conclusions are made about TBOAE mechanisms based on this model.

E. Next steps

Perhaps the most disappointing finding of this study is the discrepancy between the RETSPL and WNR thresholds at 8 kHz (see Table III). Adjustments to the present model to reduce this discrepancy are warranted. One reason for the elevated 8-kHz threshold in the model is that (for any given SPL) the amplitude of the IHC hair-bundle deflection is smaller at the 8-kHz characteristic place, compared to the characteristic place for lower frequencies. Another reason is that the excitation pattern at that place is narrower due to its sharper tuning (i.e., $Q_{erb} = 32.3$), which causes fewer IHCs to produce neural signals. The excitation pattern may be unrealistically narrow due to that lack of longitudinal bending-stiffness in the present model (e.g., Wickesberg and Geisler, 1986). Modifications to the present model that may help to reduce the 8-kHz threshold discrepancy include (1) reducing the geometric scale factor b to increase the amplitude of CP point-displacements, (2) adding longitudinal bending-stiffness to broaden the CP excitation pattern, and (3) increasing the stiffness parameter k_4 to increase the amplitude of the OHC feedback-forces. Additionally, better agreement between the model and behavioral measurements might be achieved by defining threshold in the model in a more realistic manner.

The use of WNR to define tone threshold should be reconsidered. A more realistic definition of tone threshold would combine responses from IHCs within a narrow region of the CP (~1 mm) associated with a "critical band" of frequencies. A more useful definition of threshold would incorporate (1) the random nature of neural signals and (2) time-varying factors, such as temporal integration, that cause the random processes involved in tone detection to become non-stationary.

Other cochlear modeling efforts have made progress in recent years toward simulating OAEs (e.g., Liu and Neely, 2010; Verhulst *et al.*, 2012; Moleti *et al.*, 2013). Compatibility of the present model with measurements of OAEs, other than TBOAE latency, has not yet been explored. The model should be capable of producing distortion-product OAEs without any additional elements. Modification of CP stiffness to include small, spatially-random perturbations should enable the model to also simulate stimulus-frequency OAEs. Of particular interest, because of its comprehensive nature, would be the simulation of DPOAE suppression tuning curves such as those described by Gorga *et al.* (2011).

The present study has revealed new capabilities and limitations of the transmission-line approach to cochlear modeling. For some applications, the computational efficiency of transmission-line models makes them more suitable than more detailed models of the cochlea. In addition, the limitation of the micromechanics to only two DOF has conceptual advantages for its simplicity when trying to understand

general principles of cochlear mechanics. Finite-element models of cochlear micromechanics have the advantage of providing more realistic comparisons with experimental measurements (e.g., Soons *et al.*, 2015; Elliott *et al.*, 2013; Meaud and Grosh, 2014). However, finite-element models are computationally expensive and understanding interactions among the large number of elements is more difficult conceptually. An intermediate approach to cochlear modeling would help to bridge the large gap between transmission-line models and finite-element models.

The present model has only a single fluid-filled chamber, which is made to represent both scala vestibuli and scala tympani by assuming symmetry of the two scalae. A more detailed approach could represent five distinct fluid chambers: (1) scala vestibuli, (2) scala tympani, (3) scala media, (4) the spiral sulcus (SS), and (5) the fluid space in the OC between the RL and BM that surrounds the OHCs. The separation between scala media and scala vestibuli is not thought to be important mechanically, but is conceptually important because (1) the fluids in these two chambers differ and (2) the cochlea maintains a voltage difference between them. The SS is important because shear displacement between TM and RL changes its cross-sectional area. Likewise, the OC chamber is important because OHC motility changes its cross-sectional area. This five-chamber model might allow signal power flow through the inner chamber that would not be observable in BM motion, so could help resolve questions about cochlear amplification that transmission lines lack sufficient detail to address. Results of the present study could provide a useful foundation for the development of a five-chamber model of cochlear mechanics.

V. CONCLUSIONS

- The model results indicate that ABR latency originates in cochlear mechanics by the same mechanism that slows energy to propagate in the traveling wave. The level dependence of ABR latency is mainly due to the level dependence of cochlear tuning, but additional level dependence must be attributed to synaptic adaptation to explain observed latencies, especially at low frequencies.
- The simulation of a latency decrease of about 38% per 20 dB increase in stimulus level, which is what has been observed in human ABR Wave-V data, is achievable in an active, nonlinear, transmission-line model of cochlear mechanics when it includes an adaptation stage.
- A dual role is proposed for the TM as both (1) sharpening the tuning of CP responses and (2) controlling the phase of OHC feedback forces in a manner that restricts the activation of cochlear amplification to a limited frequency range.
- Although the power gain associated with the present model is smaller than the amplitude gain, it remains too large at low levels for compatibility with recent observations of BM vibrations (van der Heijden and Versteegh, 2015) that fail to see evidence of significant power amplification. More detailed models of cochlear micromechanics are needed to resolve this issue.
- The simultaneous agreement of the model WNR with both ABR latency and psychophysical forward masking is unprecedented. The simplicity of our adaptation stage compared to other synapse models combined with its effectiveness in producing both onset enhancement and forward masking support consideration of its use in future efforts to model auditory psychophysics. However, its lack of representation of individual synapses could make it inappropriate, in its present form, for modeling phenomena such as synaptopathy (e.g., Verhulst, 2016).

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institutes of Health: R01 DC008318 (S.T.N.).

[1]In the definition of s, $i = \sqrt{-1}$ and $\omega = 2\pi f$ is the radian frequency.

[2]The description of the shape of H_c as a "high-pass filter" may be confusing because the abscissa in Fig. 2 is place and not frequency. However, a plot of H_c versus frequency at one place has a similar appearance.

[3]The model implementation that produced the results presented in this paper used an internal time-step of 2 μs. The specific method used to integrate the model equations across time was suggested by Allen and Sondhi (1979). Model inputs and outputs were each sampled at 50 kHz.

[4]The power-flow calculation is based on an integer number of cycles of the steady-state response of the time–domain model to a tone. At each location along the CP, the velocity and pressure responses are Fourier transformed. Power flow is calculated at each location as half the real part of the complex product of the tone-frequency component of (1) partition velocity and (2) the complex conjugate of pressure difference across the partition.

[5]Power gain is calculated as the ratio of the absorbed power to the total power. The total power is calculated as the sum of the power flow at all locations. The absorbed power is calculated as the negative sum of the power flow at locations where its value is negative. The existence of regions of positive power flow causes the total power to be less than the absorbed power, which makes the power gain greater than one (or greater than zero after power gain is converted to decibels).

[6]The notation peSPL means that the stimulus level is specified as the peak pressure in decibels re 28.28 μPa. In the model, the stimulus pressure was specified at the eardrum.

Allen, J. B. (**1980**). "Cochlear micromechanics—A physical model of transduction," J. Acoust. Soc. Am. **68**(6), 1660–1670.

Allen, J. B., and Neely, S. T. (**1992**). "Micromechanical models of the cochlea," Phys. Today **45**(7), 40–47.

Allen, J. B., and Sondhi, M. M. (**1979**). "Cochlear macromechanics: Time domain solutions," J. Acoust. Soc. Am. **66**(1), 123–132.

Al-Salim, S. C., Kopun, J. G., Neely, S. T., Jesteadt, W., Stiegemann, B., and Gorga, M. P. (**2010**). "Reliability of categorical loudness scaling and its relation to threshold," Ear Hear. **31**(4), 567–578.

ANSI (**2004**). ANSI S3.6-1996, *Specifications for Audiometers* (AIP, New York).

Bacon, S. P. (**1990**). "Effect of masker level on overshoot," J. Acoust. Soc. Am. **88**(2), 698–702.

Brownell, W. E., Bader, C. E., Bertrand, D., and deRibaupierre, Y. (**1985**). "Evoked mechanical responses of isolated cochlear outer hair cells," Science **227**, 194–196.

Dallos, P., and Evans, B. N. (**1995**). "High-frequency motility of outer hair cells and the cochlear amplifier," Science **267**(5206), 2006–2009.

Davis, H. (**1983**). "An active process in cochlear mechanics," Hear. Res. **9**(1), 79–90.

Dong, W., and Olson, E. S. (**2013**). "Detection of cochlear amplification and its activation," Biophys. J. **105**(4), 1067–1078.

Elliott, L. L. (**1965**). "Changes in the simultaneous masked threshold of brief tones," J. Acoust. Soc. Am. **38**, 738–746.

Elliott, S. J., Ni, G., Mace, B. R., and Lineton, B. (**2013**). "A wave finite element analysis of the passive cochlea," J. Acoust. Soc. Am. **133**(3), 1535–1545.

Gold, T. (**1948**). "Hearing. II. The physical basis of the action of the cochlea," Proc. R. Soc. Lond. B: Biol. Sci. **135**(881), 492–498.

Gorga, M. P., Neely, S. T., Kopun, J., and Tan, H. (**2011**). "Distortion-product otoacoustic emission suppression tuning curves in humans," J. Acoust. Soc. Am. **129**(2), 817–827.
Gorga, M. P., Reiland, J. K., Beauchaine, K. A., and Jesteadt, W. (**1988**). "Auditory brainstem responses to tone bursts in normal hearing subjects," J. Speech Hear. Res. **31**, 87–97.
Greenwood, D. D. (**1990**). "A cochlear frequency-position function for several species—29 years later," J. Acoust. Soc. Am. **87**(6), 2592–2605.
Harris, D., and Dallos, P. (**1979**). "Forward masking of auditory nerve fiber responses," J. Neurophys. **42**(4), 1083–1107.
Jesteadt, W., Bacon, S. P., and Lehman, J. R. (**1982**). "Forward masking as a function of frequency, masker level, and signal delay," J. Acoust. Soc. Am. **71**(4), 950–962.
Johnson, S. L., Beurg, M., Marcotti, W., and Fettiplace, R. (**2011**). "Prestin-driven cochlear amplification is not limited by the outer hair cell membrane time constant," Neuron **70**(6), 1143–1154.
Kemp, D. T. (**1978**). "Stimulated acoustic emissions from within the human auditory system," J. Acoust. Soc. Am. **64**(5), 1386–1391.
Kim, D. O., Siegel, J. H., and Molnar, C. E. (**1979**). "Cochlear nonlinear phenomena in two-tone responses," Scand. Aud. Suppl. **8**(9), 63–81.
Lamb, J. S., and Chadwick, R. S. (**2014**). "Phase of shear vibrations within cochlear partition leads to activation of the cochlear amplifier," PLoS One **9**(2), e85969.
Liu, Y. W., and Neely, S. T. (**2010**). "Distortion product emissions from a cochlear model with nonlinear mechanoelectrical transduction in outer hair cells," J. Acoust. Soc. Am. **127**(4), 2420–2432.
Meaud, J., and Grosh, K. (**2010**). "The effect of tectorial membrane and basilar membrane longitudinal coupling in cochlear mechanics," J. Acoust. Soc. Am. **127**(3), 1411–1421.
Meaud, J., and Grosh, K. (**2011**). "Coupling active hair bundle mechanics, fast adaptation, and somatic motility in a cochlear model," Biophys. J. **100**(11), 2576–2585.
Meaud, J., and Grosh, K. (**2014**). "Effect of the attachment of the tectorial membrane on cochlear micromechanics and two-tone suppression," Biophys. J. **106**(6), 1398–1405.
Meddis, R., Hewitt, M. J., and Shackleton, T. M. (**1990**). "Implementation details of a computation model of the inner hair-cell auditory-nerve synapse," J. Acoust. Soc. Am. **87**(4), 1813–1816.
Meddis, R., and O'Mard, L. P. (**2005**). "A computer model of the auditory-nerve response to forward-masking stimuli," J. Acoust. Soc. Am. **117**(6), 3787–3798.
Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (**2001**). "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**(6), 2852–2861.
Moleti, A., Al-Maamury, A. M., Bertaccini, D., Botti, T., and Sisto, R. (**2013**). "Generation place of the long- and short-latency components of transient-evoked otoacoustic emissions in a nonlinear cochlear model," J. Acoust. Soc. Am. **133**(6), 4098–4108.
Narayan, S. S., Temchin, A. N., Recio, A., and Ruggero, M. A. (**1998**). "Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae," Science **282**(5395), 1882–1884.
Neely, S. T. (**1989**). "Transient responses in an active, nonlinear model of cochlear mechanics," in *Cochlear Mechanisms*, edited by J. P. Wilson and D. O. Kemp (Plenum Press, London), pp. 106–113.
Neely, S. T. (**1990**). "Level dependence of the latency of cochlear transients," in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos *et al*. (Springer-Verlag, Berlin), pp. 411–418.
Neely, S. T., Gorga, M. P., and Dorn, P. A. (**2003**). "Cochlear compression estimates from measurements of distortion-product otoacoustic emissions," J. Acoust. Soc. Am. **114**(3), 1499–1507.
Neely, S. T., and Kim, D. O. (**1983**). "An active cochlear model showing sharp tuning and high sensitivity," Hear. Res. **9**(2), 123–130.
Neely, S. T., and Kim, D. O. (**1986**). "A model for active elements in cochlear biomechanics," J. Acoust. Soc. Am. **79**, 1472–1480.
Neely, S. T., Norton, S. J., Gorga, M. P., and Jesteadt, W. (**1988**). "Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli," J. Acoust. Soc. Am. **83**, 652–656.
Puria, S., and Allen, J. B. (**1991**). "A parametric study of cochlear input impedance," J. Acoust. Soc. Am. **89**(1), 287–309.
Rasetshwane, D. M., Argenyi, M., Neely, S. T., Kopun, J. G., and Gorga, M. P. (**2013**). "Latency of tone-burst-evoked auditory brain stem responses and otoacoustic emissions: Level, frequency, and rise-time effects," J. Acoust. Soc. Am. **133**(5), 2803–2817.
Relkin, E. M., and Turner, C. W. (**1988**). "A reexamination of forward masking in the auditory nerve," J. Acoust. Soc. Am. **84**(2), 584–591.
Rhode, W. S. (**1971**). "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique," J. Acoust. Soc. Am. **49**(4B), 1218–1231.
Rhode, W. S., and Recio, A. (**2000**). "Study of mechanical motions in the basal region of the chinchilla cochlea," J. Acoust. Soc. of Am. **107**(6), 3317–3332.
Rønne, F. M., Dau, T., Harte, J., and Elberling, C. (**2012**). "Modeling auditory evoked brainstem responses to transient stimuli," J. Acoust. Soc. Am. **131**(5), 3903–3913.
Ruggero, M. A., and Rich, N. C. (**1991**). "Furosemide alters organ of Corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane," J. Neurosci. **11**(4), 1057–1067.
Shera, C. A., Guinan, J. J., and Oxenham, A. J. (**2002**). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proc. Natl. Acad. Sci. **99**(5), 3318–3323.
Shera, C. A., and Zweig, G. (**1991**). "A symmetry suppresses the cochlear catastrophe," J. Acoust. Soc. Am. **89**(3), 1276–1289.
Shera, C. A., and Zweig, G. (**1993**). "Noninvasive measurement of the cochlear traveling-wave ratio," J. Acoust. Soc. Am. **93**(6), 3333–3352.
Soons, J., Dirckx, J., Steele, C., and Puria, S. (**2015**). "Basilar membrane and reticular lamina motion in a multi-scale finite element model of the mouse cochlea," AIP Conf. Proc. **1703**(1), 050006.
van der Heijden, M., and Versteegh, C. P. (**2015**). "Energy flux in the cochlea: Evidence against power amplification of the traveling wave," J. Assoc. Res. Otolaryngol. **16**(5), 581–597.
Verhulst, S., Dau, T., and Shera, C. A. (**2012**). "Nonlinear time–domain cochlear model for transient stimulation and human otoacoustic emission," J. Acoust. Soc. Am. **132**(6), 3842–3848.
Verhulst, S., Jagadeesh, A., Mauermann, M., and Ernst, F. (**2016**). "Individual differences in auditory brainstem response wave characteristics: Relations to different aspects of peripheral hearing loss," Trends Hear. **20**, 27837052.
Wickesberg, R. E., and Geisler, C. D. (**1986**). "Longitudinal stiffness coupling in a 1-dimensional model of the peripheral ear," in *Peripheral Auditory Mechanisms* (Springer, Berlin-Heidelberg, Germany), pp. 113–120.
Zweig, G., Lipes, R., and Pierce, J. R. (**1976**). "The cochlear compromise," J. Acoust. Soc. Am. **59**(4), 975–982.
Zwislocki, J. J., and Kletsky, E. J. (**1979**). "Tectorial membrane: A possible effect on frequency analysis in the cochlea," Science **204**(4393), 639–641.