A TWO-DIMENSIONAL NONLINEAR NONLOCAL FEED-FORWARD COCHLEAR MODEL AND TIME DOMAIN COMPUTATION OF MULTITONE INTERACTIONS*

YONGSAM KIM† AND JACK XIN‡

Abstract. A two-dimensional cochlear model is presented, which couples the classical second order partial differential equations of the basilar membrane (BM) with a discrete feed-forward (FF) outer hair cell (OHC) model for enhanced sensitivity. The enhancement (gain) factor in the model depends on BM displacement in a nonlinear nonlocal manner in order to capture multifrequency sound interactions and compression effects in a time-dependent simulation. The FF mechanism is based on the longitudinal tilt of the OHCs in feeding the mechanical energy onto the BM. A numerical method of second order accuracy in space and time is formulated by reducing the unknown variables to the BM with the representation of eigenfunction expansions. Though the nonlinear coupling with OHCs created an implicit algebraic problem at each time step, the structure of the FF mass matrix is found to permit a decomposition into a sum of a time-independent symmetric positive definite part and the remaining time-dependent part. A fast iterative method is devised and shown to converge with only the inversion of the time-independent part of the mass matrix. The time-dependent computation is studied by comparing with steady state solutions (frequency domain solutions) in the linear regime and by a convergence study in the nonlinear regime. Results are shown on OHC amplification of BM responses, compressive output for large intensity input, and nonlinear multitone interactions such as tonal suppression, noise suppression, and distortion products. Qualitative agreement with experimental data is observed.

Key words. ear model, nonlinear nonlocal feed-forward, time domain computation, multiscale signal processing

AMS subject classifications. 35Q72, 65M70, 65M12, 65F10

DOI. 10.1137/040612464

1. Introduction. Sound signal processing based on human perception has become an essential component of modern hearing devices and audio compression [27]. A closely related scientific issue is the modeling of nonlinear aspects of the ear's responses to multifrequency input, which helps to understand and extract the most relevant frequencies subsequent to their interactions. Nonlinearities are known to originate in the cochlea (inner ear) and are further modified in higher-level auditory pathways. The cochlea is accessible from classical mechanics and has partial differential equation (PDE) models of various sophistication. Our short-term goal is to develop some of these mechanical models into tools for nonlinear signal processing of complex signals. Mechanical models, though describing only the peripheral hearing process, are mathematically well defined and offer a rich playground and starting point for understanding nonlinear phenomena. Designing an efficient numerical

method with balanced performance and computational cost will be a key consideration in this paper. Our long-term goal is to formulate a mathematical framework at the perception level for sound processing, to go beyond mechanics. The existing models at the perception level are mostly ad hoc [40]. It is, however, hopeful that the mathematics being done at the cochlea level can be further developed to better model human sound perception in the future; see [36] for a work along this line.

Let us describe briefly the sound processing in the cochlea and some related nonlinear phenomena. When sound pressure waves go into the inner ear, the basilar membrane (BM) in the inner ear starts to vibrate. If the input sound contains one frequency (pure tone), BM motion reaches its maximum amplitude at a specific location (so-called characteristic place or CP) when transients die out. The frequency corresponding to the CP is called the characteristic frequency (CF). By mapping each frequency component of a sound input to its CP, the BM plays the role of a frequency analyzer. The experimental data [9, 30, 5] showed that in vivo cochlear dynamics is strongly nonlinear and nonlocal. BM motion exhibits a compressive nonlinearity, in that responses to low-level input are more sensitive and sharply tuned (frequency selective) than to high-level stimuli. Two-tone suppression and distortion products are two kinds of nonlocal phenomena observed in the cochlea. When the input contains two tones (two separate frequencies), the BM response to one tone is suppressed due to the presence of the other tone (so-called two-tone suppression). Under certain conditions, responses to frequencies other than the input frequencies are observed and are called distortion products. Part of the modeling effort is to recover these nonlinear phenomena, which interestingly are also present at the perception level [40].

Cochlear modeling has a long history [1, 3] and is constantly driven by experiments [9, 30, 34]. Solutions to linear models are well studied both analytically and numerically; see [13, 16, 18, 17] and the references therein. The so-called nonlinear active models basically contain two ingredients: one is a model of the BM, the frequency analyzer of the inner ear; and the other is a model of neighboring cell(s), e.g., the outer hair cell (OHC), known to provide the sensitivity of the ear in vivo. Nonlinearity is present in the latter. In this paper, we shall compute an initial boundary value problem for a two–space-dimensional model as such under multifrequency input. The BM part is a classical second order membrane model [24], and the active part is an adaptation of the feed-forward (FF) OHC model recently studied in [21].

The FF micromechanics OHC model [21] takes into account the tilt of the OHCs in the longitudinal direction along the cochlear duct and the resulting force on the BM. In [21], the FF model is coupled with a three-dimensional fourth order plate model of the BM. Among other findings, the coupled model recovers two-tone interactions in agreement with experimental data. Due to the complexity of the model, solutions are approximated by an asymptotic (WKB) method first, and then numerically computed, or by a hybrid WKB-numeric method. The method, however, becomes complicated and unwieldy in the nonlinear regime under multiple frequency input.

When the input sound contains multiple frequencies, an accurate way to resolve the nonlinear responses is through an efficient numerical method in the time domain. Namely, we march the model solutions in time until initial transients die out and persistent quasi-steady states appear; see [8] for an earlier work. The main computational difficulty we encounter in the discretized system of equations is the dependence of the mass matrix on the unknowns (e.g., BM displacement) as a result of the nonlinear coupling in the FF OHC model. The matrix must be inverted at each time step, which is the bottleneck in time domain computation. A direct inversion is slow and inefficient. A careful study of the intrinsic matrix structure reveals that the mass

matrix can be decomposed into a sum of the time-independent symmetric positive definite part and the remaining time-dependent part. This makes possible an iterative method to handle the inversion. Conditions on the convergence of the iteration are found and implemented. A speedup by a factor of 10 is numerically observed. The overall numerical scheme consists of this iterative method in conjunction with an eigenfunction representation to formulate equations entirely on the BM. A second order accurate discretization is used to achieve both efficiency and accuracy. The numerical convergence of the method is verified by comparing with exact steady state solutions in the linear regime and by a grid refinement study in the nonlinear regime.

The ensuing simulations allow us to study stability of the model solutions. It turns out that the local OHC nonlinearity in [21] may generate instabilities in the time-dependent solutions and spoil the desired quasi-steady states at later times. We propose a nonlocal form of the nonlinearity, by a convolution with Gaussian kernel functions. The smoothing property of the nonlocal nonlinearity helps to stabilize the computation and promotes the interaction of near neighbors on the BM. We show that the resulting model is able to capture the major nonlinear phenomena such as tonal suppressions, BM enhancement, and compression.

We remark that nonlocality of BM nonlinearity is reported in a recent finding of both forward and backward active mechanisms in the cochlea [5]. Simplified cochlear (transmission line) models with nonlocal nonlinearities in one space dimension have been utilized for both modeling and sound signal processing [15, 23, 6, 38, 37]. The two-dimensional model in this paper is a step towards developing more accurate PDE-based computational models for nonlinear sound signal processing, in similar spirit to PDE methods in image processing; see [25, 39, 33] and the references therein.

The rest of the paper is organized as follows. In section 2, we show the background and related cochlear models before presenting the model we study. In section 3, we outline main steps of the numerical method. In section 4, we specify model and numerical parameters; study the numerical convergence; and show numerical results on the BM responses under click and other multiple frequency inputs. The conclusions are in section 5. In the appendix, we present the analysis on the convergence of the iterative method in our numerical scheme.

2. Models of the cochlea.

2.1. Classical passive model. The cochlea contains an elastic structure, the BM, which is immersed in an incompressible Stokes fluid [18]. Similar to [18, 24], we simplify the upper cochlea chamber into a two-dimensional rectangle $\Omega = [0, L] \times [0, H]$, with the BM at the bottom $z = 0$; see Figure 1. The lower cochlea chamber is symmetric with respect to $z = 0$. Sound signals generate vibrations on the left boundary $(x, z) \in \{0\} \times [0, H]$, called stapes, which in turn induce fluid motion that goes skew-symmetrically from the upper chamber to the lower chamber through a small hole (helicotrema) at $(L, 0)$. The pressure difference drives the BM into motion (see the dashed line in Figure 1 for a BM displacement during its evolution), where the BM displacement is magnified several orders of magnitude for ease of viewing. The symmetry allows one to ignore details of the helicotrema and concentrate on the upper chamber where the fluid pressure obeys the Laplace equation.

In fact, the BM amplitude is on the order of nanometers (nm), while the domain size is on the order of centimeters (cm); see Table 1 in section 4. This huge difference in length scales makes negligible the effect of BM deformation on the fluid. Hence one can write the fluid pressure as a boundary integral on $\partial\Omega$. Specifically, let p represent the pressure difference across the BM; then it satisfies
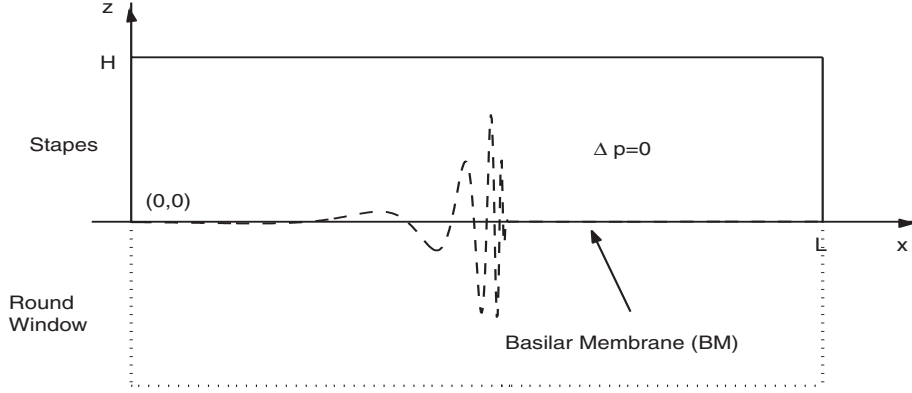
FIG. 1. *Illustration of two-dimensional cochlear model on the BM.*

$$(1) \qquad \Delta p(x, z, t) = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} = 0, \qquad (x, z) \in [0, L] \times [0, H];$$

see Figure 1.

At the boundaries the normal fluid velocity is the same as that of the immediately adjacent wall. The upper wall being rigid, we have

$$(2) \qquad \frac{\partial p}{\partial z}(x, H, t) = 0, \qquad 0 \le x \le L.$$

The apical wall (the right boundary at $x = L$) assumes a Dirichlet boundary condition for simplicity:

$$(3) \qquad p(L, z, t) = 0, \qquad 0 \le z \le H,$$

which is known to have a minimal effect on the interior BM response [24, 35]. A zero Neumann boundary condition, however, offers more insight into the dispersive nature of BM dynamics [35].

The basal wall (at $x = 0$) is moving with the stapes:

$$(4) \qquad \frac{\partial p}{\partial x}(0, z, t) = 2 \rho T_m p_e(t), \qquad 0 \le z \le H,$$

where ρ is the fluid density, $p_e(t)$ is the input sound pressure at the eardrum, and T_m is a bounded linear operator on the space of bounded continuous functions, representing the middle ear filtering. The output depends on the frequency content of $p_e(t)$. If $p_e = \sum_{j=1}^{J_m} A_j \exp(i\omega_j t) + c.c.$, a multitone input, where $c.c.$ denotes the complex conjugate and J_m a positive integer, then $T_m p_e(t) = \sum_{j=1}^{J_m} a_m(\omega_j) A_j \exp(i\omega_j t) + c.c.$, where $a_m(\cdot)$ is the following:

$$(5) \qquad a_m(\omega) = 30 \left(1/30 + 0.0605\omega^2 \left((1 - \omega^2 \omega_m^{-2})^2 + (2\phi_m \omega/\omega_m)^2 \right)^{-0.5} \right),$$

where $\omega_m = 4$ kHz, the middle ear characteristic frequency, and $\phi_m = 0.7$, the middle ear damping ratio. The formula (5) is an empirical fit of the experimental data [12] on the filtering characteristics of the middle ear.

Let BM displacement be $u(x, t)$; the last boundary condition is

$$(6) \qquad \frac{\partial p}{\partial z}(x, 0, t) = 2\rho \, u_{tt}(x, t), \qquad 0 \leq x \leq L.$$

Modeling the BM motion as a spring-mass system with the pressure p as the external force, we write the dynamical equation

$$(7) \qquad p(x, 0, t) = m \, u_{tt} + r \, u_t + s(x) \, u, \qquad 0 \leq x \leq L,$$

where m and r are the mass density of the BM and the damping term, respectively; $s(x)$ is the BM stiffness per unit length in the following form:

$$(8) \qquad s(x) = 4\pi^2 \, m \, (0.456 \exp(4.83 \, (1 - x/L)) - 0.45)^2.$$

The function $s(x)$ is based on the data [11, 19]; see also [6, 38].

For given initial data $(u(x, 0), u_t(x, 0))$, the model is complete as an initial boundary value problem with (1) and (7) subject to the boundary conditions (2)–(4) and (6). Solving the pressure Laplace equation via the Green's function, we see that p is a compact linear functional on u_{tt}. Slaving out the pressure gives a single integral-differential equation on u for the passive model.

We remark that though Table 1 lists the parameters in physical units (cgs), the equations here and hereafter can be nondimensionalized by scaling with characteristic length as 1 cm, time as 1 ms (millesecond), mass as 1 gram, and pressure as 20 μ Pa (air pressure of hearing threshold). The nondimensionalized equations are in the same form. The cgs units have been used directly in the literature [18, 21] and appear convenient in computation and comparison with experimental data.

2.2. FF active model. The passive cochlear model is to be coupled with an OHC model to account for the known sensitivity of a normal ear. We shall adopt the active FF model in [21, 20]. A few rows of OHCs act like piezo-electric actuators that push the BM; see Figure 2. The OHCs can be thought of as microstructures living on the BM. Figure 2 is an enlarged sketch of the so-called micromechanics. In the FF model, the force applied by the OHCs on the BM is assumed to be proportional to the total force acting on the BM. The (total) force acting on the BM is transmitted to the cilia, which in turn acts on the OHCs and back on the BM. The force on the BM is transmitted to the cilia by the stiff arches of Corti, as depicted in the transverse view (a) of Figure 2. The equilibrium position of the arches of Corti is located at the joint with the spiral ligament. So we write

$$(9) \qquad F_{cilia}(x, t) = C_1(x, t) \frac{F_{BM}(x, t)}{2},$$

where F_{cilia} is the force acting on the cilia, F_{BM} is the resultant force acting on the BM, and C_1 is a coefficient. The resultant force acting on the BM consists of those from the fluid in both chambers F_f and the OHCs F_{cell}:

$$(10) \qquad F_{BM}(x, t) = 2F_f(x, t) + F_{cell}(x, t).$$

Due to the longitudinal tilt of the OHCs, the force acting on the cilia at x causes the OHCs to push at a point $x + \Delta$ downstream on the BM, as shown in the longitudinal view (b) of Figure 2, so

$$(11) \qquad F_{cell}(x + \Delta, t) = C_2(x, t) F_{cilia}(x, t),$$
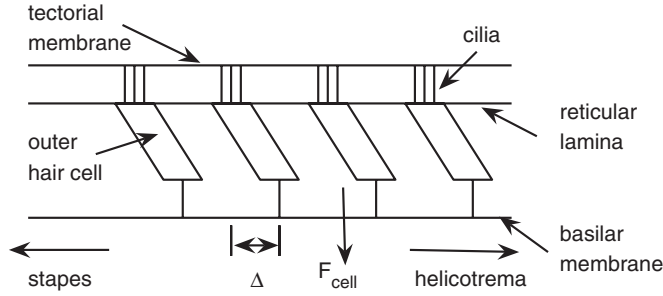
(a) Transverse View
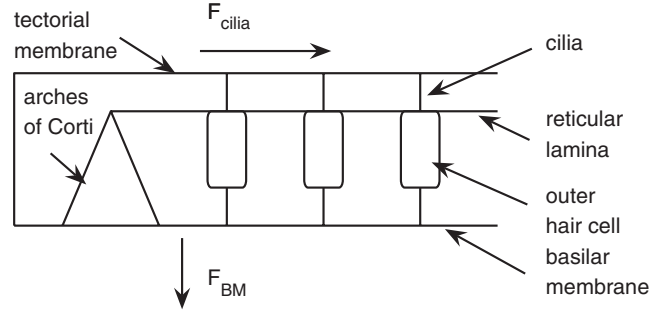
(b) Longitudinal View

FIG. 2. *FF mechanism of OHCs (adapted from* [21]*). Schematic drawings of the transverse (end) view and longitudinal (side) view of the organ of Corti are shown in panels* (a) *and* (b)*, respectively. The force acting on the BM, F_{BM}, is transmitted via the arches to the cilia, F_{cilia}. Due to the longitudinal tilt, the OHC applies a force F_{cell} on the BM at a location downstream of the sensing cilia location separated by a distance Δ.*

where C_2 is a transfer function coefficient relating the two forces, and Δ is the distance generated due to the tilt. Combining (9), (10), and (11) yields

$$F_{cell}(x + \Delta, t) = \frac{C_1(x,t)C_2(x,t)}{2}(2F_f(x,t) + F_{cell}(x,t))$$

(12)
$$= \alpha(x,t)(2F_f(x,t) + F_{cell}(x,t)),$$

where $\alpha(x,t) = C_1 C_2/2$ is referred to as the FF gain factor. The OHC force is known to saturate as the BM motion increases in magnitude [26]. In [21, 20], the gain factor α depends on the BM displacement u, with the profile shown in Figure 3. This nonlinear gain factor remains fairly constant for small BM displacement and tapers off to zero as the BM displacement increases.

2.3. Nonlinear nonlocal FF model. In repeated time domain simulations, we found it problematic, however, to have the gain factor depending pointwise on the BM displacement. An evolving BM profile tends to generate a sawtooth-shaped gain factor $\alpha(x,t)$, which in turns renders the BM profile less smooth and unstable. To fix this problem, we introduce a nonlocal form of the gain factor. Let $g(u)$ be the nonlinear function shown in Figure 3, and the nonlocal nonlinear gain factor $\alpha(x,u,t)$ is

(13)
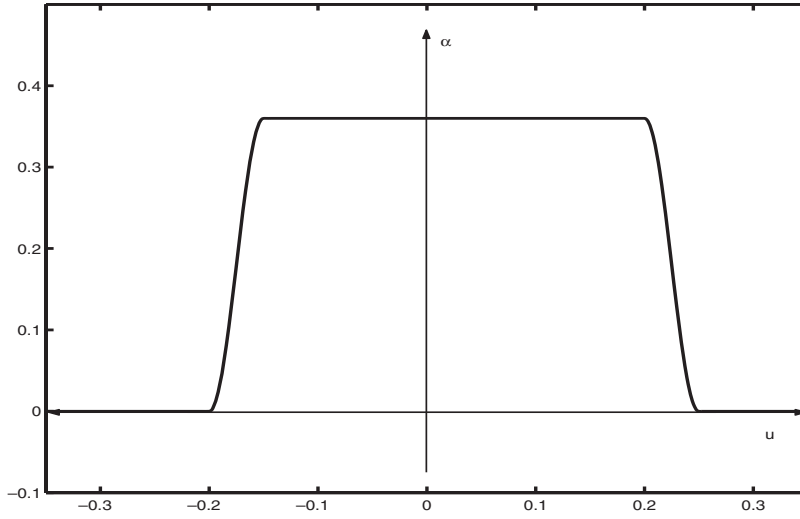$$\alpha(x,u,t) = \frac{\gamma}{\sqrt{\lambda\pi}} \int_0^L \exp(-(x-x')^2/\lambda)\, g(u(x',t))\, dx',$$

FIG. 3. *Variation of the local gain factor α vs. the BM displacement u.*

where γ and λ are constants.

Since $p(x, 0, t) = 2F_f(x, t)$, we let $q(x, t) = F_{cell}(x, t)$ and modify (7) to

$$(14) \qquad q(x, t) + p(x, 0, t) = m\, u_{tt} + r\, u_t + s(x)\, u, \qquad 0 \le x \le L,$$

and combine it with

$$(15) \qquad q(x + \Delta, t) = \alpha(x, u, t)(p(x, 0, t) + q(x, t)), \qquad 0 \le x \le L - \Delta,$$

and $q(x, t) = 0$ for $0 \le x < \Delta$, which accounts for the absence of energy gain from the OHCs in this region due to the tilt of the OHCs.

To summarize, the active nonlinear nonlocal system consists of (13)–(15) and (1)–(6). We shall numerically solve an initial boundary value problem for the system in the next section.

3. Numerical method. In this section we discuss a second order discretization method. The basic idea is to write an integral representation of pressure p in terms of boundary data involving u_{tt} and then work with the remaining equations on the BM.

We first make a change of variable so that the pressure boundary condition at $x = 0$ becomes homogeneous. Let $r(x, z, t) \equiv 2\rho\, T_m\, p_e(t)\, (L - x)$, and write $p' = p + r$. Since r is linear, p' satisfies a Laplace equation with homogeneous boundary data except on $z = 0$. Writing p for p' as a notation, we have

$$(16) \qquad \Delta p(x, z, t) = 0, \qquad (x, z) \in [0, L] \times [0, H],$$

$$(17) \qquad p(L, z, t) = 0, \qquad \frac{\partial p}{\partial x}(0, z, t) = 0, \qquad \frac{\partial p}{\partial z}(x, H, t) = 0,$$

$$(18) \qquad \frac{\partial p}{\partial z}(x, 0, t) = 2\rho\, u_{tt}(x, t),$$

$$(19) \qquad q(x, t) + p(x, 0, t) = m\, u_{tt} + r\, u_t + s(x)\, u + 2\rho\, T_m p_e(t)\, (L - x),$$

$$(20) \qquad q(x + \Delta, t) = \alpha(x, u, t)(p(x, 0, t) + q(x, t) - 2\rho\, T_m p_e(t)\, (L - x)).$$

By the method of separation of variables, $p(x, z, t)$ has the expansion

$$(21) \qquad p(x, z, t) = \sum_{n=1}^{\infty} Q^n(t) \cosh(\xi_n(H - z)) \, e^n(x),$$

where $e^n(x) = \sqrt{\frac{2}{L}} \cos(\xi_n x)$, $\xi_n = \frac{(n-1/2)\pi}{L}$, and $Q^n(t)$ is the coefficient. The $p(x, z, t)$ in (21) satisfies the Laplace equation and the boundary conditions in (17). Expanding $u_{tt}(x, t)$ in basis functions $e^n(x)$'s, we have from (18)

$$(22) \qquad -\xi_n Q^n(t) \sinh(\xi_n H) = 2 \rho u_{tt}^n(t),$$

where $u_{tt}^n(t)$ is the nth coefficient of the series expansion of $u_{tt}(x, t)$. Let us substitute $Q^n(t)$ into (21), and set $z = 0$; the pressure on the BM equals

$$(23) \qquad p(x, 0, t) = \sum_{n=1}^{\infty} \frac{-2 \rho e^n(x)}{\xi_n \tanh(\xi_n H)} u_{tt}^n(t).$$

Substituting $p(x, 0, t)$ into (19) and (20), we get a closed system of equations for the BM displacement $u(x, t)$ and the force $q(x, t)$.

Let us turn to discretization and the associated numerical method. Suppose there are $N + 1$ nodes on the BM. At fixed time t, $u(x_i, t)$ represents $u(i\Delta x, t)$ for $i = 0, \ldots, N$, where $\Delta x = L/N$. By the trapezoidal rule, we have

$$(24) \qquad u_{tt}^n(t) = \sum_{j=0}^{N} u_{tt}(x_j, t) \, e^n(x_j) \, d(j) \, \Delta x \, + \, O((\Delta x)^2),$$

where $d(j)$ equals 1 on the interior points and $1/2$ at the end points.

Now we approximate the series in (23) by the M term partial sum. Let $U(t)$ and $P(t)$ be column vectors representing $u(x_j, t)$ and $p(x_j, 0, t)$, respectively, where $j = 0, \ldots, N$. Setting $x = x_i$ in (23), and plugging (24) into (23), we obtain a relation $P(t) = A \, U_{tt}(t)$, where A is an $(N + 1) \times (N + 1)$ matrix with the following (i, j)th element:

$$(25) \qquad A_{i,j} = \sum_{n=1}^{M} \frac{-2\rho \, e^n(x_i) \, e^n(x_j)}{\xi_n \tanh(\xi_n H)} \, d(j) \, \Delta x$$

for $i = 0, \ldots, N$ and $j = 0, \ldots, N$. Hereafter 0 can be a column or row label for notational convenience.

For simplicity, select $\Delta = K \Delta x$, where $K < N$ is an integer. Then (20) gives

$$(26) \qquad q(x_i + K\Delta x, t) - \alpha(x_i, u, t) \, q(x_i, 0, t) = \alpha(x_i, u, t) \, (p(x_i, 0, t) - F(x_i, t)),$$

where $F(x, t) = 2\rho T_m p_e(t) (L - x)$. Let $Q(t)$ denote the column vector for $q(x_i, t)$; it follows that $B \, Q(t) = C \, (P(t) - F(t))$. The matrix B has 1's on its diagonal, and $B(i + K, i) = -\alpha(x_i, u, t)$ for $i = 0, \ldots, N - K$; otherwise, its elements equal zero. The matrix C is a lower triangular matrix with 0 as its entries other than $C(i + K, i) = \alpha(x_i, u, t)$ for $i = 0, \ldots, N - K$. The column vector $F(t)$ from the forcing term has the ith element $2\rho T_m p_T(t) (L - x_i)$. Since B is invertible, (20) in matrix form is $Q(t) = B^{-1} C \, (P(t) - F(t))$.

Define the gain matrix G as $B^{-1}C + Id$, where Id is the identity matrix. By $P(t) = AU_{tt}(t)$, we express (19) as the discretized equation for U:

$$(27) \qquad (G(U)A - M)U_{tt} = RU_t + SU + G(U)F,$$

where the matrices M, R, and S are diagonal matrices with their diagonal entries m, r, and $S(i,i) = s(x_i)$, respectively.

For the time discretization, we use the second order backward differencing formula [14]. Define $V(t) = U_t(t)$, which forms a first order system with $(G(U)A - M)V_t = RV + SU + G(U)F$. Let U^n and V^n denote $U(n\Delta t)$ and $V(n\Delta t)$, respectively, with Δt the time step; then the two-step method reads

$$(28) \qquad U^{n+2} = \frac{4}{3}U^{n+1} - \frac{1}{3}U^n + \frac{2\Delta t}{3}V^{n+2},$$

(29)
$$V^{n+2} = \frac{4}{3}V^{n+1} - \frac{1}{3}V^n + \frac{2\Delta t}{3}(G(U^{n+2})A - M)^{-1}(RV^{n+2} + SU^{n+2} + G(U^{n+2})F^{n+2}),$$

where the inverse $(G(U^{n+2})A - M)^{-1}$ is proved in the appendix. Plugging (28) into (29), we get the equation for V^{n+2}:

$$(G(U^{n+2})A - \tilde{M})V^{n+2} = (G(U^{n+2})A - M)\left(\frac{4}{3}V^{n+1} - \frac{1}{3}V^n\right)$$
$$(30) \qquad\qquad + \frac{2\Delta t}{3}\left(S\left(\frac{4}{3}U^{n+1} - \frac{\Delta t}{3}U^n\right) + G(U^{n+2})F^{n+2}\right),$$

where $\tilde{M} = M + \frac{2\Delta t}{3}R + (\frac{2\Delta t}{3})^2 S$.

To solve (30) for V^{n+2}, we treat matrix G explicitly; namely, we approximate the U^{n+2} variable in G by $U^{n+1} + \Delta t V^{n+1} + \frac{\Delta t}{2}(V^{n+1} - V^n)$, the second order Taylor expansion of U^{n+2}. Then we solve (30) for V^{n+2} and (28) for U^{n+2}. The above two-step method is initiated by a backward Euler method.

System (30) is solved by an iterative method [32] at each time step of evolution; see the appendix for details of the method and analysis. The iterative method avoids calculating the inverse of the entire *mass matrix* $(G(U^{n+2})A - \tilde{M})$ at each time step. The idea is to decompose the mass matrix into time-dependent and time-independent parts. In the iterative procedure, only the time-independent part needs to be inverted, and we have to calculate its inversion only once in the first time step. A further analysis will be carried out in the appendix on the matrix structures induced by the OHC coupling that allow such a decomposition. We estimate the spectral radius of the iteration matrix which is the multiplication of the inverse of the time-independent part and the time-dependent part. It depends on the order of the gain factor α and the time duration Δt and can be made less than 1 to guarantee the fast convergence of the iteration. The iterative method solves (30) nearly 10 times faster than a standard direct solver in our numerical experiments.

4. Numerical results. In this section we present numerical results from the model and the numerical method in the previous two sections. We first introduce the model and numerical parameters used in the simulations and verify that the numerical method is convergent with the second order accuracy. We show that the OHC active mechanism induces the sharply tuned responses of the BM (sensitivity) and the nonlinear phenomena such as compressive responses, multiple tonal suppressions, and distortion products.

Parameters	Symbol	Magnitude	Unit
Membrane density	m	0.07	g/cm^2
Fluid density	ρ	1	g/cm^3
Length of cochlea	L	3.5	cm
Height of cochlea	H	0.1	cm
Resistance	r	0.07	g/(cm$^2 \cdot$ ms)
Meshwidth	Δx	0.005	cm
Time duration	Δt	0.0025	ms

4.1. Model and numerical parameters. We tabulate model and numerical parameters for our simulations in Table 1 (ms short for millisecond). For the number of terms of the series in (25), we choose $M = 700$ so that higher order truncation makes no appreciable difference in the results. The OHC tilt Δ is 0.005 cm, the same as the meshwidth Δx. For the gain factor α, we consider two different cases: in the linear case, $\alpha(x, u, t) = 0.36$; and, in the nonlinear case, the gain factor is as described in (13) with $\gamma = 0.36$ and $\lambda = 0.12$ cm^2. Note that λ is chosen to guarantee that BM motion is smooth and that γ is fixed to generate a realistic amount of amplification. As γ increases, so does the gain factor α generally, which means that BM responses are more amplified.

We remark that the length L and height H in Table 1 are close to those of the human cochlea; however, OHC tilt Δ and BM resistance are chosen to match with experimental findings [30]. Likewise, our middle ear and BM stiffness parameters are close to cat data [12, 19]. On the other hand, our model and solution method are robust under parameter variations. A knowledge-based parameter adjustment may lead to better fit with certain targeted experimental data. Yet we do not pursue this task here.

The input pressure wave is

$$(31) \qquad T_m p_e(t) = (1 - \mathrm{e}^{-t/t_0}) \sum_{j=1}^{J_m} a_m(\omega_j)\, A_j \sin(\omega_j t),$$

where $t_0 = 2.0$ ms is chosen so that (31) approaches a steady wave after a 12 ms transient period. The initial BM displacement and velocity are both zero.

If $t_0 = 0^+$, an undesirable dispersive tail often appears in the BM displacement, which lasts for some time and makes convergence to steady states slow. The gradually increasing input (31) is a common remedy. For a recent approach using nonzero initial data to suppress the dispersive tail and speed up temporal convergence to steady states, see [35].

We observed that, after the BM response went through an initial transient period, it reached a steady oscillatory motion between 20 and 40 ms regardless of the input frequency and the type of stimulus (e.g., noise). The maximum of 40 ms duration is equivalent to 16000 time steps in our choice of parameters. This translates into an upper limit of 25 minutes for a simulation to complete on an Athlon XP (1616 MHz) processor. Higher frequencies take less time.

4.2. Convergence test for the numerical method. We verify that the numerical method in section 3 is a convergent scheme. Consider the linear model with

constant gain factor in which the steady states reduce to the so-called frequency do-
main solutions below. For a pure tone $p_e = A\,e^{i\omega t}$, the steady response of the cochlea
takes the form (known as *frequency domain solutions*)

$$(32) \qquad u(x,t) = U(x)\,e^{i\omega t}, \qquad q(x,t) = Q(x)\,e^{i\omega t}, \qquad p(x,z,t) = P(x,z)\,e^{i\omega t}.$$

This leads to a Laplace equation on P with the boundary conditions

$$(33) \qquad P(L,z) = 0, \qquad \frac{\partial P}{\partial x}(0,z) = 2\,\rho\,a_m(\omega)\,A, \qquad \frac{\partial P}{\partial z}(x,H) = 0,$$

$$(34) \qquad \frac{\partial P}{\partial z}(x,0) = -2\,\rho\,\omega^2\,U(x),$$

along with

$$(35) \qquad Q(x) + P(x,0) = (-m\,\omega^2 + i\,r\,\omega + s(x))U(x),$$

$$(36) \qquad Q(x+\Delta) = \alpha(P(x,0) + Q(x)).$$

We numerically solve for (P,U) by the standard second order finite difference
discretization and LU decomposition. The first row of Figure 4 shows the profiles
of the imaginary part of the BM complex displacement $U(x)$. The input wave is
4.5 kHz (kilo-Hertz) at 40 decibel (dB) in sound pressure level (SPL), with 20 μ Pa as
reference. The two columns are for the constant gain factors: 0 (left) and 0.36 (right).

For comparison, the profile of the time domain solution $u(x,t)$ at $t = 28$ ms is
drawn in the second row for each gain factor. We see that both methods give similar
BM profiles. The third row shows the BM displacement at a BM point as a function
of time. The BM point is chosen at the location of the maximum BM motion, which
is called the characteristic place (CP), corresponding uniquely to the input frequency.
For a given CP, the corresponding frequency to generate a maximum displacement
peak at the CP is called the characteristic frequency (CF). Here CF = 4.5 kHz
has its CP around $x = 1.73$ cm. The panels in the bottom row represent the BM
displacements in time at CF = 4.5 kHz for the two gain factors. We see that, as the
time proceeds, the time domain method achieves a steady oscillatory pattern. The
gain factor at 0.36 enhances the BM displacement magnitude about 100 times higher
than that without gain, realizing the usual difference of 40 dB between impaired and
normal ears [9, 21].

Now we demonstrate the convergence of the time domain solutions. Choose
$N = M = 175 \times k$, $k = 1, 2, 4, 8$, and 16. For each k, vary the time step as
$\Delta t = 4 \times 10^{-3}/k$ ms. That is, as we refine the meshwidth Δx by a factor of 2, the
time step Δt goes down by the same factor. The convergence test is done with the
nonlinear OHC gain factor. We fix $\gamma = 0.36$ and $\lambda = 0.12$ cm for the gain factor and
6 kHz at 50 dB as the input stimulus. Because we want to treat the same problem
under grid refinement, we fix the OHC tilt $\Delta = L/175$. This value is different from
the one in simulations discussed in later sections, where $\Delta = \Delta x = 0.0025$ cm.

Table 2 shows the convergence ratios of the BM displacement and velocity. Since
we do not know the exact solutions of the problem, one convergence ratio needs three
numerical solutions for three consecutive N's. Let u_N and v_N be the BM displacement
and the velocity, respectively, at the chosen time $t = 24$ ms, and let $||\cdot||_2$ and $||\cdot||_\infty$
be L_2 and L_∞ norms, respectively. Table 2 illustrates that our numerical method
for the time domain solution has the convergence ratio roughly equal to 4, implying
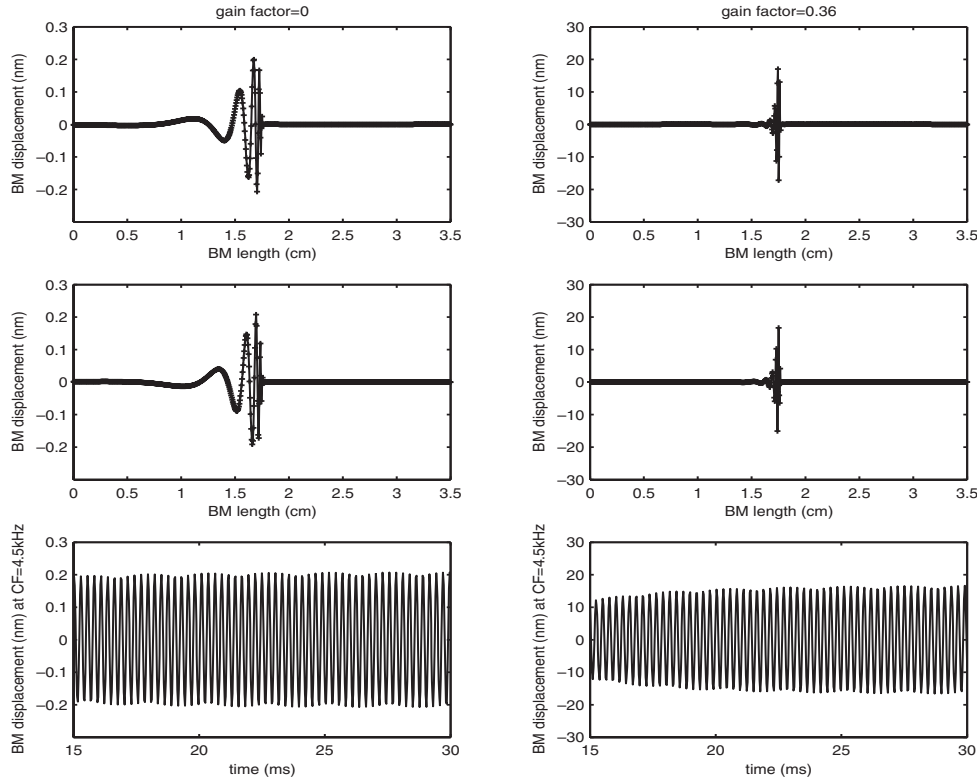second order accuracy of the method.

FIG. 4. *The comparison between the steady state and the time domain solutions. The first two rows show the profiles of the BM displacement: the imaginary part of $U(x)$ (the first row) and $u(x,t)$ at $t = 28$ ms (the second row). The input stimulus 4.5 kHz at 40 dB is the same for all the plots, and the two columns are for the constant gain factors: 0 (left) and 0.36 (right). Both methods give similar profiles of the BM displacement. The third row shows as a function of time the BM displacement at $CF = 4.5$ kHz whose CP is around $x = 1.73$ cm from the basal end. As the time proceeds, the time domain solution achieves a steady oscillatory pattern. The gain factor 0.36 has increased the BM displacement nearly 100 times.*

TABLE 2
Convergence ratios in L_2 and L_∞ of the BM displacement and velocity.

N	$\dfrac{\|u_N - u_{2N}\|_2}{\|u_{2N} - u_{4N}\|_2}$	$\dfrac{\|u_N - u_{2N}\|_\infty}{\|u_{2N} - u_{4N}\|_\infty}$	$\dfrac{\|v_N - v_{2N}\|_2}{\|v_{2N} - v_{4N}\|_2}$	$\dfrac{\|v_N - v_{2N}\|_\infty}{\|v_{2N} - v_{4N}\|_\infty}$
175	3.7079	3.5832	3.7393	3.9525
350	3.4950	3.8903	3.8571	3.4950
700	3.8029	3.7844	3.8166	3.8029

4.3. BM response to single frequency tones. We show that our model with the nonlinear gain factor exhibits amplification and compressive nonlinearity due to the saturation effect. Experimental data have shown that responses of the ear in vivo are more sensitive and sharply tuned to low-level input.

In the literature, experiments have usually been conducted close to the basal (left end) and apical (right end) sites [29, 30, 31], due to difficulties in probing in vivo cochlea. In the computational study, we shall observe several sites, including the basal and apical sites of the BM to show comparable results. Those sites (CPs) are
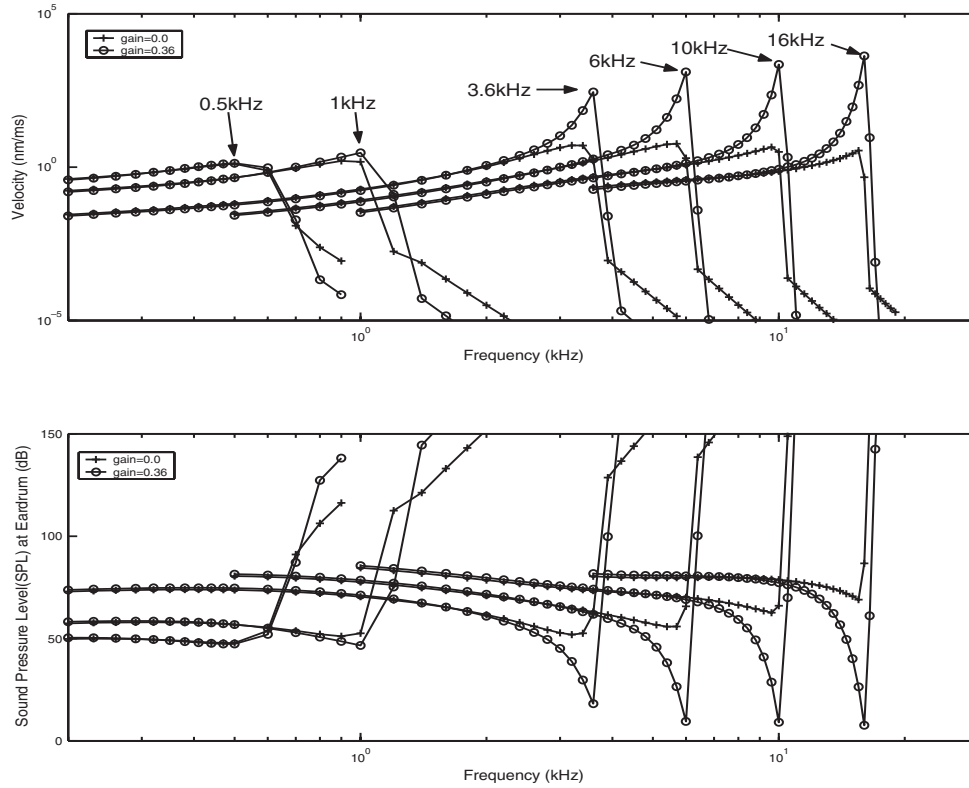
FIG. 5. *Isointensity curves (top) and isodisplacement curves (bottom) for the frequency domain solution without and with a gain factor. The isointensity curve is the variation of the BM velocity (or displacement) as a function of stimulus frequency and intensity. The BM velocities at the six chosen CPs are presented as the frequency of the input wave varies. The isodisplacement curves are computed with 1 nm threshold BM displacement. The cases with the OHC gain are sharply tuned and more sensitive around each CP. The sensitivity gets higher as each CP goes to the basal end, or has more gain at higher frequencies.*

$x = 0.875, 1.21, 1.56, 1.89, 2.54$, and 2.84 cm, respectively, with the corresponding CFs: 16, 10, 6, 3.6, 1.0, and 0.5 kHz.

The top of Figure 5 contains six pairs of curves illustrating the steady state BM velocity as a function of input frequency variable f on the log scale at fixed input intensity of 50 dB. These curves are the so-called isointensity curves, each pair at two different gain factor values: 0.0 and 0.36. Let the modulus of the BM velocity be $|v| = |v|(x; A_{in}, f_{in})$, where (A_{in}, f_{in}) is the input amplitude and frequency and x the BM length variable. The time variable t disappears as it is in the phase of v. The BM frequency to location map is

$$x = cp(f) = 3.5 \cdot (1 - (4.83)^{-1} \cdot \log((f + 0.45)/0.456)).$$

Let $x_i = cp(f_i)$, and let f_i, $i = 1{:}6$, attain values $(0.5, 1, 3.6, 6, 10, 16)$ kHz. The isointensity curves plot $|v|$ vs. f_{in} with $x = x_i$, $A_{in} = 50$ dB. The results in this figure are obtained with the frequency domain solutions. As we are in the linear regime, the change of input intensity level causes only a vertical shift. We see from the isointensity curves that the FF OHC mechanism enhances the BM sensitivity by

raising and sharpening the peaks, similar to the results in [20, 21]. The sensitivity or sharpness grows as the input frequency gets higher. When we compare the active case ($\alpha = 0.36$) with the passive one ($\alpha = 0$), the gain amount is 1.5 times at 1 kHz, 600 times at 6.0 kHz, and 8×10^4 times at 16.0 kHz. Interestingly, hearing loss often occurs near high frequency range. The loss can be further modeled for specific individuals by making α into a function of x to be determined from measurement data (such as an audiogram).

The curves in the bottom of Figure 5 are called isodisplacement curves. They are obtained as follows. Let the modulus of the BM displacement be $|u| = |u|(x; A_{in}, f_{in})$. Set a threshold value for the BM displacement, here 1 nm. Then we find A_{in} so that $|u|(x; A_{in}, f_{in}) = 1$ nm for each $x_i = cp(f_i)$, $i = $ 1:6; plot A_{in} as a function of f_{in} on the log scale. The profile of the BM displacement (Figure 4) is asymmetric about the peak (CP): the absolute displacement increases slowly up to each CP and decays rapidly to zero beyond each CP. Thus it takes a higher input level to stimulate a point on the right of each CP to 1 nm than to stimulate an equally distanced point on the left of each CP to 1 nm. So an isodisplacement curve has a minimum at each f_i, rising sharply over $f > f_i$ and gradually over $f < f_i$. With the gain factor equal to 0.36, these curves have lower and sharper tips, indicating higher sensitivity. The plot is similar to the isodisplacement curves shown in [30, 31, 24, 38].

With the nonlinear gain, we expect not only the sharply tuned response to low-level inputs but also the saturation at high input levels. Hereafter the steady state BM responses are measured at long enough times (from 20 ms on high frequency input to 40 ms on low frequency input) so that the amplitudes have settled and transients die out. In order to see both, we present two types of isointensity curves in Figure 6: The left panels represent the velocity of BM responses to tones as a function of frequency, and the right panels show the sensitivity function defined as the BM displacement divided by the stimulus pressure at the ear drum [30]. The amplitudes of velocity and displacement of the BM are recorded at CF = 1 kHz (top), 6 kHz (middle), and 16 kHz (bottom), respectively, when the frequency and intensity level of the input stimulus vary. If responses grew linearly to the input level, the family of graphs in the left panels would be the translation from one graph, and the sensitivity curves in the right measured at various intensities would be identical. Near each CP, however, the growth of the response to the increasing stimulus level is compressive, and the sensitivity grows systematically larger as the stimulus level decreases. Sensitivity also grows with increasing frequencies. The curves nearly superimpose at frequencies away from each CF. Besides the compressive nonlinearity, we can observe the asymmetrical decrease away from each peak point as well. Figure 6 is similar to Figures 4 and 5 in [30].

Figure 7 shows the isodisplacement curves for several CFs. In addition to the CFs used for Figure 5 (linear case), two more CFs, 2 and 3 kHz, are added. The curves are obtained at a large enough time in the same way as the graphs in the bottom of Figure 5. They also show almost the same pattern as those in Figure 5 (the case of a nonzero gain factor), except that here the overall (y-) value of each curve is larger than before. This is due to the saturation of the active mechanism which generally prevents the full gain allowed when the linear gain is used.

4.4. Transient response. In order to see some of the characteristic nonlinear behavior of the BM, the model is used to simulate the BM responses due to click inputs of various levels at the stapes. A click with a profile in the top panel of Figure 8 is applied at the stapes in a time interval of 10 ms. The subsequent plots show the
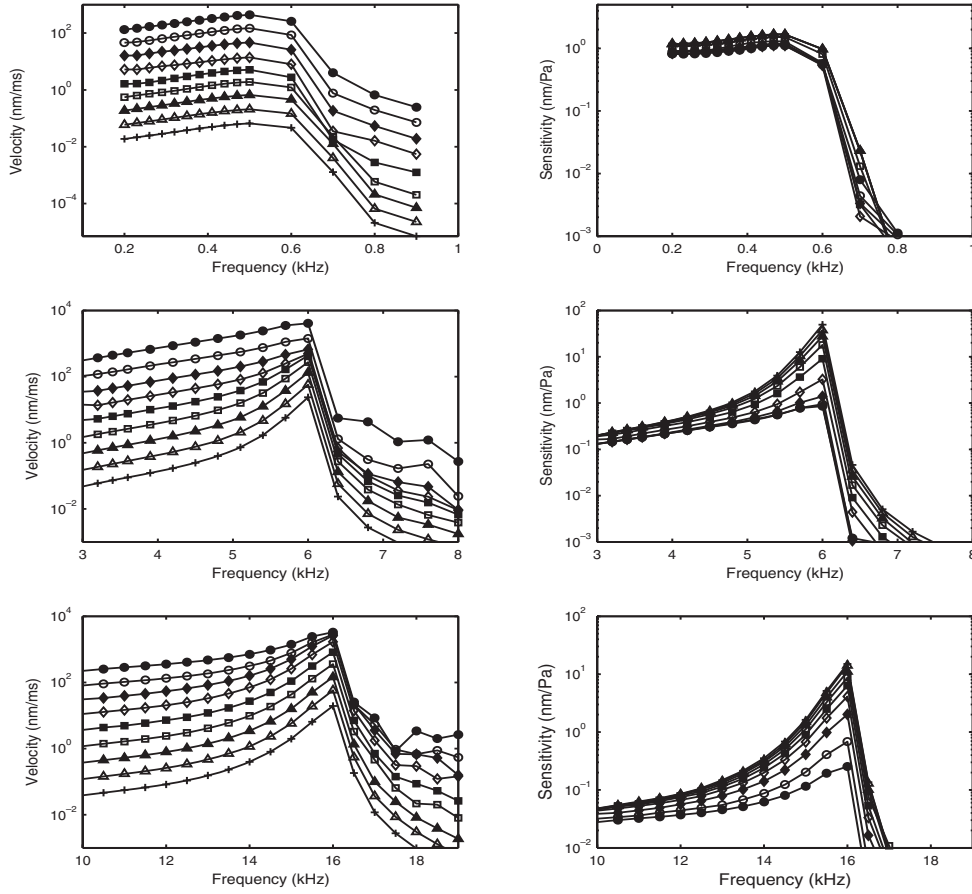
FIG. 6. *Families of isointensity curves; left panels represent the velocity of BM responses to tones as a function of frequency, and right ones show the sensitivity of BM responses. The sensitivity is defined as the BM displacement at each CF divided by the stimulus pressure level at the ear drum. The intensities of the input are -+ = 10, -△ = 20, -▲ = 30, -□ = 40, -■ = 50, -◇ = 60, -♦ = 70, -○ = 80, and -● = 90 in unit dB SPL, decibel in sound pressure level. The velocity and displacement of the BM are recorded at CF = 1 kHz (top), 6 kHz (middle), and 16 kHz (bottom), respectively, when the frequency and intensity of the input stimulus vary. We see the compressive nonlinearity near each CP and the asymmetrical decrease of the BM response away from each peak point.*

transient responses of the velocity of the BM at the point 1.21 cm (CF = 10 kHz) from the stapes. The intensity levels of 20, 60, and 80 dB SPL are used. The BM responses are normalized by input intensities, and they are also known as sensitivity. Similar to the single frequency response, the BM transient responses show a higher amplification at the low stimulus level. The nonlinear behavior of the BM is also reflected by the differences in the shape of the transient responses. At low stimulus level (20 dB here), there is usually a second lobe after the first response envelope decreases to zero. The sustained "ringing" effect in the BM response is clearer at low input level. This is because of the saturation of the OHC force. The same phenomenon is also observed in experiments by Recio et al. [28] and by Lim and Steele [22] using a different solution method.
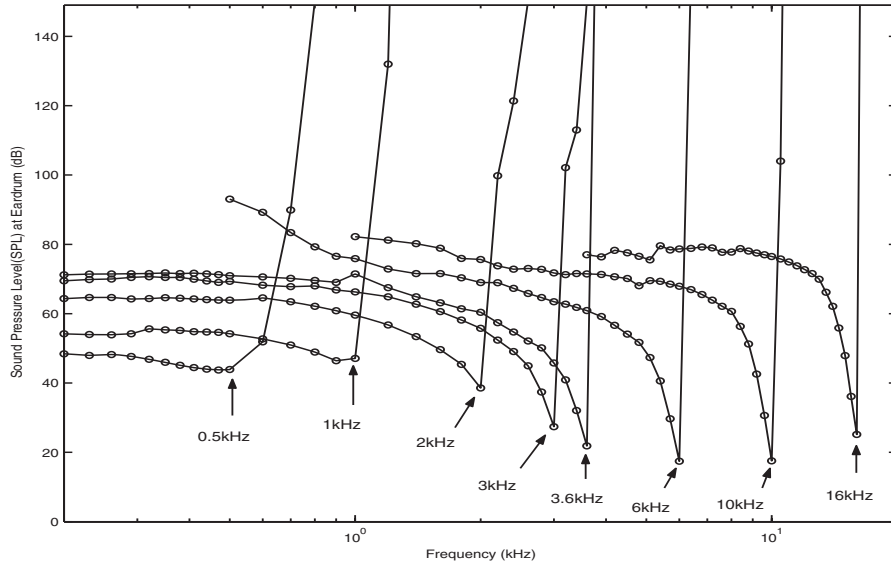
FIG. 7. *Isodisplacement curves for the time domain solution with the nonlinear gain factor. The isodisplacement curves are computed with 1 nm BM displacement as threshold. The isodisplacement curves resemble those in Figure* 5.
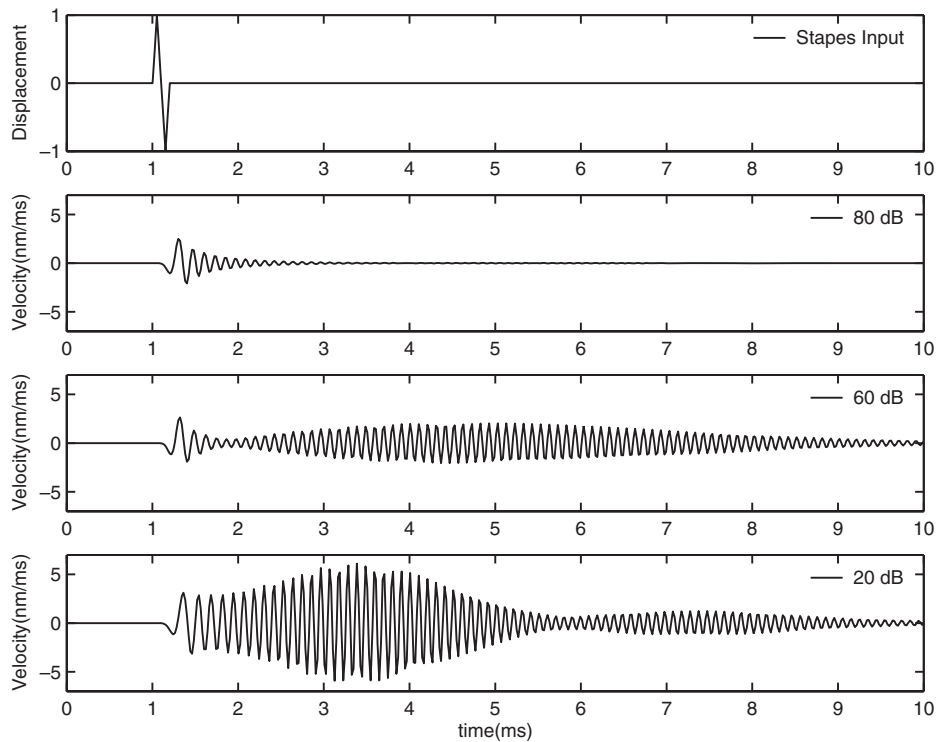

FIG. 8. *Transient response of click. The top panel shows the input click of duration* 0.2 *ms in a time interval of* 10 *ms at the stapes. The BM response (velocity), at the location with CF* 10 *kHz, is given in the subsequent plots for input levels* 20, 60, *and* 80 *dB (normalized by the input level). A lower-level stimulus gives a higher response and longer "ringing" than a higher-level stimulus does.*
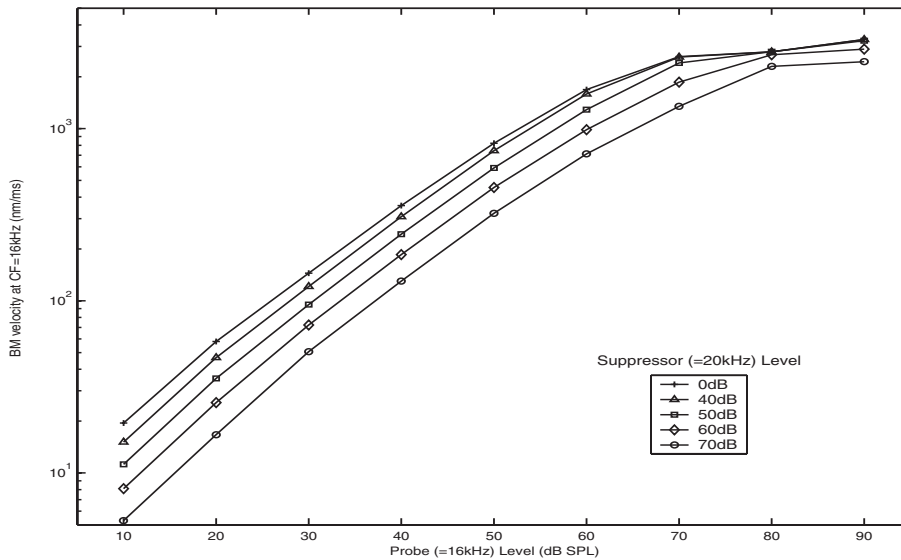
FIG. 9. *Two-tone suppression: The velocity-intensity functions for BM responses to a CF =
16 kHz probe tone presented simultaneously with a 20 kHz suppressor tone. The x-axis indicates the
intensity of the probe tone, and the y-axis is the velocity at the CP of the probe tone. Each graph has
a different suppressor level, as shown in the legend. As the input level of the suppressor increases,
the responses to the CF tones are reduced. The reduction becomes weaker at the high enough level
CF tones.*

4.5. Two-tone interactions. In the nonlinear regime, BM responses to tone
pairs often include interactions between the tones such as suppressions and distor-
tions [30]. Both have been observed in the early stage of the hearing process (cochlea)
and are well documented in the literature [2, 6, 9, 30, 31]. Here we illustrate that the
time domain model with the nonlinear active mechanism gives similar results.

Two-tone suppression consists of the reduction of the response to one tone by
the simultaneous presence of another. Figure 9 illustrates the suppression of the
BM response to a CF (= 16 kHz) tone by a suppressor tone (= 20 kHz) at various
intensities. The x-axis indicates the level of the CF tone (probe) from 10 dB to 90 dB
SPL, and each graph has a different suppressor level. The velocities at the CP of the
probe tone are recorded. In the absence of the suppressor the response to the CF tone
grows at compressive rates, especially at the high intensity level. As the suppressor
level increases, the responses to the CF tones are reduced; however, the reduction
becomes weaker at the high enough levels of CF tones. This result is comparable to
the experimental data; see Figure 16 in [31].

In addition to the high side suppression shown in Figure 9, the low side suppression
is also known to exist. Figure 10 compares the two suppressions. The probe CF =
6 kHz tone which has various input levels from 10 dB to 90 dB (x-axis) is presented
with the suppressors of the input frequencies 10 kHz (high side) and 3.4 kHz (low
side) at three different levels: 30 dB, 50 dB, and 70 dB, respectively. The CP of
6 kHz has the same distance from CPs of both 10 and 3.4 kHz. We record the
velocity at CF = 6 kHz, discrete Fourier transform it, and plot the velocity component
corresponding to 6 kHz. While both panels show the suppression of the probe tone
due to the high and low side suppressor tones, the low side suppression is bigger than
that of the high side one, which is a well-known fact [9, 38]. Also the graphs are
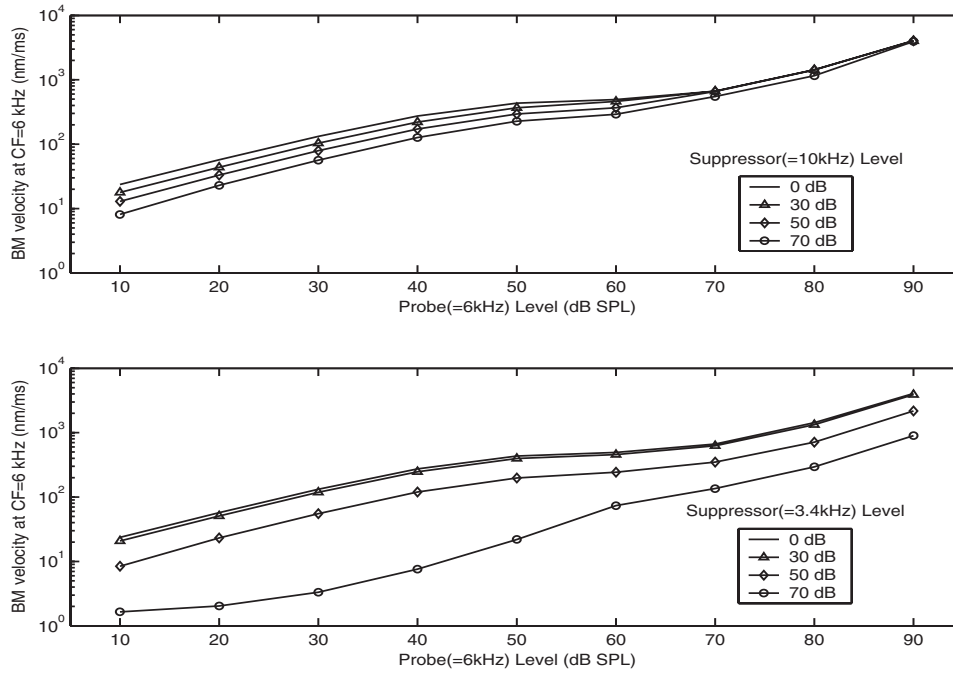
FIG. 10. *Two-tone suppressions with a suppressor of higher (top) and lower (bottom) frequencies than the probe tone, respectively. The velocities of the BM response to the probe tone 6 kHz are saved when the tone is interacting with each suppressor. When the probe level varies from 10 dB to 90 dB, the suppressor level is chosen at 30 dB, 50 dB, and 70 dB in both the low and high side suppressions. The velocity component corresponding to 6 kHz is plotted. Both panels show the suppression of the probe tone, with the low side reduction larger than from the high side.*

comparable to the result in Figures 10 and 11 of [21].

When two or more tones are presented simultaneously, humans can hear additional tones not present in the acoustic stimulus. For two-tone stimuli, denoted as F_1 and F_2 ($F_2 > F_1$), these so-called distortion products have pitches corresponding to linear combinations of the primary tones, such as F_2-F_1, $2F_1$-F_2, and $2F_2$-F_1. The largest, most easily perceived distortion product among them occurs at a frequency equal to $2F_1 - F_2$; this is known as the cubic difference tone (CDT) [2, 31]. Figure 11 shows the CDTs in the BM vibrations. The primary tones are 10 kHz and 8 kHz, both at 80 dB. The top shows the BM displacements at CF = 10 kHz (left) and CF = 6 kHz (right) in terms of time. The bottom presents the spectrum of the waves shown in the top. At CF = 10 kHz we can see the frequency of the CDT as well as the frequencies of the primary tones. At CF = 6 kHz, which is the frequency of the CDT, while we cannot observe the frequency of the primary tones, the displacement of the BM is clearly observed at the CP corresponding to the CDT frequency, and the frequency of the BM displacement is the same as the CDT frequency. Since the BM displacement at CF = 6 kHz is above the critical threshold (= 1 nm, for example), the inner hair cell, and hence the auditory nerve fibers, may detect it, and humans may hear the sound with the CDT frequency. CDT signals are used to test ear functions of newborn babies for early abnormality detection [9].

Figure 12 shows the CDT responses when the input levels of two primary tones are increased simultaneously and by equal amounts. The two primary tones F_1 and F_2
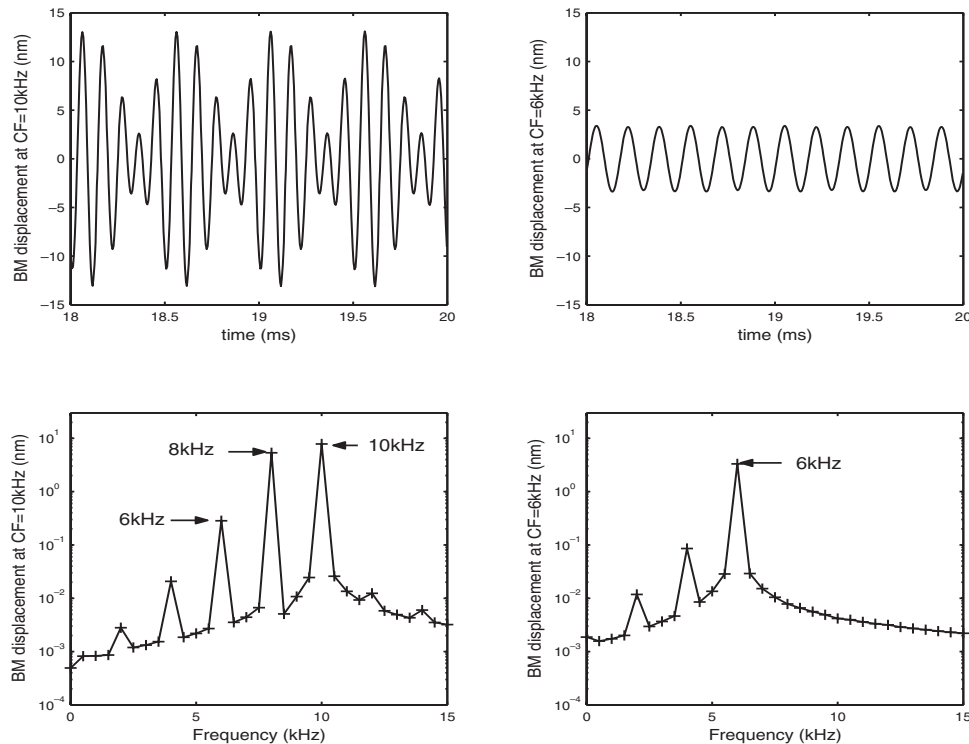
FIG. 11. *CDTs in BM vibrations. The primary tones are 10 kHz and 8 kHz, both at 80 dB. The top shows the BM displacements at CF = 10 kHz (left) and 6 kHz (right) in terms of time. The bottom presents the spectrum of the time-dependent functions on the top. At CF = 10 kHz we see the frequency of the CDT as well as the frequencies of the primary tones. The site at CF = 6 kHz, a CDT frequency, has a nonlinearly generated displacement which may be perceived by humans.*

change with $2F_1$-F_2 equal to a fixed CDT frequency and F_1 approaching the CDT frequency. The top has CDT CF = 0.5 kHz, and the bottom has CF = 6 kHz. The x-axis represents the level of the primary tones, and the y-axis represents the BM responses at CF = CDT. Each graph has different frequencies of the primary tones, as shown in the legend. From the results in both panels, we recover two facts observed in experiments [2, 31]: (1) As the levels of the two primary tones increase, the BM displacement at CF = CDT as a distortion product also increases. However, the distortion product levels relative to the levels of the primary tones decrease; see the compressive increase of the graphs. (2) The magnitude of the CDT on the BM grows with the decreasing frequency ratio F_2/F_1; i.e., the closer the CDT CF is to the primary tones, the bigger is the response of the CDT CF. For comparison, see Figure 4 in [2].

4.6. Multitone interactions. The advantage of the time domain approach is that the time and complication involved in the computation do not depend on the number of the input tones. The input can be any sound signal. We show here a few examples of responses of the BM to more than two tones. The number of tones makes little difference in the time domain computation in terms of computational costs. From the results of the two-tone interactions, we can expect the suppressions and distortion products from more than three-tone interactions [7, 29, 38]. Figure 13
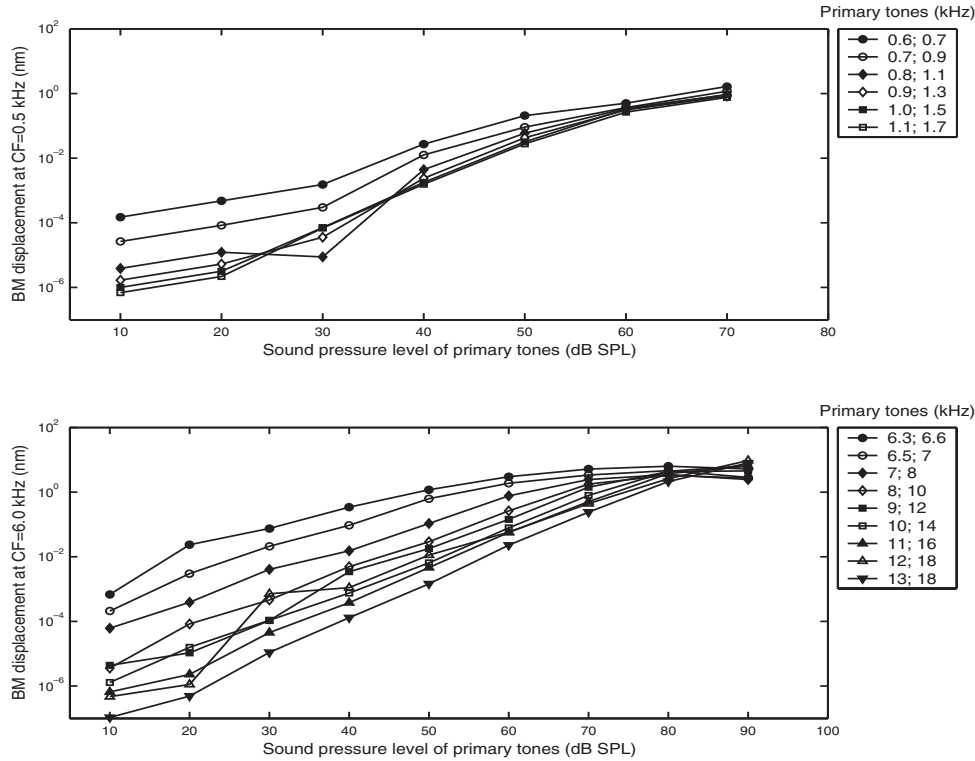
FIG. 12. *CDTs in BM vibrations. The top has CDT CF = 0.5 kHz, and the bottom has CF = 6 kHz. The x-axis represents the level of the primary tones, and the y-axis represents the BM responses at CF = CDT. Each graph has different primary tones F_1 and F_2 which change with $2F_1$-F_2 equal to a fixed CDT frequency and F_1 approaching the CDT frequency. As the levels of the two primary tones increase, the BM displacement at CF = CDT increases with a compressive rate. The magnitude of the CDT on the BM grows with decreasing frequency ratios F_2/F_1; i.e., the closer the CDT CF is to the primary tones, the larger is the response of the CDT CF.*

illustrates the BM displacement (nm) along the BM length at a fixed time 30 ms and the envelope of the BM response when there are five frequencies in the input stimulus. The envelope is obtained by taking the absolute maximum value of Hilbert transform (MATLAB tool) of the BM displacement over time. We see five peaks due to the five-tone stimulus from the figure.

Now we illustrate the suppression of three-tone interactions. Tow tones are fixed at 3.6 kHz and 4.4 kHz, both at 55 dB. The third tone varies in frequency and in magnitude. In Figure 14, we plot the velocities of the BM response of two tones (3.6 kHz, plus line; 4.4 kHz, circle line) as the third tone (line) increases its level from 10 dB to 70 dB. The suppressor (third) tone has three different frequencies; 3.2 kHz (top), 4.0 kHz (middle), and 4.2 kHz (bottom). Here we do not use the real velocity itself for the plots. Instead we compute the frequency components using FFT and draw only the velocity component corresponding to each probe tone. The reason for this treatment is to achieve more accuracy in identifying responses at desired frequencies. The figure shows that the responses of the two tones decrease with the increasing amplitude of the third (suppressor) tone, especially when the level of the suppressor tone is larger than 60 dB. The figure is qualitatively comparable to Figure 3 in [7] and Figure 4.6 in [38].
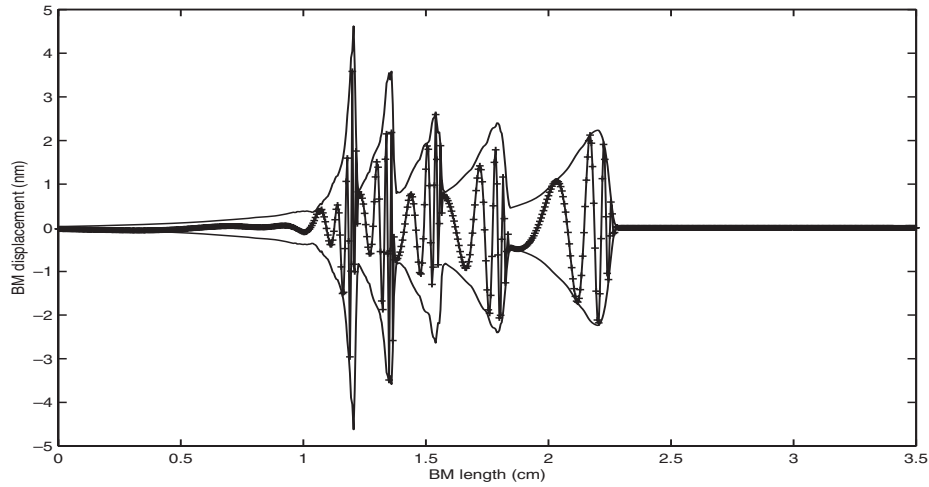
FIG. 13. *BM displacement (nm) along the BM length at a fixed time 30 ms and the envelope of the BM response when the input stimulus is a sum of tones at five frequencies. The envelope is obtained by taking the absolute maximum value of Hilbert transform (MATLAB tool) of the BM displacement over time. Five envelope peaks appear due to the five-frequency stimulus.*
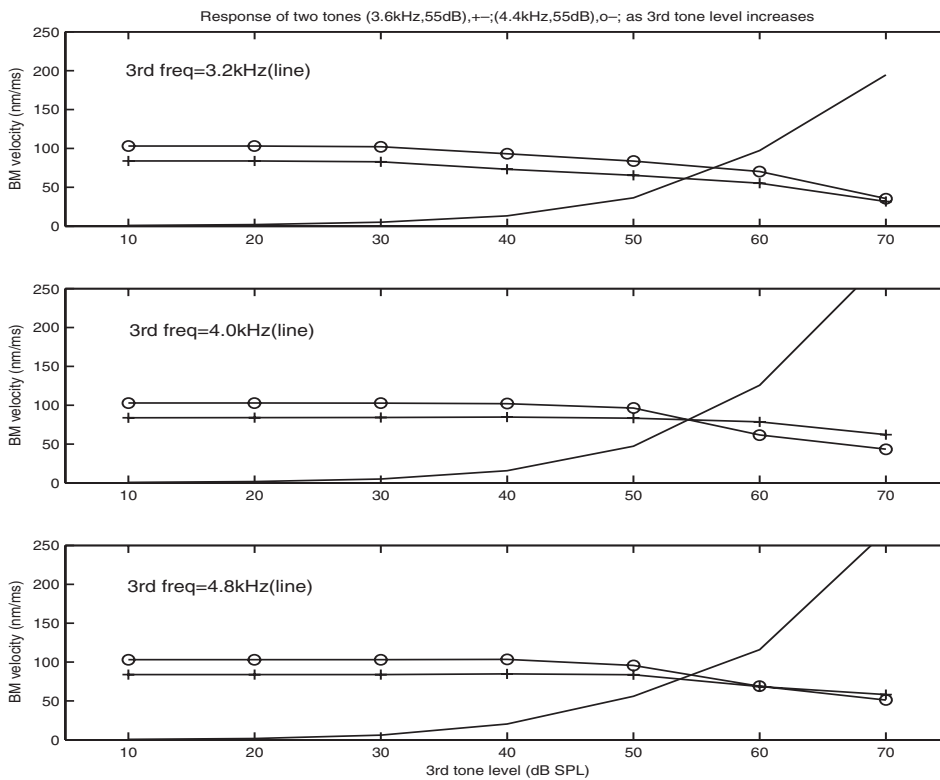

FIG. 14. *Velocities of the BM response of two tones (3.6 kHz, plus line; 4.4 kHz, circle line) as the third tone (line) increases its level from 10 dB to 70 dB. The third tone has three different frequencies; 3.2 kHz (top), 4.0 kHz (middle), and 4.8 kHz (bottom). The responses of the two tones decrease with the increasing amplitude of the third tone, especially when the level of the suppressor tone is above 60 dB.*
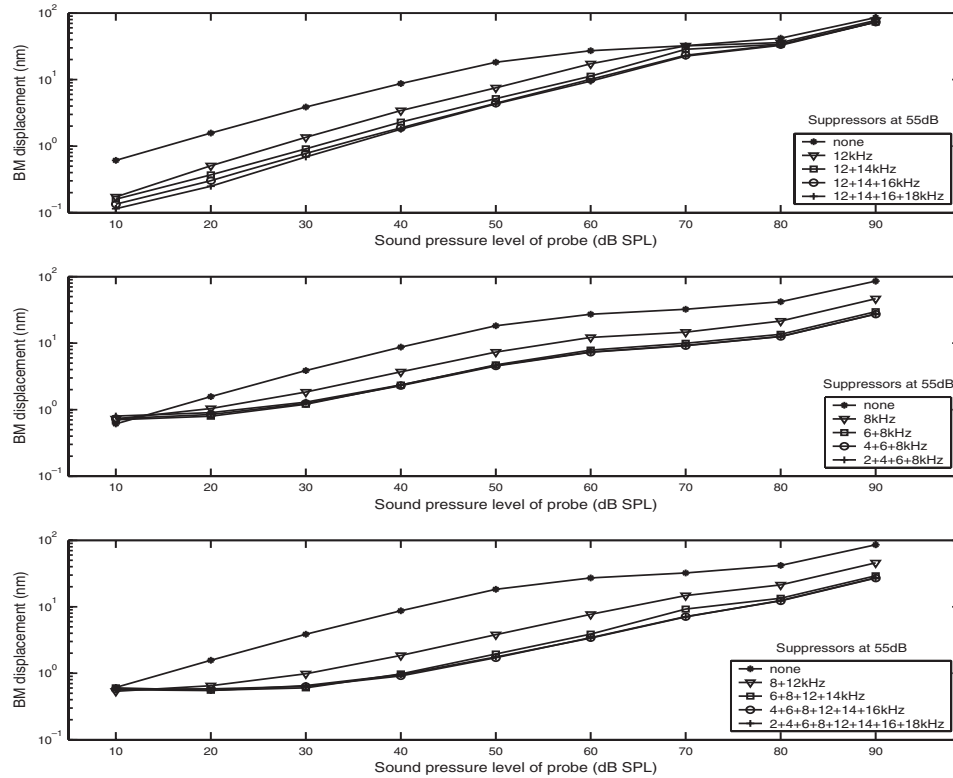
FIG. 15. *Suppression of a tone due to multitone suppressors. The probe tone is* 10 *kHz with the level from* 10 *dB to* 90 *dB. The number of suppressors, all at* 55 *dB, increases in each figure: the suppressors are* 12*,* $(12, 14)$*,* $(12, 14, 16)$*, and* $(12, 14, 16, 18)$ *kHz in the top,* 8*,* $(6, 8)$*,* $(4, 6, 8)$*, and* $(2, 4, 6, 8)$ *kHz in the middle, and* $(8, 12)$*,* $(6, 8, 12, 14)$*,* $(4, 6, 8, 12, 14, 16)$*, and* $(2, 4, 6, 8, 12, 14, 16, 18)$ *kHz in the bottom. When the number of the suppressor tones increases, the probe tone has more suppression up to a critical number of the suppressors. Beyond the critical number, however, the suppression rate may decrease.*

Figure 15 represents the suppression of a tone due to multitone suppressors. The probe tone is 10 kHz with the level from 10 dB to 90 dB. The number of suppressors, all at the 55 dB level, increases in each figure: the suppressors are 12, $(12, 14)$, $(12, 14, 16)$, and $(12, 14, 16, 18)$ kHz in the top; 8, $(6, 8)$, $(4, 6, 8)$, and $(2, 4, 6, 8)$ kHz in the middle; $(8, 12)$, $(6, 8, 12, 14)$, $(4, 6, 8, 12, 14, 16)$, and $(2, 4, 6, 8, 12, 14, 16, 18)$ kHz in the bottom. The frequencies of the suppressors are all higher in the top and lower in the middle than those of the probe tone. The graphs in the top and middle are similar to those in Figure 10 which show the high and low side suppressions. In all three panels, when the number of the suppressor tones increases, the probe tone goes through more suppression up to some number of the suppressors. Beyond the critical number of the suppressors, however, the suppression rate may decrease. The BM response to the multitone input contains distortion products. We omit the details, as the results are similar to those shown in Figure 11.

4.7. Tone and noise interactions. The mutual suppression (e.g., simultaneous masking) of tone and noise is well documented in the psychoacoustic literature [40] though seldom investigated in the study of the cochlea. We show that our model
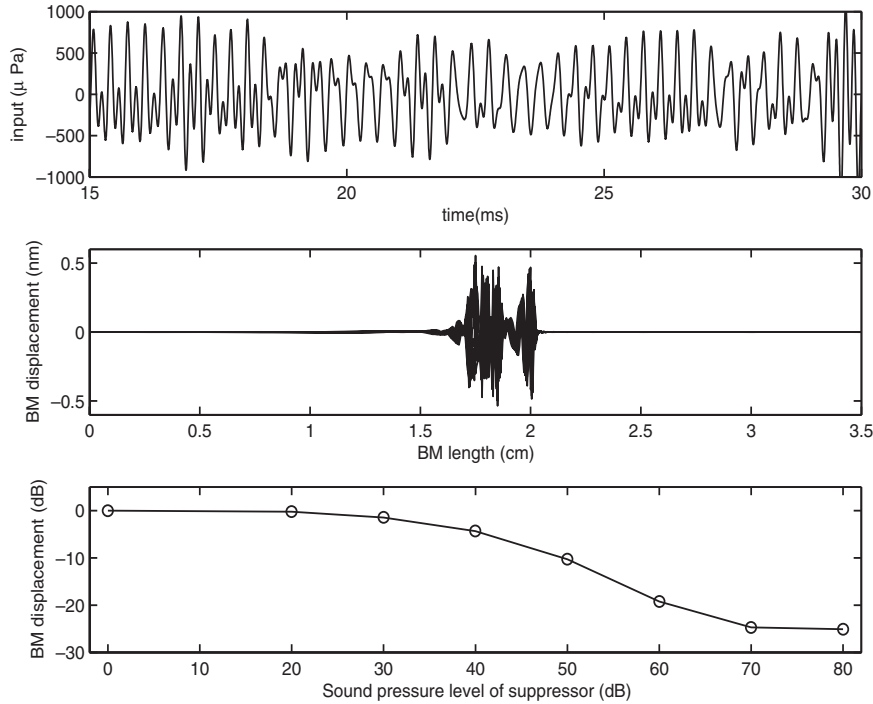
FIG. 16. *Noise suppression by a single tone: The top panel shows the input sound which is the sum of a tone of 3 kHz and a banded noise with center frequency at 4.2 kHz and bandwidth 0.84 kHz, both at 20 dB with respect to 20 μ Pa. The middle panel plots the overlapped BM responses at five equally spaced time slices in the interval $[29.875, 30]$ ms, from the input signal in the top panel. The peaks appear around the two locations corresponding to 4.2 kHz and 3 kHz. The bottom panel shows the reduction of the noise as the intensity of the tone increases.*

captures such a suppression phenomenon. The result demonstrates the advantage of our time-dependent solution method and the model's capability to process arbitrary sound input in a consistent manner with existing data [40].

The signal processing procedure is as follows. Let $u_S(x,t)$ and $u_N(x,t)$ be the BM responses of a pure tone S and a noise N, respectively; and let $u_{S+N}(x,t)$ represent the BM response when the input is the sum of tone S and noise N. The top panel of Figure 16 shows a segment of the input signal when S is 3 kHz pure tone and N is a banded noise with its center frequency at 4.2 kHz and bandwidth 0.84 kHz, both at 20 dB (with reference dB value equal to 20 μ Pa). The S and N are one critical band apart. The middle panel is the overlapped BM responses at five equally spaced time slices in the interval $[29.875, 30]$ ms, to illustrate the dynamics of BM vibration patterns. Two groups of peaks appear around the BM locations corresponding to 4.2 kHz and 3 kHz. The left one contains multiple peaks in the response envelope as a result of the banded noise input.

We calculate the root mean square (rms) of the difference of two BM responses $u_{S+N}(x,t) - u_S(x,t)$ as

$$(37) \qquad \left(\frac{1}{I \cdot J} \sum_{i=1}^{I} \sum_{j=1}^{J} |u_{S+N}(x_i, t_j) - u_S(x_i, t_j)|^2 \right)^{1/2},$$

where x_i's range over $[1.65, 1.85]$ cm along which the first group of peaks resides, and t_j's range over $[14, 30]$ ms. The time interval $[0, 14]$ ms is skipped for the initial transients to die out. With the BM response to the pure tone input removed from the one to the combined signal input, rms (37) measures the BM response to the noise input. The rms measure is chosen for its robustness in treating noisy signals.

The bottom panel of Figure 16 shows the rms measure as the level of the pure tone S increases from 0 dB to 80 dB. The reference (zero dB) value is the rms value when the input is the noise only. The y-axis shows the amount of the suppression in dB. Before the suppressor (pure tone) level reaches 30 dB, there is almost no reduction of the noise. As tone intensity further increases, the noise is suppressed by as much as 25 dB when the pure tone level is 70 dB. After that, the suppression amount changes rather little, partly because the nonlinearity is saturated in the high intensity regime.

5. Conclusions. Time domain computation of a two–space-dimensional cochlear model with nonlinear nonlocal FF OHC coupling is shown to capture BM nonlinear responses to multifrequency sound input. Exploiting the matrix structure of the discrete FF model, we develop a fast iterative scheme in a second order accurate semi-implicit method. Numerical convergence of the method and its application to the study of nonlinear phenomena in hearing are demonstrated, such as tone-tone (noise) interaction, compression, and distortion products. The simulation results are in qualitative agreement with experimental findings.

A significant advantage of the time domain computation is that the time and complication involved in the computation do not depend on the number of input tones. In fact, the input can be any sound signal, e.g., noise. So the time domain approach readily applies to computing nonlinear responses when the input contains both tones and noise. The computational framework presented here may be upscaled as in [36] to model nonlinear masking of tones and noises in psychoacoustics.

Several simplified assumptions are made in the model: two-dimensional, fictitious mass and damping attached to the BM, and neglect of fluid viscosity and coiled shape of the cochlea. A comprehensive time-dependent model computed in [10] included these additional realistic factors though without OHC amplification. The price to pay is a more complicated model and longer computing time on the order of hours on parallel computers, while our run time is on the order of minutes on desktop computers.

6. Appendix: Iterative method and convergence. Denoting the right-hand term of (30) in section 3 by b, we write the linear system (30) as $(G(U)A - \tilde{M})x = b$, where $\tilde{M}$ is a diagonal matrix with the following diagonal entry:

$$(38) \qquad \tilde{M}(i,i) = m + \frac{2\Delta t}{3}r + \left(\frac{2\Delta t}{3}\right)^2 s(x_i) \qquad \text{for} \quad i = 0, \ldots, N.$$

The mass m and resistance r of the BM are constants, and the stiffness $s(x)$ is a function of the BM length. Let us also write $\tilde{M}(x_i)$ for $\tilde{M}(i,i)$.

We solve the linear equation $(G(U)A - \tilde{M})x = b$ by an iterative method. The method also applies to (27) in section 3. In fact, the matrix A is a product of a symmetric matrix A^s times a diagonal matrix D, where

$$(39) \qquad A^s_{i,j} = \sum_{n=1}^{M} \frac{-2\rho \, e^n(x_i) \, e^n(x_j)}{\xi_n \tanh(\xi_n H)} \Delta x,$$

and D is a diagonal matrix with its diagonal entries $D(i,i) = d(i)$. We see that $-DA^sD = -DA$ is symmetric and nonnegative. Recall from section 3 that $G(U)$ is invertible and that $B^{-1}C + Id = G(U)$ and $B + C = Id$. We have $B^{-1} = G(U)$.

Multiplying both sides of (30) by B, we get $(A^sD - B\tilde{M})x = Bb$. Let $y = Dx$; then the system can be written as $(A^s - B\tilde{M}D^{-1})y = Bb$. Subtracting $\tilde{M}D^{-1}y$ from both sides, we obtain

$$(40) \qquad (A^s - \tilde{M}D^{-1})y = (B\tilde{M}D^{-1} - \tilde{M}D^{-1})y + Bb = -C\tilde{M}D^{-1}y + Bb,$$

where we used $B + C = Id$ for the second equality. The iterative method now reads

$$(41) \qquad\qquad y^{k+1} = (A^s - \tilde{M}D^{-1})^{-1}(-C\tilde{M}D^{-1}y^k + Bb),$$

where k is the iteration step. Note that $-(A^s - \tilde{M}D^{-1})$ is a symmetric positive definite matrix and is independent of the time. The time-dependent property of $G(U)A - \tilde{M}$ is moved into the right-hand side which can be calculated simply by matrix multiplications. Since we have to compute only the inversion of $A^s - \tilde{M}D^{-1}$ once, the iterative procedure helps to save the computation time considerably.

For the iterative method to converge, we estimate the spectral radius of the iteration matrix $(A^s - \tilde{M}D^{-1})^{-1}C\tilde{M}D^{-1}$. Let $\rho(P)$ be the spectral radius of a matrix P, and let $||P||_2$ be the 2-norm, so $||P||_2 = \sqrt{\rho(P^TP)}$. Then

$$\rho((A^s - \tilde{M}D^{-1})^{-1}C\tilde{M}D^{-1}) \leq ||(A^s - \tilde{M}D^{-1})^{-1}C\tilde{M}D^{-1}||_2$$
$$\leq ||(A^s - \tilde{M}D^{-1})^{-1}||_2||C\tilde{M}D^{-1}||_2$$
$$(42) \qquad\qquad\qquad\qquad \leq \rho((A^s - \tilde{M}D^{-1})^{-1})||C\tilde{M}D^{-1}||_2.$$

By direct calculation,

$$(43) \qquad ||C\tilde{M}D^{-1}||_2 = \max\{2\alpha(x_0)\tilde{M}(x_0), \alpha(x_i)\tilde{M}(x_i), i = 1, \ldots, N - K - 1\},$$

where $\alpha(x_i)$ is the gain factor at $x = x_i$, and $\tilde{M}(x_i)$ is defined above. As $\alpha(x_0)$ is zero, i.e., there is no gain at the basal end point of the cochlea, we get

$$||C\tilde{M}D^{-1}||_2 = \max\{\alpha(x_i)\tilde{M}(x_i), i = 1, \ldots, N - K - 1\}$$
$$(44) \qquad\qquad\qquad \leq \max\{\alpha(x_i), i = 1, \ldots, N - K - 1\}\tilde{M}(x_1).$$

The last inequality holds because of the decreasing property of $\tilde{M}(x_i)$.

Since $-(A^s - \tilde{M}D^{-1})$ is positive definite, the minimum value of the eigenvalues is greater than 0. Let λ_m be the smallest eigenvalue; then $1/\lambda_m$ becomes the largest eigenvalue, and hence the spectral radius of $-(A^s - \tilde{M}D^{-1})^{-1}$. Now since $-A^s$ is nonnegative, for all the $v \in \Re^{(N+1)}$, $v^T(-A^s + \tilde{M}D^{-1})v \geq v^T\tilde{M}D^{-1}v$. Especially, substituting for v the unit length eigenvector v_m corresponding to λ_m, we get

$$(45) \qquad\qquad \lambda_m = v_m^T(-A^s + \tilde{M}D^{-1})v_m \geq v_m^T\tilde{M}D^{-1}v_m \geq \min\{\tilde{M}D^{-1}\}.$$

The last expression represents the smallest value of the diagonal entry of $\tilde{M}D^{-1}$. Using the property of $\tilde{M}(x_i)$ that it decreases along x_i and considering the multiplication by D^{-1} which doubles the first and last columns, we find $\min\{\tilde{M}D^{-1}\} = \tilde{M}(x_{N-1})$.

It follows that

$$
\begin{aligned}
\rho((A^s - \tilde{M}D^{-1})^{-1}C\tilde{M}D^{-1}) &\leq \frac{\max\{\alpha(x_i), i = 1, \ldots, N - K\}\tilde{M}(x_1)}{\lambda_m} \\
&\leq \frac{\max\{\alpha(x_i), i = 1, \ldots, N - K\}\tilde{M}(x_1)}{\tilde{M}(x_{N-1})} \\
&= \frac{\max\{\alpha(x_i)\}(m + \frac{2\Delta t}{3}r + (\frac{2\Delta t}{3})^2 s(x_1))}{m + \frac{2\Delta t}{3}r + (\frac{2\Delta t}{3})^2 s(x_{N-1})},
\end{aligned}
$$

(46)

which is less than 1 if Δt is small enough and $|\alpha(x_i)| \leq \bar{\alpha}_0 < 1$, for all i, for some positive constant $\bar{\alpha}_0$. For given model parameters, the upper bound of the spectral radius depends only on the time step Δt. In our simulation, the gain factor α is maximally 0.36. With the other parameters in Table 1, $\Delta t = 0.0045$ ms already makes the spectral radius less than 1, which guarantees that the iterative method converges. At $\Delta t = 0.0025$ ms used in the simulations, the spectral radius is about 0.54, and the iteration takes only $k = 30$ iteration steps to reach the error tolerance of 10^{-8}.

REFERENCES

[1] J. B. ALLEN, *Cochlear modeling*-1980, in Mathematical Modeling of the Hearing Process, Lecture Notes in Biomath. 43, M. Holmes and L. Rubenfeld, eds., Springer-Verlag, New York, 1980, pp. 1–8.

[2] N. P. COOPER AND W. S. RHODE, *Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea*, J. Neurophysiol., 78 (1997), pp. 261–270.

[3] E. DE BOER, *Mechanics of the cochlea: Modeling efforts*, in The Cochlea, Springer Handbook of Auditory Research 8, P. Dallos, A. Popper, and R. Fay, eds., Springer-Verlag, New York, 1996, pp. 258–317.

[4] E. DE BOER, *On active and passive cochlear models: Towards a generalized analysis*, J. Acoust. Soc. Amer., 73 (1983), pp. 574–576.

[5] E. DE BOER AND A. L. NUTTALL, *Properties of amplifying elements in the cochlea*, in Biophysics of the Cochlea: From Molecules to Models, A. W. Gummer, ed., World Scientific, Singapore, 2003, pp. 331–342.

[6] L. DENG, *Processing of acoustic signals in a cochlear model incorporating laterally coupled suppressive elements*, Neural Networks, 5 (1992), pp. 12–34.

[7] L. DENG AND C. D. GEISLER, *Responses of auditory-nerve fibers to multiple-tone complexes*, J. Acoust. Soc. Amer., 82 (1987), pp. 1989–2000.

[8] R. J. DIEPENDAAL AND M. A. VIERGEVER, *Nonlinear and active two-dimensional cochlear models: Time-domain solution*, J. Acoust. Soc. Amer., 85 (1989), pp. 803–812.

[9] C. D. GEISLER, *From Sound to Synapse*, Oxford University Press, Oxford, UK, 1998.

[10] E. GIVELBERG AND J. BUNN, *A comprehensive three-dimensional model of the cochlea*, J. Comput. Phys., 191 (2003), pp. 377–391.

[11] D. D. GREENWOOD, *A cochlear frequency-position function for several species: 29 years later*, J. Acoust. Soc. Amer., 87 (1990), pp. 2592–2605.

[12] J. J. GUINAN AND W. T. PEAKE, *Middle-ear characteristics of anesthetized cats*, J. Acoust. Soc. Amer., 41 (1967), pp. 1237–1261.

[13] M. HOLMES AND J. D. COLE, *Cochlear mechanics: Analysis for a pure tone*, J. Acoust. Soc. Amer., 76 (1984), pp. 767–778.

[14] A. ISERLES, *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, Cambridge, UK, 1997.

[15] Y. Jau and C. D. Geisler, *Results from a cochlear model utilizing longitudinal coupling*, in Mechanics of Hearing, E. de Boer and M. Viergever, eds., Martinus Nijhoff, Delft University Press, The Netherlands, 1983, pp. 169–176.

[16] J. B. Keller and J. C. Neu, *Asymptotic analysis of a viscous cochlear model*, J. Acoust. Soc. Amer., 77 (1985), pp. 2107–2110.

[17] R. J. Leveque, C. S. Peskin, and P. D. Lax, *Solution of a two-dimensional cochlea model using transform techniques*, SIAM J. Appl. Math., 45 (1985), pp. 450–464.

[18] R. J. Leveque, C. S. Peskin, and P. D. Lax, *Solution of a two-dimensional cochlea model with fluid viscosity*, SIAM J. Appl. Math., 48 (1988), pp. 191–213.

[19] M. C. Liberman, *The cochlear frequency map for the cat: Labeling auditory nerve fibers of known characteristic frequency*, J. Acoust. Soc. Amer., 72 (1982), pp. 1441–1449.

[20] K. M. Lim, *Physical and Mathematical Cochlear Models*, Ph.D. thesis, Stanford University, Stanford, CA, 2000.

[21] K. M. Lim and C. R. Steele, *A three-dimensional nonlinear active cochlear model analyzed by the WKB-numeric method*, Hearing Research, 170 (2002), pp. 190–205.

[22] K. M. Lim and C. R. Steele, *Response suppression and transient behavior in a nonlinear active cochlear model with feed-forward*, Internat. J. Solids Structures, 40 (2003), pp. 5097–5107.

[23] R. Lyon, *A computational model of filtering, detection, and compression in the cochlea*, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Paris, France, 1982, pp. 1282–1285.

[24] S. Neely, *Mathematical model of cochlear mechanics*, J. Acoust. Soc. Amer., 78 (1985), pp. 345–352.

[25] S. Osher and L. I. Rudin, *Feature-oriented image enhancement using shock filters*, SIAM J. Numer. Anal., 27 (1990), pp. 919–940.

[26] J. O. Pickles, *An Introduction to the Physiology of Hearing*, 2nd ed., Academic Press, London, 1988.

[27] K. Pohlmann, *Principles of Digital Audio*, 4th ed., McGraw–Hill Video/Audio Professional, New York, 2000.

[28] A. Recio, N. C. Rich, S. S. Narayan, and M. A. Ruggero, *Basilar-membrane responses to clicks at the base of the chinchilla cochlea*, J. Acoust. Soc. Amer., 103 (1998), pp. 1972–1982.

[29] W. S. Rhode and A. Recio, *Multicomponent stimulus interactions observed in basilar-membrane vibration in the basal region of the chinchilla cochlea*, J. Acoust. Soc. Amer., 110 (2001), pp. 3140–3154.

[30] L. Robles and M. A. Ruggero, *Mechanics of the mammalian cochlea*, Physiological Review, 81 (2001), pp. 1305–1352.

[31] M. A. Ruggero, N. C. Rich, A. Recio, S. S. Narayan, and L. Robles, *Basilar membrane responses to tones at the base of the chinchilla cochlea*, J. Acoust. Soc. Amer., 101 (1997), pp. 2151–2163.

[32] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth and Brooks, Pacific Grove, CA, 1989.

[33] R. Tsai and S. Osher, *Level set methods and their applications in image sciences*, Commun. Math. Sci., 1 (2003), pp. 623–656.

[34] G. von Békésy, *Experiments in Hearing*, McGraw–Hill, New York, 1960.

[35] J. Xin, *Dispersive instability and its minimization in time domain computation of steady state responses of cochlear models*, J. Acoust. Soc. Amer., 115 (2004), pp. 2173–2177.

[36] J. Xin and Y. Qi, *A PDE based two level model of the masking property of the human ear*, Commun. Math. Sci., 1 (2003), pp. 833–840.

[37] J. Xin and Y. Qi, *Global well-posedness and multi-tone solutions of a class of nonlinear nonlocal cochlear models in hearing*, Nonlinearity, 17 (2004), pp. 711–728.

[38] J. Xin, Y. Qi, and L. Deng, *Time domain computation of a nonlinear nonlocal cochlear model with applications to multitone interaction in hearing*, Commun. Math. Sci., 1 (2003), pp. 211–227.

[39] H. K. Zhao, S. Osher, B. Merriman, and M. Kang, *Implicit and non-parametric shape reconstruction from unorganized points using a variational level set method*, Computer Vision and Image Understanding, 80 (2000), pp. 295–319.

[40] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, 2nd ed., Springer Ser. Inform. Sci. 22, Springer-Verlag, Berlin, 1999.