JASA ARTICLE

Fast time-domain solution of a nonlinear three-dimensional cochlear model using the fast Fourier transform

Yasuki Murakami[a)]
Faculty of Design, Kyushu University, 4-9-1 Shiobaru, Minamiku, Fukuoka 815-8540, Japan

ABSTRACT:
A fast numerical time-domain solution of a nonlinear three-dimensional (3D) cochlear model is proposed. In dynamical systems, a time-domain solution can determine nonlinear responses, and the human faculty of hearing depends on nonlinear behaviors of the microscopically structured organs of the cochlea. Thus, time-domain 3D modeling can help explain hearing. The matrix product, an n^2 operation, is a central part of the time-domain solution procedure in cochlear models. To solve the cochlear model faster, the fast Fourier transform (FFT), an $n \log n$ operation, is used to replace the matrix product. Numerical simulation results verified the similarity of the matrix product and the FFT under coarse grid settings. Furthermore, applying the FFT reduced the computation time by a factor of up to 100 owing to the computational complexity of the proposed approach being reduced from n^2 to $n \log n$. Additionally, the proposed method successfully computed 3D models under moderate and fine grid settings that were unsolvable using the matrix product. The 3D cochlear model exhibited nonlinear responses for pure tones and clicks under various gain distributions in a time-domain simulation. Thus, the FFT-based method provides fast numerical solutions and supports the development of 3D models for cochlear mechanics.
 https://doi.org/10.1121/10.0006533
(Received 7 February 2021; revised 12 August 2021; accepted 13 September 2021; published online 11 October 2021)
[Editor: Sunil Puria] Pages: 2589–2599

I. INTRODUCTION

Computational modeling studies can help explain the processes underlying dynamical systems, and time-domain numerical methods are powerful tools for solving nonlinear models. The active and nonlinear nature of cochlear responses to sound characterize the human faculty of hearing (Kemp, 1978; Hudspeth, 2008). By definition, an active system has an external power source while a nonlinear system does not obey the superposition principle leading to distortions Allen 2020. Outer hair cells (OHCs) are thought to be the source of the active mechanism Brownell 1985 while there are multiple sources of non-linearity within the cochlea (e.g., Kemp 1978). The sense of hearing depends on complex behaviors of the microscopically structured organs of the cochlea, which present nonlinearities depending on sound pressure levels and frequency (Olson and Strimbu, 2020). The results of prior experiments suggest that complex mechanical motions of the cochlea through space lead to the nonlinearities; however, some disagreement remains on this point (Cooper *et al.*, 2018; Dewey *et al.*, 2019; Gao *et al.*, 2014; He *et al.*, 2018; Lee *et al.*, 2015). This fact suggests that time-domain three-dimensional (3D) modeling can help explain the mechanism of hearing because it can predict the nonlinear cochlear responses caused by motions of the microscopically structured organs.

One-dimensional (1D) models of the cochlea have been used to investigate the cochlear nonlinearities, including Epp *et al.* (2010), Ku *et al.* (2008), Liu and Neely (2010), and Verhulst *et al.* (2012). Modern computers can easily solve 1D models of the cochlea with several hundred segments (Pan *et al.*, 2015). However, some modeling studies require greater computational resources to improve the reproducibility of the experimental data using two-dimensional (2D) or 3D models (Elliott *et al.*, 2017; Murakami, 2018; Prodanovic *et al.*, 2019; Vencovský *et al.*, 2020). As a result, a high computational cost is required to solve these cochlear models owing to their increased numbers of spatial grids.

Solving the Poisson equation represents a central task in computational fluid dynamics. Therefore, the 1D cochlear model is formulated using the Poisson equation as the governing equation with boundary conditions (Diependaal *et al.*, 1987; Elliott *et al.*, 2007; Neely and Kim, 1986). To solve the Poisson equation numerically, the spatial derivative must be converted into a difference. The finite difference (FD) approach approximates the derivative most directly to solve the cochlear boundary value problem (Elliott *et al.*, 2007; Neely and Kim, 1986). In addition, the Laplace equation describes a 2D or 3D model of the cochlea similar to the 1D model (Neely, 1981). Because the Laplace equation is a homogeneous version of the Poisson equation, the 2D cochlear model can be rewritten using Poisson's equation in the process of discretization (Neely, 1981).

A fast solver for the Poisson equation has been proposed using the fast Fourier transform (FFT) (Schumann and Sweet, 1988). Discretizing the Poisson equation eventually creates another problem of obtaining a solution to the simultaneous equations. This approach is called the direct

[a)]Electronic mail: murakami@design.kyushu-u.ac.jp

method. The matrix product is a major part of this method, and it is performed with a calculation complexity on the order of at least n^2. Thus, higher-dimensional models swiftly increase the computational cost depending on their spatial grid settings. In contrast, the FFT-based method can reduce the computational complexity to $n \log n$ (Schumann and Sweet, 1988). The FFT-based method is a method to obtain the solution of partial differential equations and is known as a pseudo-spectral method (Orszag, 1972). This method is closely related to the spectral method. However, a pseudo-spectral basis simplifies the evaluation of certain operators and allows for faster computation when using the FFT.

Allen and Sondhi (1979) first applied the FFT to solve the cochlear integral equation in the time domain. Their idea was based on the convolutional property of the Fourier transform for two place-dependent variables g and p as follows:

$$g(x) * p(x) \leftrightarrow G(k)P(k),$$

where $*$ and $\leftrightarrow$ denote convolution and the Fourier transform, respectively, and x and k, respectively, represent place and spatial frequency. The place-domain convolution follows directly from the Green's function in the cochlear integral equation. By using FFT, Allen and Sondhi (1979) converted the place-domain convolution into a multiplication in the spatial frequency domain to solve the equation more rapidly. This approach is well documented in the signal processing field, where it is denoted as the fast convolution. Subsequently, Diependaal *et al.* (1987) developed a robust and efficient method for nonlinear time-domain 1D models of the cochlea. The fundamental ideas of this method can be extended to 2D and 3D models, and they can be applied to model the motion of compressive and viscous fluids. Furthermore, the computational speed of this solution method is similar to that of the solution method based on sparse matrix manipulation, which is often used for fast calculation of large-scale models (Pan *et al.*, 2015).

To develop a fast solution method for nonlinear time-domain cochlear models, in this study, we propose a method that replaces the direct solver in the method proposed by Diependaal *et al.* (1987) with an FFT-based solver. Furthermore, the proposed method employs a 3D cochlear model including two-degree-of-freedom (2DoF) models that are nonlinearized by a saturating function.

The remainder of this paper is organized as follows. Section II describes the FFT-based solution of the nonlinear time-domain model for 3D cochlear mechanics. Section III discusses the evaluation of the proposed method's performance by time-domain simulations, and a final discussion and conclusions are presented in Sec. IV.

II. MODEL

A. FFT solution for the cochlear model

The cochlear model consists of macro- and micro-mechanical models representing the fluid motion and mechanical behavior of the cochlear organs, respectively (Allen, 1977, 1980; Neely and Kim, 1986). In this study, we aim to adapt an FFT-based solver for the macro-mechanical model, which is the central component of the time-domain solution (Diependaal *et al.*, 1987). The starting point introduces the Laplace equation, a homogeneous version of the Poisson equation, governing the 3D macro-model of the cochlea. In the present work, a rectangular box-shaped cochlear chamber was assumed, which was discretized in uniform grids

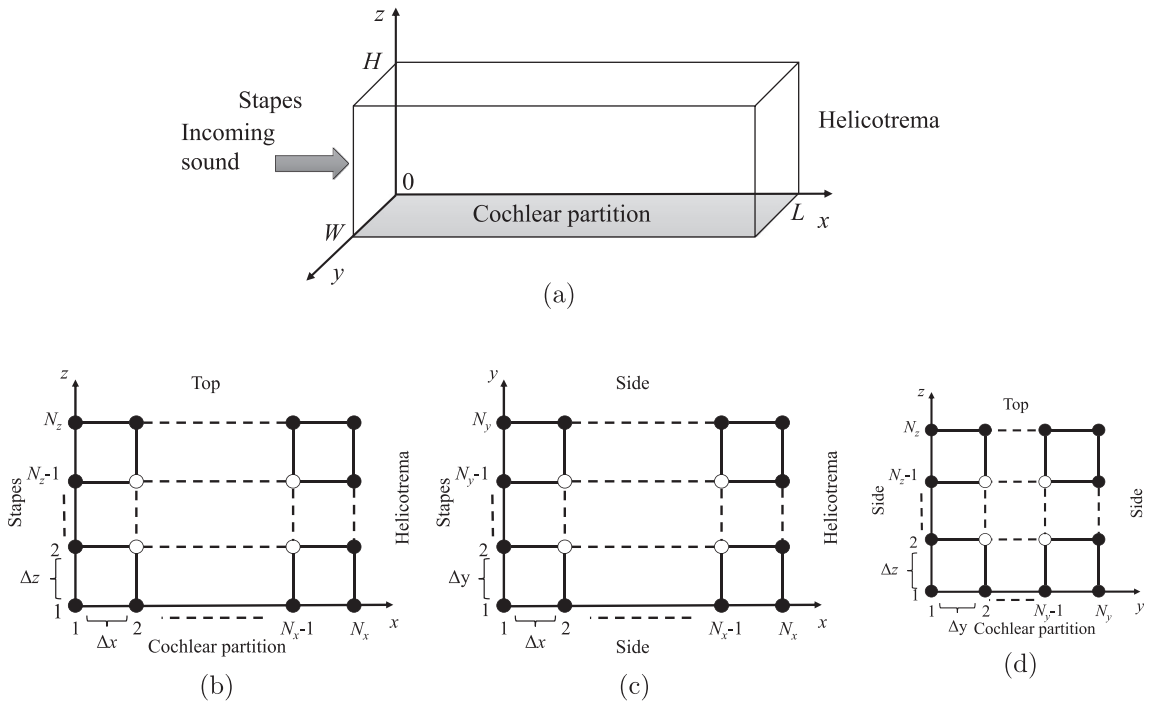

FIG. 1. (a) Illustrated scheme of a box-shaped 3D macro-mechanical model of the cochlea. (b)–(d) Computational grids for 3D macro-mechanical models of the cochlea. Filled and open circles denote boundary and internal points, respectively.

https://doi.org/10.1121/10.0006533

as shown in Fig. 1; an ideal fluid was assumed to fill the chamber. Under these assumptions, the pressure difference p is required to satisfy Laplace's equation in the cochlear fluid:

$$\frac{\partial^2 p(x,y,z,t)}{\partial x^2}+\frac{\partial^2 p(x,y,z,t)}{\partial y^2}+\frac{\partial^2 p(x,y,z,t)}{\partial z^2}=0, \quad (1)$$

where t is time, and x, y, and z are the longitudinal, width, and height location, respectively. The boundary conditions are given as

$$\left.\frac{\partial p(x,y,z,t)}{\partial x}\right|_{x=0}=2\rho\ddot{u}_s(t), \quad (2)$$

$$p(x,y,z,t)|_{x=L}=0, \quad (3)$$

$$\left.\frac{\partial p(x,y,z,t)}{\partial y}\right|_{y=0}=0, \quad (4)$$

$$\left.\frac{\partial p(x,y,z,t)}{\partial y}\right|_{y=W}=0, \quad (5)$$

$$\left.\frac{\partial p(x,y,z,t)}{\partial z}\right|_{z=0}=2\rho\ddot{u}(t), \quad (6)$$

$$\left.\frac{\partial p(x,y,z,t)}{\partial z}\right|_{z=H}=0, \quad (7)$$

where $\ddot{u}_s$ and $\ddot{u}$ are the acceleration of the stapes and the cochlear partition (CP), respectively, L and H are the length and height of the chamber, respectively, and ρ is the density of the cochlear fluid.

The Poisson equation can be solved by the eigenfunction expansion method (Swarztrauber, 1977); the eigenfunctions are composed of sine and cosine expressions. Hence, a Fourier transform can be applied to solve the Poisson equation in a uniform grid (Schumann and Sweet, 1988). The advantage of the FFT-based algorithm is that it reduces the computational complexity of the problem. This algorithm uses the discrete cosine transform (DCT) depending on the boundary conditions. According to the literature (Schumann and Sweet, 1988), the boundary conditions should not contain the pressure variable p when applying the FFT-based solver. However, as described in Eq. (6), the boundary condition in the z-direction includes the pressure variable p because it is a term used to denote the acceleration of the CP $\ddot{u}$ in micro-mechanical models. Thus, in the present work, an attempt was made to apply the FFT solver without reference to the z-direction. This approach is suitable because there were more segments in the x-direction than in the z-direction. In addition, a nonuniform grid setting can be set (Prodanovic *et al.*, 2019). For these reasons, this study employed the conventional FD scheme (Neely, 1981) that discretizes the macro-mechanical model in the z-direction.

By approximating the FD scheme for Eq. (1) in the z-direction with spatial step Δz and number of segments N_z, the governing equation of the cochlear model can be written as

$$\frac{\partial^2 p_{n_z}(x,y,t)}{\partial x^2}+\frac{\partial^2 p_{n_z}(x,y,t)}{\partial y^2}+D_{zz}\left[p_{n_z}(x,y,t)\right]=0, \quad (8)$$

where the operator D_{zz} represents the second-order numerical derivative in the z-direction, and the suffix $n_z=1\ldots N_z$ represents the discrete location for the z-direction. By applying the boundary conditions in Eqs. (6) and (7), the discretized Eq. (8) can be written as

$$\frac{\partial^2 \boldsymbol{p}(x,y,t)}{\partial x^2}+\frac{\partial^2 \boldsymbol{p}(x,y,t)}{\partial y^2}+\boldsymbol{A}\boldsymbol{p}(x,y,t)=\boldsymbol{g}(x,y,t), \quad (9)$$

where

$$\boldsymbol{A}=\frac{-1}{(\Delta z)^2}\begin{pmatrix} 2+\alpha & -2 & & & & 0 \\ -1 & 2 & -1 & & & \\ & & \ddots & & & \\ & & & -1 & 2 & -1 \\ 0 & & & & -2 & 2 \end{pmatrix}, \quad (10)$$

$$\boldsymbol{p}(x,y,t)=\left(p_1(x,y,t)\ldots p_{N_z}(x,y,t)\right), \quad (11)$$

$$\boldsymbol{g}(x,y,t)=(-\alpha g(x,y,t)\,0\cdots 0)^T. \quad (12)$$

In Eqs. (10) and (12), the variables α and g depend on a micro-mechanical model of the cochlea. A specific example presented below explains how the values of α and g are set.

The time-domain solution (Diependaal *et al.*, 1987) solves Eq. (9) in each time step. The values of the vector $\boldsymbol{g}$ are determined before solving Eq. (9) by calculating the micro-mechanical model of the cochlea. To adapt the FFT-based solver for the macro-mechanical model, Eq. (9) is divided for both the x- and y-direction with spatial steps Δx and Δy and number of segments N_x and N_y by using the FD scheme. The model equation can be written as

$$D_{xx}\left[\boldsymbol{p}_{n_x n_y}(t)\right]+D_{yy}\left[\boldsymbol{p}_{n_x n_y}(t)\right]+\boldsymbol{A}\boldsymbol{p}_{n_x n_y}(t)=\boldsymbol{g}_{n_x n_y}(t), \quad (13)$$

where $n_x=1,\ldots,N_x$ and $n_y=1,\ldots,N_y$ represent the discrete location for the x- and y-direction, respectively.

Mixed boundary conditions defined at the position $n_x=1$ and N_x are given by Eqs. (2) and (3), respectively, as expressed below:

$$\frac{-\boldsymbol{p}_{1n_y}(t)+\boldsymbol{p}_{2n_y}(t)}{\Delta x}=2\rho\ddot{\boldsymbol{u}}_s(t), \quad (14)$$

$$\boldsymbol{p}_{N_x n_y}=[0\cdots 0]^T. \quad (15)$$

In addition, the Neumann boundary condition defined at the position $n_y=1$ and N_y is expressed by Eqs. (4) and (5), respectively,

$$\frac{-\boldsymbol{p}_{n_x 1}(t)+\boldsymbol{p}_{n_x 2}(t)}{\Delta y}=[0\cdots 0]^T, \quad (16)$$

$$\frac{\boldsymbol{p}_{n_xN_y-1}(t)-\boldsymbol{p}_{n_xN_y}(t)}{\Delta y}=[0\cdots 0]^T. \tag{17}$$

The cochlear model includes the mixed and Neumann boundary conditions, as expressed in the preceding equations. DCTs have types from I to IV, and the type of DCT to be applied depends on the boundary conditions (Schumann and Sweet, 1988). In the mixed and Neumann boundary conditions, the proposed method employs DCT-III and DCT-I to solve the cochlear model in the x and y dimensions, respectively.

The details of the FFT-based solution can be divided into three stages: analysis, computation, and synthesis (Schumann and Sweet, 1988). First, the analysis stage obtains the DCT coefficients $\bar{\boldsymbol{g}}_{n_xi_y}(t)$ for the y-direction using DCT-I,

$$\begin{aligned}\bar{\boldsymbol{g}}_{n_xi_y}(t)=&\frac{1}{N_y-1}\boldsymbol{g}_{n_x1}(t)+\frac{2}{N_y-1}\sum_{n_y=1}^{N_y}\boldsymbol{g}_{n_xn_y}(t)\\&\times\cos\frac{(n_y-1)(i_y-1)\pi}{N_y-1}\\&-\frac{1}{N_y-1}\boldsymbol{g}_{n_xN_y}(t)(-1)^{i_y},\end{aligned} \tag{18}$$

where $i_y=1,\ldots,N_y$. For the x-direction, the DCT coefficient $\bar{\bar{\boldsymbol{g}}}_{i_xi_y}(t)$ is obtained using DCT-III,

$$\begin{aligned}\bar{\bar{\boldsymbol{g}}}_{i_xi_y}(t)=&\frac{1}{N_x}\bar{\boldsymbol{g}}_{1i_y}(t)\\&+\frac{2}{N_x}\sum_{n_x=1}^{N_x-1}\bar{\boldsymbol{g}}_{n_xi_y}(t)\cos\frac{(n_x-1)(2i_x-1)\pi}{2N_x}.\end{aligned} \tag{19}$$

The next stage computes the DCT coefficients of the pressure vector $\bar{\bar{\boldsymbol{p}}}_{i_xi_y}(t)$ from the DCT coefficients $\bar{\bar{\boldsymbol{g}}}_{i_xi_y}(t)$,

$$\left(\frac{\lambda_{i_x}}{(\Delta x)^2}\boldsymbol{I}+\frac{\lambda_{i_y}}{(\Delta y)^2}\boldsymbol{I}+\boldsymbol{A}\right)\bar{\bar{\boldsymbol{p}}}_{i_xi_y}(t)=\bar{\bar{\boldsymbol{g}}}_{i_xi_y}(t), \tag{20}$$

where λ_{i_x} and λ_{i_y} represent the eigenvalues that can be calculated by

$$\lambda_{i_x}=-4\sin^2\frac{(2i_x-1)\pi}{4N_x}, \tag{21}$$

$$\lambda_{i_y}=-4\sin^2\frac{(i_y-1)\pi}{2(N_y-1)}. \tag{22}$$

The final stage synthesizes the DCT coefficients $\bar{\bar{\boldsymbol{p}}}_{i_xi_y}(t)$ to obtain the desired pressure vector $\boldsymbol{p}_{n_xn_y}(t)$ from the inverse transform,

$$\boldsymbol{p}_{n_xn_y}(t)=\sum_{i_x=1}^{N_x}\bar{\boldsymbol{p}}_{i_xn_y}(t)\cos\frac{(n_x-1)(2i_x-1)\pi}{4N_x}, \tag{23}$$

where

$$\begin{aligned}\bar{\boldsymbol{p}}_{i_xn_y}(t)=&\frac{1}{2}\bar{\bar{\boldsymbol{p}}}_{i_x1}(t)+\sum_{i_y=2}^{N_y-1}\bar{\bar{\boldsymbol{p}}}_{i_xi_y}(t)\\&\times\cos\frac{(n_y-1)(i_y-1)\pi}{N_y-1}-\frac{1}{2}\bar{\bar{\boldsymbol{p}}}_{i_xN_y}(t)(-1)^{n_y}.\end{aligned} \tag{24}$$

B. Specific example

The macro-cochlear model is linear and assumes an ideal fluid, as described in Sec. II A. The CP includes the source of the cochlear nonlinearity that transmits to the entire cochlea through coupling with the fluid facing the boundary condition described in Eq. (6). This perspective was predicted in Goldstein's series of works (Goldstein, 1967; Goldstein and Kiang, 1968) and was proven correct in the early 1970s (Kim *et al.*, 1973; Rhode, 1971). To date, the cochlea is considered a nonlinear system, and the active processing via hair bundle and/or somatic motilities of OHCs in micro-mechanics of the cochlea is a major hypothesis to explain the observed cochlear nonlinearity (Ashmore, 2019; Hudspeth, 2008).

The effectiveness of the proposed FFT-based solver for the macro-mechanical model was demonstrated by combining it with the micro-mechanical model proposed by Neely and Kim (1986). This micro-mechanical model is a 2DoF model including active feedback, as shown in Fig. 2, that can better account for the experimental data than a one-degree-of-freedom (1DoF) model (Elliott *et al.*, 2017). The pressure difference p in the macro-mechanical model drives the lower mass m_1 at $z=0$. Furthermore, active pressure γp_a acts on the mass m_1, and γ denotes the feedback gain coefficient. The lower and upper mass m_1 and m_2 represent the basilar membrane (BM) and the tectorial membrane (TM), respectively. The equations of motion can be written as

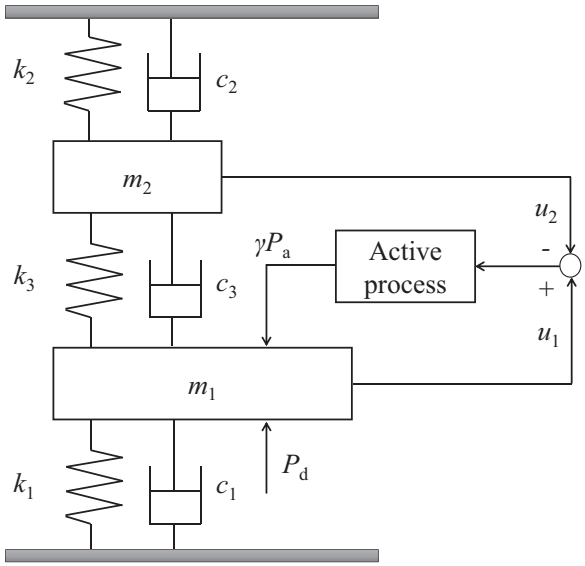

FIG. 2. Micro-mechanical model proposed by Neely and Kim (1986). Masses m_1 and m_2 represent the BM and TM, respectively.

https://doi.org/10.1121/10.0006533

$$p+\gamma p_a = m_1\ddot{u}_1+(c_1+c_3)\dot{u}_1+(k_1+k_3)u_1-c_3\dot{u}_2-k_3u_2, \tag{25}$$

$$0 = -c_3\dot{u}_1-k_3u_1+m_2\ddot{u}_2+(c_2+c_3)\dot{u}_2+(k_2+k_3)u_2, \tag{26}$$

where u_1 and u_2 represent the displacement of the two masses, $\dot{u}_1$ and $\dot{u}_2$ indicate their velocities, and $\ddot{u}_1$ and $\ddot{u}_2$ are their accelerations. The symbols c and k correspond to the damping and stiffness coefficients, and their subscripts are shown in Fig. 2.

The displacement of the CP $u(x,y,t)$ is assumed in the same bending mode, which is independent of x and y and is proportional to the displacement of the BM $u_1(x,y,t)$, as given below:

$$u(x,y,t) = bu_1(x,y,t), \tag{27}$$

where b defines the ratio of the displacement of the CP to the displacement of the BM.

In the macro-mechanical model, Eqs. (10) and (12) introduce variables α and g, which depend on the macro-mechanical model. By defining α and g as

$$\alpha = \frac{4\rho b}{m_1\Delta z}, \tag{28}$$

$$g = (c_1+c_3)\dot{u}_1+(k_1+k_3)u_1-c_3\dot{u}_2-k_2u_2+\gamma p_a, \tag{29}$$

the model Eq. (9) can be determined using Eqs. (28) and (29).

The active pressure p_a is proportional to the gap between two masses m_1 and m_2 in the original model and is also proportional to the damping and stiffness components c_4 and k_4. In the proposed method, the active pressure is nonlinearized using a saturation function to develop a nonlinear cochlear model as follows:

$$p_a = c_4\dot{u}_c + k_4u_c, \tag{30}$$

with

$$u_c = \beta\tanh\left(\frac{u_1-u_2}{\beta}\right), \tag{31}$$

where β represents the maximum value of u_c.

From Eqs. (25) and (29), the displacement and velocity of the BM must satisfy the system of the ordinary differential equation

$$\ddot{\boldsymbol{u}}_1(t) = \frac{\boldsymbol{p}(t)-\boldsymbol{g}(t)}{m_1}, \tag{32}$$

where

$$\boldsymbol{u}_1(t) = \begin{pmatrix} u_1(0,0,t) & u_1(\Delta x,0,t) & \dots & u_1(L,0,t) \\ u_1(0,\Delta y,t) & u_1(\Delta x,\Delta y,t) & \dots & u_1(L,\Delta y,t) \\ \vdots & \vdots & \vdots & \vdots \\ u_1(0,W,t) & u_1(\Delta x,W,t) & \dots & u_1(L,W,t) \end{pmatrix}, \tag{33}$$

$$\boldsymbol{p}(t) = \begin{pmatrix} p(0,0,t) & p(\Delta x,0,t) & \dots & p(L,0,t) \\ p(0,\Delta y,t) & p(\Delta x,\Delta y,t) & \dots & p(L,\Delta y,t) \\ \vdots & \vdots & \vdots & \vdots \\ p(0,W,t) & p(\Delta x,W,t) & \dots & p(L,W,t) \end{pmatrix}, \tag{34}$$

$$\boldsymbol{g}(t) = \begin{pmatrix} g(0,0,t) & g(\Delta x,0,t) & \dots & g(L,0,t) \\ g(0,\Delta y,t) & g(\Delta x,\Delta y,t) & \dots & g(L,\Delta y,t) \\ \vdots & \vdots & \vdots & \vdots \\ g(0,W,t) & g(\Delta x,W,t) & \dots & g(L,W,t) \end{pmatrix}. \tag{35}$$

In addition, from Eq. (26), by defining g_2 as

$$g_2 = -c_3\dot{u}_1-k_3u_1+(c_2+c_3)\dot{u}_2+(c_2+k_2)u_2, \tag{36}$$

the displacement and velocity of the tectorial membrane also satisfy

$$\ddot{\boldsymbol{u}}_2(t) = -\frac{\boldsymbol{g}_2(t)}{m_2}, \tag{37}$$

where

$$\boldsymbol{g}_2(t) = \begin{pmatrix} g_2(0,0,t) & g_2(\Delta x,0,t) & \dots & g_2(L,0,t) \\ g_2(0,\Delta y,t) & g_2(\Delta x,\Delta y,t) & \dots & g_2(L,\Delta y,t) \\ \vdots & \vdots & \vdots & \vdots \\ g_2(0,W,t) & g_2(\Delta x,W,t) & \dots & g_2(L,W,t) \end{pmatrix}. \tag{38}$$

C. Procedures for time-domain simulation

The proposed method employs the time-domain solution for the cochlear model proposed by Diependaal *et al.* (1987), and it uses the 3D cochlear model with the FD approach, as described in Sec. II A. The original version of the time-domain solution was applied to the 1D cochlear model with the finite element (FE) approach (Diependaal *et al.*, 1987). However, no modification of the procedure is required to solve the model presented in the time domain.

Diependaal *et al.* (1987) outlines the procedure as follows.

(i) Discretize time series t into sequence t_i with the time step Δt.
For $i = 1, \dots, I$, perform steps (ii)–(v):
(ii) Compute at time t_{i-1} the vectors $\boldsymbol{g}$ and $\boldsymbol{g_2}$ in Eqs. (35) and (38).
(iii) Solve $\boldsymbol{p}$ from Eqs. (18)–(23).
(iv) Compute Eqs. (32) and (37).
(v) Integrate Eqs. (32) and (37) from time t_{i-1} to t_i.

The method presented here replaces the matrix product with the FFT in step (iii) from the original version, called

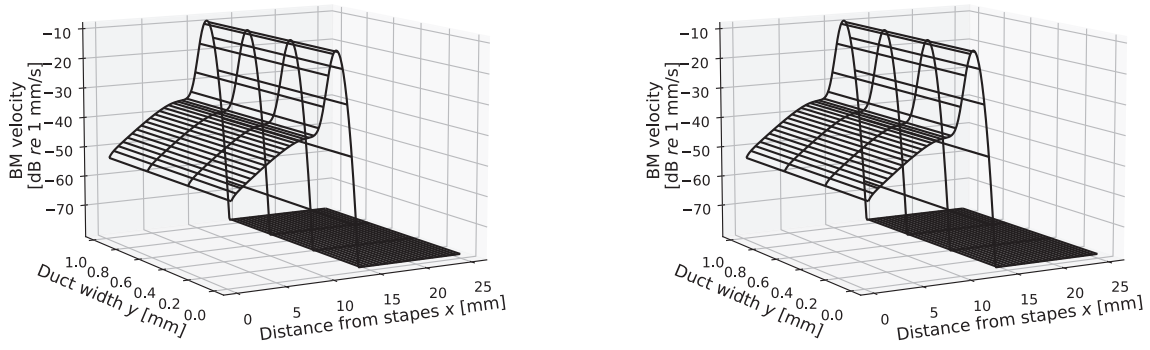

FIG. 3. Envelopes of BM velocity $\dot{u}_b$ for pure tones. The input frequency and intensity were 4 kHz and 40 dB. The value of the gain factor γ was set to 0.7.

the direct method. This step has the highest computational cost in the simulation. Replacing it with the FFT reduces an apparently $(N_xN_y)^2$ calculation to the order of $N_xN_y \log N_xN_y$ operations because the FFT-based solver obtains the pressure vector $\boldsymbol{p}$ from Eqs. (18)–(23) with the $N_xN_y \log N_xN_y$ operation. In contrast, the direct method computes the matrix product with the $(N_xN_y)^2$ operation. Furthermore, the inverse matrix computed using the Gaussian elimination method as a preprocessing step is used in the calculation in the direct method.

D. Simulation conditions

In this study, we chose the fourth-order Runge–Kutta method for the system of ordinary differential equations in step (v) of the algorithm with constant time step $\Delta t = 2$ [μs]. The simulation was performed on a Linux-based computer with an Intel Core-i9 9900k central processing unit (CPU) and 64 GB RAM. The programs were written in the PYTHON programming language.

The values of the parameters were set to those of the original model representing a cat's auditory nerve responses (Neely and Kim, 1986). However, to adjust the amount of amplification by active feedback, the values of the damping factors were increased by twice those of the original model.

III. RESULTS

In this study, we evaluated whether the FFT-based method increased the computational speed of the nonlinear time-domain cochlear model. In addition, to investigate the advantages of the 3D model, we performed time-domain simulations for various gain distributions.

A. Evaluation of computation time

The BM responses for pure tones is a typical feature in cochlear mechanics and is generated by the pressure difference p. Figure 3 shows envelopes of the BM velocity $\dot{u}$ distribution for pure tones when the response had reached a steady state. The FFT-based solver and the direct solver show similar envelopes of BM velocity $\dot{u}$. Furthermore, to evaluate quantitative differences between the FFT-based solution and the direct solution, the L2 norm was defined as

$$L_2 = \sqrt{\frac{\sum_{n_x}\sum_{n_y}\left(\dot{u}^{direct}_{n_xn_y}(t) - \dot{u}^{fft}_{n_xn_y}(t)\right)^2}{\sum_{n_x}\sum_{n_y}\left(\dot{u}^{direct}_{n_xn_y}(t)\right)^2}}, \tag{39}$$

where $\dot{u}^{direct}_{n_xn_y}$ and $\dot{u}^{fft}_{n_xn_y}$ represent the BM velocities obtained from the direct solver and the FFT-based solver, respectively. Table I shows the L2-norms of the calculated BM velocities between the FFT-based solutions and the direct solutions. Under all grid conditions, the values of the differences were sufficiently small because the number of significant digits was approximately 14 when the calculation was conducted with double-precision floating-point values. In the present

TABLE I. Difference of the BM velocity calculated between the FFT-based solution and the direct solution defined by the L2 norm in Eq. (39). A dash indicates that the direct method was not able to obtain the solution owing to an out-of-memory error during the computation.

		N_x								
		300			600			1200		
		N_y			N_y			N_y		
		4	8	16	4	8	16	4	8	16
N_z	4	6.0×10^{-14}	9.2×10^{-14}	6.6×10^{-13}	5.9×10^{-14}	1.5×10^{-14}	7.2×10^{-13}	1.7×10^{-9}	1.3×10^{-12}	—
	8	2.0×10^{-13}	4.1×10^{-13}	8.3×10^{-13}	1.1×10^{-7}	1.3×10^{-11}	—	—	—	—
	16	6.5×10^{-7}	1.6×10^{-8}	—	—	—	—	—	—	—

https://doi.org/10.1121/10.0006533

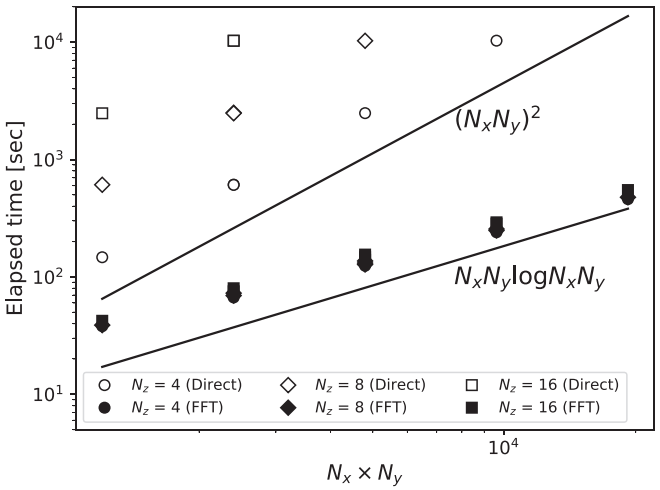

FIG. 4. Comparison of computation times for the model under different grid settings. The computations were performed for 10 ms of a pure tone. The filled and non-filled symbols indicate the results obtained from the FFT-based solver and the direct solver, respectively. The line represents the theoretically estimated increment. The values of the intercepts of the line are arbitrary.

work, the software program developed used double-precision floating-point values for all computations. Thus, this result indicates that there was no significant difference between the two solutions.

Computational speed is an important feature in adapting the FFT-based solver for the cochlear model. Figure 4 shows that the computation times followed two tendencies. First, the times were smaller for the FFT-based method than for the direct method under all conditions. Second, the computation times increased monotonically with the number of segments. The increase in computation time can be explained theoretically. The computation times in the direct method increase by approximately $(N_xN_y)^2$ because step (iii) of the algorithm multiplies the matrix and the vector. By contrast, in the FFT-based method, step (iii) of the algorithm uses the FFT, which is an $N_xN_y \log N_xN_y$ computation, instead of the matrix product. In the program, step (iii) of the algorithm is a bottleneck for fast calculation and is the only difference between the FFT-based method and the direct method. The modification in step (iii) of the algorithm helps realize fast calculation in the proposed cochlear model.

B. Responses for pure tone with gain distribution simulating alignment of OHCs

The response to pure tones is a fundamental feature in cochlear mechanics. To test the advantage of the 3D model, the gain distribution was set as shown in Fig. 5(a). The value of the gain factor γ was set to 0.8 at the middle site to

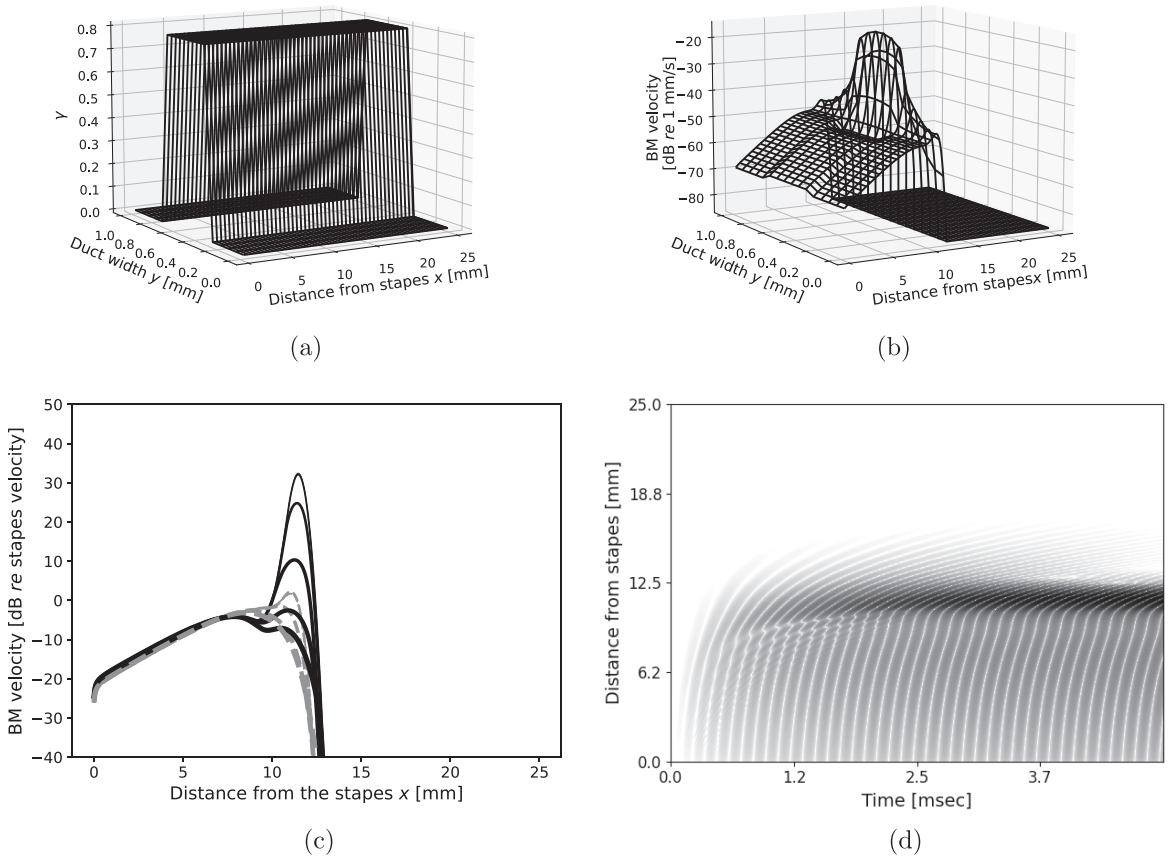

FIG. 5. (a) Gain distribution for simulating alignment of OHCs. (b) Envelope of BM velocity $\dot{u}_b(x, y)$ when input frequency and intensity were 4 kHz and 40 dB, respectively. (c) Sensitivity of BM velocity $\dot{u}_b(x, y)$ along the cochlear length. The solid and dashed lines represent sensitivity at $y = 0.5$ and 0 mm, respectively. The intensity was varied from 0 to 100 dB with a 20 dB step. Thin and thick lines indicate the responses for 0 and 100 dB tones, respectively. (d) Temporal pattern in BM velocity at $y = 0.5$ mm for tone as used in (b). The shading represents the magnitude of the absolute value of the BM velocity.

simulate OHC activity. The OHCs align near the middle of the BM in the transverse direction in the real cochlea.

The spatial pattern of the envelope of BM velocity for pure tone peaks at the characteristic frequency (CF) site, as shown in Fig. 5(b). In this case, this trend clearly appeared around the middle of the BM in the transverse direction. Then the BM velocity decreased toward the edge.

Compression is a typical feature of cochlear nonlinearities (Hudspeth, 2008). This feature indicates that the sensitivity of cochlear responses decreases with increasing input sound pressure level. The model included the saturation function in Eq. (31) as the source of nonlinearities. Figure 5(c) reproduces the compression in the BM velocity at the CF site. The sensitivity reached 30 dB for lower-intensity tones and dropped below −10 dB at high-intensity tones. However, at the edge site $y = 0$, the sensitivity was constant.

The temporal pattern of BM velocity propagates from the base to the apex region along the cochlear length, as shown in Fig. 5(d). The input sound wave reached the CF position within 0.5 ms. However, the magnitude of the BM velocity was still weak at this point and became stronger after 2.5 ms.

Figure 6 shows the maximum pressure distribution in the chamber. On the x–z plane, higher pressures were obtained at the cochlear entrance, which monotonically decreased toward the apex, as shown in Fig. 6(a). However, where x was approximately 12 mm, the pressure increased around $z = 0$. As shown in Fig. 5(c), this location was sensitive to the pure tone. Additionally, Fig. 6(b) shows the pressure distribution on the y–z plane, where x was the most sensitive location for the pure tone. In this plane, the higher pressure distribution was concentrated around the center at $z = 0$, which locates the OHCs as shown in Fig. 5(a).

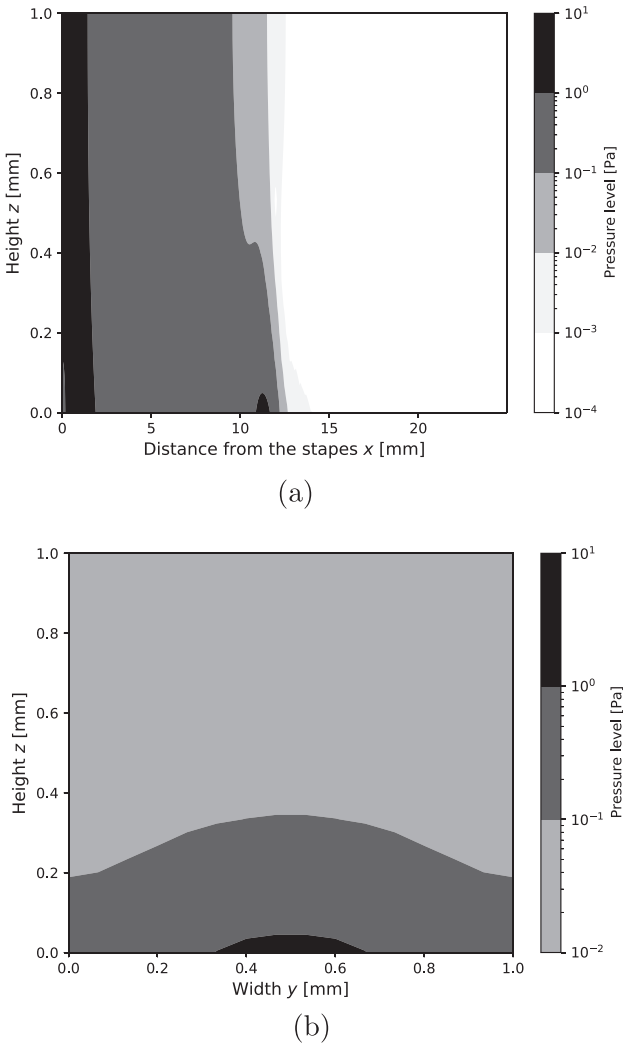

FIG. 6. Maximum pressure distribution in the chamber for the pure tone under the same condition of Fig. 5(b). (a) x–z plane at $y = 0.5$ mm. (b) y–z plane where the location x was most sensitive location for the pure tone.

C. Responses to clicks with stepped gain distribution

Natural sounds contain broadband spectral components, and click stimuli are useful to investigate BM responses for spectral component spread. Particularly, a specific frequency can be enhanced by spatial irregularity in the impedance of the CP, known as spontaneous otoacoustic emission (Shera, 2003). In the proposed model, spatial variations in the values of the feedback gain coefficient γ generated spatial irregularity of the CP impedance. The stepped gain location was set at $x = 10$ mm with an OHC simulated gain distribution, as shown in Fig. 7(a).

Figure 7(b) shows that clicks induced a steady response in the pressure-difference variation at the cochlear entrance site $x = 0$. Immediately after a click was input, there was no fluctuation in the pressure difference. However, after 10 ms without variations, the amplitude of the pressure difference began to fluctuate and reached a steady state. This trend was captured by a 1D model (Elliott *et al.*, 2007).

To describe the response to clicks inside the cochlear model in more detail, we note that the spatial pattern of the BM velocity peaked at the stepped location $x = 10$ mm and spread from the stepped site to the base, as shown in Fig. 7(c). Although the peaks occurred at positions with large values of the gain coefficient, the peaks also appeared at places with zero values of the gain coefficient. Furthermore, Fig. 7(d) shows a back-propagating and standing wave in BM velocity at $y = 0.5$ mm, the center of the duct width. The back-propagating wave was generated at the stepped site, $x = 10$ mm, and the standing wave occurred around the cochlear entrance. In addition, a weaker wave travelled to the cochlear apex site. For these phenomena, laser amplification theory can explain that an impedance mismatch site such as a stepped gain location causes the back-propagating and standing wave (Shera, 2003).

Figure 8 shows the maximum pressure distribution in the cochlear chamber. On the x–z plane, higher pressure was obtained at the cochlear base site, as shown in Fig. 8(a), and it monotonically decreased around the far side from the CP. By contrast, there were three regions of higher pressure near the CP. These three regions were closer to the cochlea entrance than the location of the stepped gain where $x = 10$ mm. In addition, as shown in Fig. 8(b), the high

https://doi.org/10.1121/10.0006533

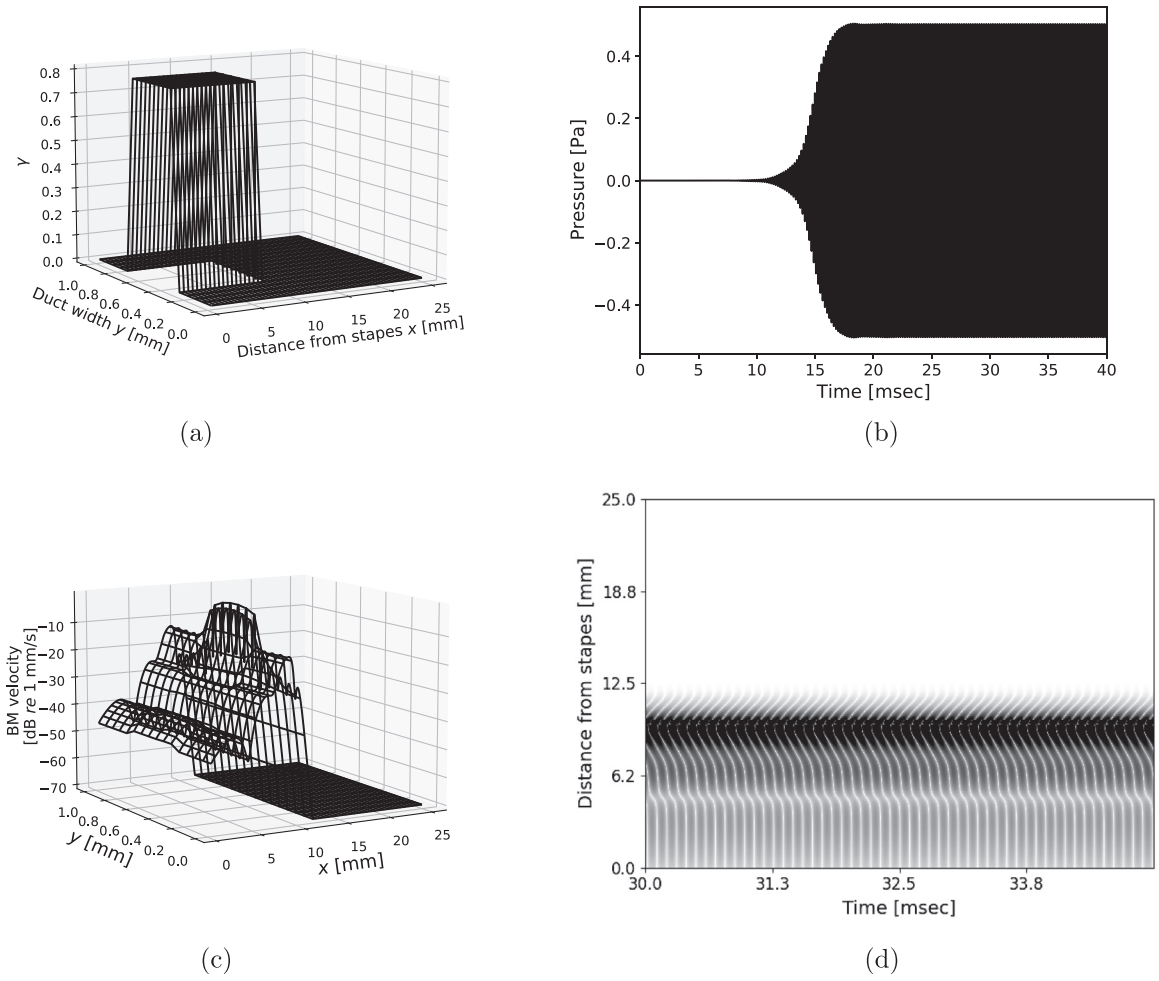

FIG. 7. (a) Gain distribution for simulated alignment of OHCs. (b) Temporal variation in pressure p for clicks at $x = 0$ and $y = 0.5$ mm. (c) Spatial pattern of BM envelope for clicks in steady state. (d) Temporal variation in BM velocity along the cochlear length at $y = 0.5$ mm.

pressure concentrated on the center of the cochlear width that was assigned OHCs.

IV. DISCUSSION

The main contribution of this study is the development of a fast time-domain solution for nonlinear models of the cochlea using the FFT. In the method proposed herein, we have applied the FFT-based solver to the macro-mechanical model, as described in Sec. II A. The direct solver included in the method proposed by Diependaal *et al.* (1987) was replaced with the FFT-based solver (Schumann and Sweet, 1988) as described in Sec. II C. Based on the results of the numerical simulations, Fig. 3 shows equivalent distributions of the BM velocities between the two methods. Notably, the computation times were reduced by applying the FFT-based solution, as shown in Fig. 4. In addition, the direct method could not solve the model under large number of grid segments because of the memory limitations of the computer used in the experiment.

The advantage of the proposed method is that it enables the modeling of 3D cochlear mechanics, as shown in Figs. 5–8, which has the potential to facilitate the discovery of unknown nonlinear phenomena evoked by natural sounds. The linear response characteristics of the BM and auditory nerve are approximately equal; however, the nonlinear response trends may differ (Geisler and Nuttall, 1997; Lin and Guinan, 2004). Moreover, recent optical coherent tomography (OCT) techniques have shown that velocity in the organ of Corti (OoC) varies with location (Olson and Strimbu, 2020). These results emphasize that BM tuning transforms neural tuning through the TM movement, acting as a second filter (Allen, 1980). Because the proposed method can simulate the 3D motion of the cochlea in the time domain, it is possible to investigate the TM motion as a second filter for generating cochlear nonlinearities.

Although observations by experiments as well as modeling studies have shown the necessity of 3D models, the spatial dimension in a fluid chamber influences numerical results when using the same macro-mechanical model of the cochlea (Elliott *et al.*, 2017). Higher-dimension models, such as 2D or 3D models, can account for experimental data better than 1D models (Elliott *et al.*, 2017). In addition, fluid viscosity can enhance the frequency selectivity of sound in the cochlear model (Prodanovic *et al.*, 2019). To obtain higher computational accuracy, nonuniform grid settings

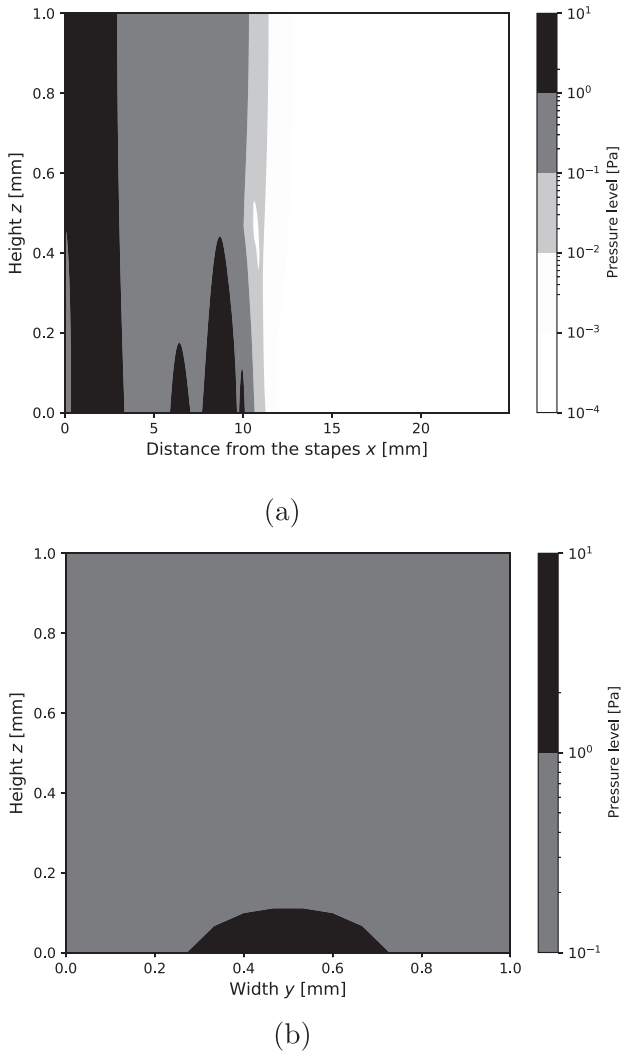

FIG. 8. Maximum pressure distribution in the chamber for the click under same condition of Fig. 7. (a) x–z plane at $y = 0.5$ mm. (b) y–z plane at $x = 10$ mm, which is the stepped gain location.

have been employed for simulating the dynamics of viscous fluids (Prodanovic *et al.*, 2019). These reasons suggest that large-scale cochlear modeling, such as 3D structures with viscous fluids, have the potential to reproduce experimental data more accurately and discover unknown nonlinear phenomena evoked by natural sounds.

Allen and Sondhi (1979) had proposed a solution method for the cochlear model using the FFT. This solution method was similar in concept to the proposed method solved in the spatial frequency domain. The main difference lies in the way the given governing equation, the Laplace equation, is solved. While the proposed method solves the given Laplace equation directly, Allen and Sondhi (1979) solved the Laplace equation by transforming the differential equation into the form of an integral equation. Although an ideal fluid is assumed in this paper, the proposed method can extend to simulate a real fluid, such as a viscous fluid, to solve the governing equations directly. In addition, the nonlinearized 2DoF model based on Neely and Kim (1986) is used as the micro-mechanical model of the cochlea. The method presented here can be modified to any micro-mechanical model with nonlinear elements, provided that the variables α and g obtained from the equations of motion shown in Eqs. (28) and (29) are set.

The proposed method does involve some limitations. The FFT-based method assumes a box-shaped cochlear duct because of the necessity to achieve a solution with equally spaced spatial bins using the FFT. However, a real cochlear duct has a coiled and tapered shape. Therefore, Altoè and Shera (2020) report that the tapered duct affects the tuning property in the cochlear responses. In this case, to mimic the complex shape of the duct, it is necessary to use the domain decomposition method (Tang *et al.*, 2021) to accurately model the boxed shape.

ACKNOWLEDGMENTS

This work was supported by Japan Society for the Promotion of Science KAKENHI Grant No. 21K17765. The author is greatly indebted to three anonymous reviewers for their helpful comments.

Allen, J. B. (**1977**). "Cochlear micromechanics–a mechanism for transforming mechanical to neural tuning within the cochlea," J. Acoust. Soc. Am. **62**(4), 930–939.
Allen, J. B. (**1980**). "Cochlear micromechanics—A physical model of transduction," J. Acoust. Soc. Am. **68**(6), 1660–1670.
Allen, J. B. "Invitation to mathematical physics and its history" (Springer-Nature, New York, 2020).
Allen, J. B., and Sondhi, M. M. (**1979**). "Cochlear macromechanics: Time domain solutions," J. Acoust. Soc. Am. **66**(1), 123–132.
Altoè, A., and Shera, C. A. (**2020**). "The cochlear ear horn: Geometric origin of tonotopic variations in auditory signal processing," Sci. Rep. **10**(1), 20528.
Ashmore, J. (**2019**). "Outer hair cells and electromotility," Cold Spring Harbor Perspectives Med. **9**(7), a033522.
Brownell, W. E., Bader, C. R., Bertrand, D., and De Ribaupierre, Y. (**1985**). "Evoked mechanical responses of isolated cochlear outer hair cells," Science **227**(4683), 194–196.
Cooper, N. P., Vavakou, A., and van der Heijden, M. (**2018**). "Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea," Nat. Commun. **9**(1), 3054.
Dewey, J. B., Applegate, B. E., and Oghalai, J. S. (**2019**). "Amplification and suppression of traveling waves along the mouse organ of Corti: Evidence for spatial variation in the longitudinal coupling of outer hair cell-generated forces," J. Neurosci. **39**(10), 1805–1816.
Diependaal, R., Duifhuis, H., Hoogstraten, H., and Viergever, M. (**1987**). "Numerical methods for solving one-dimensional cochlear models in the time domain," J. Acoust. Soc. Am. **82**, 1655–1699.
Elliott, S. J., Ku, E. M., and Lineton, B. (**2007**). "A state space model for cochlear mechanics," J. Acoust. Soc. Am. **122**(5), 2759–2771.
Elliott, S. J., Ni, G., and Sun, L. (**2017**). "Fitting pole-zero micromechanical models to cochlear response measurements," J. Acoust. Soc. Am. **142**(2), 666–679.
Epp, B., Verhey, J. L., and Mauermann, M. (**2010**). "Modeling cochlear dynamics: Interrelation between cochlea mechanics and psychoacoustics," J. Acoust. Soc. Am. **128**(4), 1870–1883.
Gao, S. S., Wang, R., Raphael, P. D., Moayedi, Y., Groves, A. K., Zuo, J., Applegate, B. E., and Oghalai, J. S. (**2014**). "Vibration of the organ of corti within the cochlear apex in mice," J. Neurophysiol. **112**(5), 1192–1204.
Geisler, C. D., and Nuttall, A. L. (**1997**). "Two-tone suppression of basilar membrane vibrations in the base of the guinea pig cochlea using 'low-side' suppressors," J. Acoust. Soc. Am. **102**(1), 430–440.
Goldstein, J. (**1967**). "Auditory nonlinearity," J. Acoust. Soc. Am. **41**(3), 676–699.

https://doi.org/10.1121/10.0006533

Goldstein, J., and Kiang, N. (**1968**). "Neural correlates of the aural combination tone $2f_1–f_2$," Proc. IEEE **56**(6), 981–992.
He, W., Kemp, D., and Ren, T. (**2018**). "Timing of the reticular lamina and basilar membrane vibration in living gerbil cochleae," eLife **7**, e37625.
Hudspeth, A. J. (**2008**). "Review making an effort to listen: Mechanical amplification in the ear," Neuron **59**, 530–545.
Kemp, D. T. (**1978**). "Stimulated acoustic emissions from within the human auditory system," J. Acoust. Soc. Am. **64**(5), 1386–1391.
Kim, D. O., Molnar, C. E., and Pfeiffer, R. R. (**1973**). "A system of nonlinear differential equations modeling basilar–membrane motion," J. Acoust. Soc. Am. **54**(6), 1517–1529.
Ku, E. M., Elliott, S. J., and Lineton, B. (**2008**). "Statistics of instabilities in a state space model of the human cochlea," J. Acoust. Soc. Am. **124**(2), 1068–1079.
Lee, H. Y., Raphael, P. D., Park, J., Ellerbee, A. K., Applegate, B. E., and Oghalai, J. S. (**2015**). "Noninvasive *in vivo* imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea," Proc. Natl. Acad. Sci. U.S.A. **112**(10), 3128–3133.
Lin, T., and Guinan, J. J. (**2004**). "Time–frequency analysis of auditory-nerve-fiber and basilar-membrane click responses reveal glide irregularities and non-characteristic-frequency skirts," J. Acoust. Soc. Am. **116**(1), 405–416.
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.
Murakami, Y. (**2018**). "Influence of discretization error on instability of cochlear model," J. Acoust. Soc. Am. **144**(6), 3593–3602.
Neely, S. T. (**1981**). "Finite difference solution of a two-dimensional mathematical model of the cochlea," J. Acoust. Soc. Am. **69**(5), 1386–1393.
Neely, S. T., and Kim, D. (**1986**). "A model for active elements in cochlear biomechanics," J. Acoust. Soc. Am. **79**, 1472–1480.
Olson, E. S., and Strimbu, C. E. (**2020**). "Cochlear mechanics: New insights from vibrometry and optical coherence tomography," Curr. Opin. Physiol. **18**, 56–62.
Orszag, S. A. (**1972**). "Comparison of pseudospectral and spectral approximation," Stud. Appl. Math. **51**(3), 253–259.
Pan, S., Elliott, S. J., Teal, P. D., and Lineton, B. (**2015**). "Efficient time-domain simulation of nonlinear, state-space, transmission-line models of the cochlea," J. Acoust. Soc. Am. **137**(6), 3559–3562.
Prodanovic, S., Gracewski, S. M., and Nam, J. H. (**2019**). "Power dissipation in the cochlea can enhance frequency selectivity," Biophys. J. **116**(7), 1362–1375.
Rhode, W. S. (**1971**). "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique," J. Acoust. Soc. Am. **49**(4), 1218–1231.
Schumann, U., and Sweet, R. A. (**1988**). "Fast Fourier transforms for direct solution of Poisson's equation with staggered boundary conditions," J. Comput. Phys. **75**(1), 123–137.
Shera, C. A. (**2003**). "Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves," J. Acoust. Soc. Am. **114**(1), 244–262.
Swarztrauber, P. N. (**1977**). "The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle," SIAM Rev. **19**(3), 490–501.
Tang, H. S., Haynes, R. D., and Houzeaux, G. (**2021**). "A review of domain decomposition methods for simulation of fluid flows: Concepts, algorithms, and applications," Arch. Comput. Methods Eng. **28**(4), 841–873.
Vencovský, V., Vetešník, A., and Gummer, A. W. (**2020**). "Nonlinear reflection as a cause of the short-latency component in stimulus-frequency otoacoustic emissions simulated by the methods of compression and suppression," J. Acoust. Soc. Am. **147**(6), 3992–4008.
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.