Fast time-domain solution of the cochlear transmission line model in real-time applications

Yasuki Murakami

Check for updates

View Online   Export Citation

Fast time–domain solution of the cochlear transmission line model in real–time applications

Yasuki Murakami[a] iD

Faculty of Design, Kyushu University, 4-9-1 Shiobaru, Minamiku, Fukuoka 815-8540, Japan

murakami@design.kyushu-u.ac.jp

Abstract: A fast numerical time-domain solution for a one-dimensional cochlear transmission-line model was proposed for real-time applications. In this approach, the three-dimensional solver developed by Murakami [J. Acoust. Soc. Am. **150**(4), 2589–2599 (2021)] was modified to develop a solution for the one-dimensional model. This development allows the solution to accurately and quickly calculate cochlear responses. The present solution can solve the model in real-time under coarse grid conditions. However, under fine-grid conditions, the computation time is significantly longer than the duration of the signal. Nevertheless, calculations can be performed under the fine grid condition, which previously required much computation time. This fact is essential to applications. © *2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).*

1. Introduction

In our auditory system, the cochlea functions as a natural time–frequency analyzer, segregating sounds into constituent frequencies and analyzing their temporal evolution. Researchers have developed a time–frequency analyzer similar to the cochlea, successfully emulating cochlear processing to electronically analyze sound signals (Baby *et al.,* 2021; Lyon and Mead, 1988; Mandal *et al.,* 2009; Xu *et al.,* 2018; Zwicker, 1986). This approach can potentially improve audio signal-processing systems in two principal domains: preprocessing and hearing aids. The preprocessing aspect could see significant improvements in frequency separation and temporal analysis. The analyzer enhances the performance of pre-processing stages in audio systems where signals are typically readied for further analysis or manipulation. In hearing aids, a bio-inspired analyzer can be used to improve fitting and potentially aid in diagnosing hearing impairments. Mimicking the natural processing of the cochlea can improve the effectiveness of hearing aids. A crucial criterion for this analyzer is real-time operation, ensuring that sound signals are processed promptly without significant delays. In essence, this bio-inspired analyzer aims to harness cochlear strengths for improved audio processing across diverse applications, with real-time functionality serving as a key determinant of its efficacy.

Cochlear models fall into two main categories: transmission-line (TL) models and filter bank models. TL models are grounded in the physical principles of the cochlea, treating it akin to an equivalent electrical circuit whose components are governed by specific equations. This approach encapsulates the processes involved in the propagation of sound waves through the cochlea and their subsequent conversion into electrical signals, which the brain perceives as sound. Conversely, filter bank models concentrate on the function or output of the cochlea, sidestepping the detailed inner mechanics. These models employ signal processing techniques to mimic how the cochlea filters segregate them into various frequencies, providing insights into the cochlear ability to achieve frequency selectivity without delving into the underlying causes. Comparisons of cochlear models, including both categories, reveal similar frequency selectivities but slightly different temporal feature actions (Osses Vecchi *et al.,* 2022; Saremi *et al.,* 2016). Given the differences in temporal characteristics observed, it is evident that there is not a singular, definitive model of the cochlea that is entirely precise. While both models excel in capturing certain aspects of cochlear function, neither accurately encapsulates the full complexity.

The strength of the TL model is based on the actual physical properties of the cochlea, making it particularly useful for comprehending and simulating hearing loss. Because it is rooted in the physics of the cochlea, the TL model is capable of incorporating the underlying mechanisms associated with otoacoustic emission (OAE) (Liu and Neely, 2010; Verhulst *et al.,* 2012; Wen and Meaud, 2022) and hearing loss (Verhulst *et al.,* 2018). However, solving the TL model incurs high computational costs (Diependaal *et al.,* 1987). To address this challenge, attempts have been made toward hardware implementations (Lyon and Mead, 1988; Mandal *et al.,* 2009; Xu *et al.,* 2018; Zwicker, 1986) and the use of neural network-based estimations (Baby *et al.,* 2021). Despite these advancements, challenges persist in terms of model

---

[a]Author to whom correspondence should be addressed.

construction efficiency and the complexity of reconstructing these models, especially when considering variations in model parameters that correspond to different degrees of hearing loss. Addressing these challenges often requires adjustments to the model equations and computational configurations to accurately reflect the condition being modeled.

Murakami (2021) proposed a fast Fourier transform (FFT)-based solution for fast solving the three-dimensional (3D) TL model in the time domain. This software-based approach is adaptable to a range of micromechanical models, offering flexibility in terms of altering parameters and reconstructing models. Building upon this foundational concept, the current study suggests a modification that transitions this fast-solving approach from the complex 3D TL model to a more simplified one-dimensional (1D) variant. Significantly, the method put forward in this study delivers both accurate and timely performance results.

2. Methods

2.1 FFT-based solution of a cochlear transmission–line model

The cochlear model consists of macromechanical and micromechanical parts. A TL model, representing macromechanics is derived from the law of conservation of mass and Newton's second law, governing fluid motion within the cochlear chamber. The proposed method assumes a rectangular-shaped cochlear chamber with a uniform grid setting. Equation of a 1D cochlear TL model that expresses the interaction between fluid pressure p and acceleration of the cochlear partition $\ddot{u}$ can be written as follows:

$$\frac{d^2p(t)}{dx^2} - \frac{2\rho}{H}\ddot{u}(t) = 0, \tag{1}$$

where ρ and H represent the fluid density and chamber height and t and x represent time and the longitudinal location. The Poisson equation for fluid pressure p is created to compute the cochlea model in time domain, derived from Eq. (1) (Diependaal *et al.*,1987). We then separate the variables related to fluid pressure and other pressures, which are included depending on the micromechanical model:

$$\frac{d^2p(t)}{dx^2} + \alpha p(t) + \alpha g(t) = 0, \tag{2}$$

where the variables α and g depend on the micromechanical model and are described in Sec. 2.2.

The boundaries are set at the cochlear entrance ($x = 0$) and the helicotrema ($x = L$). At the cochlear entrance site ($x = 0$), the boundary value is obtained under the Neumann condition:

$$\left.\frac{dp(t)}{dx}\right|_{x=0} = -2\rho\ddot{u}_s(t), \tag{3}$$

where $\ddot{u}_s$ denotes the acceleration of the stapes as the input stimulus. The boundary value at the helicotrema ($x = L$) is represented by the Dirichlet condition:

$$p(t)|_{x=L} = 0. \tag{4}$$

Diependaal *et al.* (1987) used the matrix production to solve the Poisson equation in Eq. (2). Murakami (2021) replaced the 3D model with the FFT-based Poisson solver (Schumann and Sweet, 1988). This FFT-based algorithm uses the discrete cosine transform (DCT) depending on the boundary conditions. In this study, we modified the FFT-based method from a 3D model to a 1D model.

By approximating the finite difference scheme in Eq. (2) in the longitudinal direction with spatial step Δ and number of segments N, the governing equation can be expressed as follows:

$$D_{xx}\big[p_n(t)\big] + \alpha p_n(t) + \alpha g_n(t) = 0, \tag{5}$$

where the operator D_{xx} represents the second-order numerical derivative in the x direction and the suffix n indicates the discrete location.

The mixed boundary conditions defined at $n = 1$ and N are given by Eqs. (3) and (4), respectively, as follows:

$$\frac{-p_1(t) + p_2(t)}{\Delta} = 2\rho\ddot{u}_s(t), \tag{6}$$

$$p_N = 0. \tag{7}$$

The cochlear model includes the mixed boundary conditions described in Eqs. (3) and (4). DCTs are categorized from type from I to IV, with the choice depending on the boundary conditions (Schumann and Sweet, 1988). Under mixed boundary conditions, the proposed method employs DCT-III.

The FFT-based method consists of three stages: analysis, computation, and synthesis stages (Schumann and Sweet, 1988). In the analysis stage, the DCT coefficient $\bar{g}_i(t)$ is obtained using DCT–III as follows:

$$\bar{g}_i(t) = \frac{1}{N} g_1(t) + \frac{2}{N} \sum_{n=1}^{N} g_n(t) \cos \frac{(n-1)(2i-1)\pi}{2N}, \tag{8}$$

where $i = 1, \dots, N$. The computation stage obtains the DCT coefficient of the pressure $\bar{p}_i$ from the DCT coefficient $\bar{g}_i$,

$$\frac{\lambda_i}{\Delta^2} \bar{p}_i(t) + \alpha \bar{p}_i(t) + \alpha \bar{g}_i(t) = 0, \tag{9}$$

where λ_i is an eigenvalue that can be calculated as follows:

$$\lambda_i = -4 \sin^2 \frac{(2i-1)\pi}{4N}. \tag{10}$$

Finally, the synthesis stage computes the DCT coefficient $\bar{p}_i(t)$ to the pressure difference $p_n(t)$ obtained from the inverse transforms,

$$p_n = \sum_{i=1}^{N} \bar{p}_i(t) \cos \frac{(n-1)(2i-1)\pi}{4N}. \tag{11}$$

2.2 Specific example

The proposed FFT-based solution for the 1D cochlear TL model was evaluated using the micromechanical model proposed by Neely and Kim (1986). This micromechanical model, a 2-degree-of-freedom (2DoF) model, includes active feedback to account for the fine frequency selectivity obtained from experimental data. The two masses m_1 and m_2 in the 2DoF system represent the basilar membrane (BM) and tectorial membrane, respectively. In particular, the pressure difference p in the macromechanical model drives the mass m_1, and the active pressure p_a affects the mass m_1 to amplify motion. The equation of motion derived from the micromechanical model is expressed as follows:

$$p + p_a = m_1 \ddot{u}_1 + (c_1 + c_3)\dot{u}_1 + (k_1 + k_3)u_1 - c_3\dot{u}_2 - k_3 u_2, \tag{12}$$

$$0 = -c_3\dot{u}_1 - k_3 u_1 + m_2 \ddot{u}_2 + (c_2 + c_3)\dot{u}_2 + (k_2 + k_3)u_2, \tag{13}$$

where $\ddot{u}_1 (=\ddot{u})$ and $\ddot{u}_2$ represent the accelerations of the two masses, $\dot{u}_1$ and $\dot{u}_2$ indicate their velocities, and u_1 and u_2 are their displacements. The symbols c and k correspond to the damping and stiffness coefficients, respectively. The active and nonlinear pressure p_a is defined as

$$p_a = c_4 \dot{u}_{nl} + k_4 u_{nl}, \tag{14}$$

where

$$u_{nl} = \tanh(u_2 - u_1). \tag{15}$$

Variables α and g in Eq. (2) depend on the macromechanical model, which can be derived from Eq. (12) as follows:

$$\alpha = \frac{2\rho}{m_1 H}, \tag{16}$$

$$g = (c_1 + c_3)\dot{u}_1 + (k_1 + k_3)u_1 - c_3\dot{u}_2 - k_2 u_2 + p_a. \tag{17}$$

As demonstrated in Eqs. (12) and (17), the displacement and velocity of the BM must satisfy the ordinary differential equation:

$$\ddot{u}_1(t) = \frac{p(t) - g(t)}{m_1}. \tag{18}$$

In addition, based on Eq. (13), variable g_2 can be defined as follows:

$$g_2 = -c_3\dot{u}_1 - k_3 u_1 + (c_2 + c_3)\dot{u}_2 + (c_2 + k_2)u_2, \tag{19}$$

where the displacement and velocity of the tectorial membrane satisfy

$$\ddot{u}_2(t) = -\frac{g_2(t)}{m_2}. \tag{20}$$

2.3 Numerical procedures

The proposed FFT-based method constitutes a modified time-domain solution for the cochlear TL model, as proposed by Diependaal et al. (1987), and it performs discretization using the finite difference approach, as described in Sec. 2.1.

Diependaal *et al.* (1987) employed the finite element approach for solving numerical problems. However, no modification to the numerical procedure was required to solve the cochlear TL model in the time domain.

The numerical procedure is outlined as follows:

(i) Discretize time series t into sequence t_i with the time step Δt.
(ii) Compute g in Eq. (17).
(iii) Solve p using DCT described in Eqs. (8)–(11).
(iv) Compute $\ddot{u}_1$ and $\ddot{u}_2$ in Eqs. (18) and (20), respectively.
(v) Integrate $\ddot{u}_1$ and $\ddot{u}_2$ to obtain $\dot{u}_1$, $\dot{u}_2$, u_1, and u_1 from t_{i-1} to t_i.

The proposed method is implemented using the FFT described in step (iii). The original version of the direct method uses the matrix product in step (iii). However, this step entails the highest computational cost of all steps. By replacing the matrix operation with an FFT, the computational complexity reduces from N^2 to $N \log N$ operations. This reduction is pivotal for the highly efficient computation of the proposed method. In addition, the direct method incurs a high computational cost in the preprocessing step, which calculates the inverse matrix using the Gaussian elimination method.

The computations were performed on A Linux-based personal computer with an Intel Corei9 12900 central processing unit and 64 GB of RAM. The programs were written in C++ programming language. In step (iii) of the FFT-based algorithm, FFTW_REDFT10 and FFTW_REDFT01 routines in the FFTW3 library were used. The DGEMV routine in OpenBLAS performs the matrix product in step (iii) of the original direct algorithm. In step (v) of the algorithm, we employed the fourth-order Runge–Kutta method for solving ordinary differential equations with a constant time step $\Delta t = 10[\mu s]$.

The values of the parameters were set to fit human data (Greenwood, 1990) based on the values of the original model proposed by Neely and Kim (1986), which was used to reproduce the auditory nerve response in cats.

3. Results

3.1 Accuracy

The BM responses for pure tones are used to evaluate frequency selectivity in cochlear mechanics. Figure 1 illustrates that the BM responses calculated using the FFT-based method and direct methods used by Diependaal *et al.* (1987) are similar. As an essential feature, the cochlear model exhibits a peak at the BM location depending on the tone input frequency. Moreover, the active pressure p_a causes the BM model to improve frequency selectivity and sensitivity compared to the low input level. This result matches the original simulation conducted by Neely and Kim (1986). Decreased frequency selectivity and sensitivity are observed when the input levels are increased. This phenomenon is called compressive nonlinearity in the experimental measurements (Cooper *et al.*, 2018; Fallah *et al.*, 2021).

To evaluate the quantitative difference between the FFT-based method and the direct method, the L2 norm was defined as follows:

$$L_2 = \sqrt{\frac{\sum \left(\dot{u}_n^{direct}(t) - \dot{u}_n^{fft}(t) \right)^2}{\sum \left(\dot{u}_n^{direct}(t) \right)^2}}, \tag{21}$$
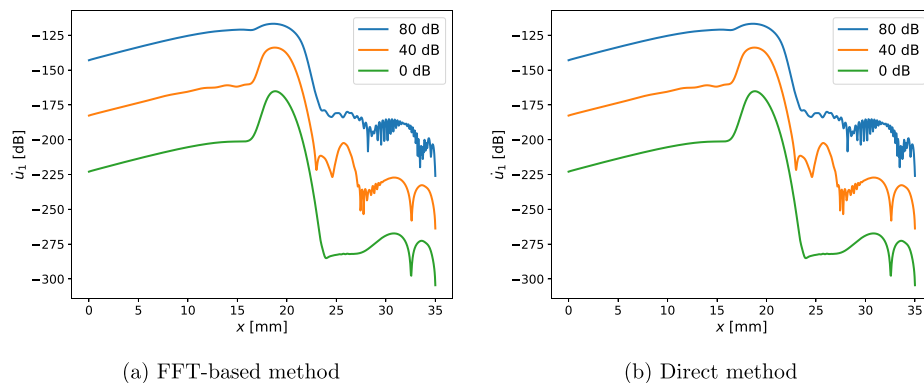

(a) FFT-based method   (b) Direct method

Fig. 1. Envelopes of BM velocity u_1 for 1000-Hz tone when the response had reached a steady state. Relative input level is varied with 0 to 80 dB in 40 dB increments.

Table 1. Difference in the BM velocity obtained from the FFT-based solution and the direct solution. The L_2 norm is defined by Eq. (21).

Input level (dB)	$N = 256$	$N = 512$	$N = 1024$
0	8.20×10^{-12}	5.11×10^{-11}	5.93×10^{-10}
40	3.39×10^{-11}	4.38×10^{-11}	5.20×10^{-10}
80	4.94×10^{-12}	8.91×10^{-12}	5.71×10^{-11}

where $\dot{u}_n^{direct}$ and $\dot{u}_n^{fft}$ denote the BM velocities obtained from the direct and FFT-based solutions, respectively. Table 1 shows the L2 norms of the calculated BM velocities by both solutions. The grid settings selected were 256 ($=2^8$), 512 ($=2^9$), and 1024 ($=2^{10}$) because several hundred segments were used in cochlear modeling and the OAE simulations conducted under the 1024 segments condition (Liu and Neely, 2010; Verhulst *et al.*, 2012), and the number of FFTs must be a power of 2. The differences were sufficiently small because the number of significant digits was 10 when the calculation was conducted using double-floating-point values. The number of significant digits in the floating-point representation is determined by the IEEE standard for floating-point arithmetic and is approximately 7 and 15 for single and double precision, respectively (Monniaux, 2008). This implies that the error is not negligible if the simulations were performed with double precision. However, it is worth noting that our simulations achieved better accuracy than those performed with single precision.

3.2 Computational cost

The computational speed is a crucial evaluation criterion for measuring the practicality of the FFT-based solution, which should focus on real-time computations. As illustrated in Fig. 2, the FFT-based solver is faster and more scalable than the direct solver. Under all conditions, the times obtained from the FFT-based solution were smaller than those obtained by the direct solution. The elapsed time is smaller than the input signal duration when the number of segments N is 256. The increments of the times are monotonical, approximately orders of $N \log N$ and N^2 for the FFT-based and direct solutions, respectively. This discrepancy arises from step (iii) of the algorithm, which uses the FFT or matrix product depending on the solutions. Thus, the FFT-based method can solve a large number of segment settings.

4. Discussion

The primary contribution of this study lies in the development of a fast time-domain solution for the 1D cochlear model, enabling real-time computation. This breakthrough is an extension of the 3D model solution introduced by Murakami (2021). The improved efficiency of our model is primarily due to the substitution of matrix multiplication with FFT computations, an approach initially used by Diependaal *et al.* (1987), as described in Sec. 2.1. While FFT has a computationally complexity of $N \log N$, matrix multiplication scales with a complexity of N^2. Theoretically, the FFT-based method produces equivalent outputs to the direct method, as demonstrated in Fig. 1 and Table 1. The FFT-based approach yields faster calculation speeds and operates in real-time under coarse conditions, as depicted in Fig. 2. However, under normal and fine grid scenarios, computational times still exceed the duration of the input signal. Despite this, the method has not yet achieved the ideal of spare-time real-time computation.

The equation for the TL model is the same for all cochlear models. It is derived from Newton's second law and the law of conservation of mass concerning the fluid flowing through the cochlear duct, as shown in Eq. (1), by assuming
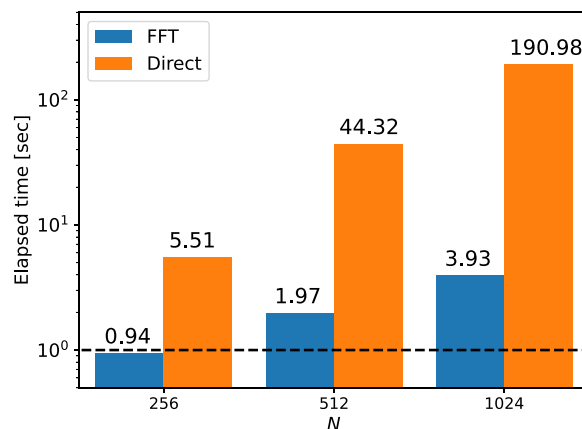

Fig. 2. Comparison of computation times. The computations were performed for 1 s of a tone.

an incompressible fluid with no viscosity. The cochlear model is constructed by combining this macro fluid model with a microcochlear model, as shown in Sec. 2.2. The microcochlear model can be a 1DoF oscillating system (Verhulst *et al.*, 2012), a 2DoF model used here, or a complicated model including an electrophysiological mechanism (Liu and Neely, 2010). It is important to emphasize that the proposed method can calculate any cochlear model, including the macro model, which is the TL form shown in Eq. (1).

The software-based method proposed here allows for easy modification of the cochlear model configuration. This straightforward modification is a significant advantage compared to hardware (Lyon and Mead, 1988; Mandal *et al.*, 2009; Xu *et al.*, 2018; Zwicker, 1986) and neural network-based methods (Baby *et al.*, 2021), which can be computationally expensive to retrain for different configurations. Another strong point is the ability to predict OAEs. Predicting OAEs can provide valuable information about the internal state of the cochlea (Liu and Neely, 2010; Verhulst *et al.*, 2012; Wen and Meaud, 2022). Combining these two aspects leads to the possibility of constructing personalized cochlear models.

Supplementary Material

See the supplementary material for the source code for the computer program required for the simulation.

Acknowledgments

Author Declarations
Conflict of Interest

The author declares no conflict of interest.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Baby, D., Van Den Broucke, A., and Verhulst, S. (**2021**). "A convolutional neural-network model of human cochlear mechanics and filter tuning for real-time applications," Nat. Mach. Intell. **3**(2), 134–143.

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**, 3054.

Diependaal, R. J., Duifhuis, H., Hoogstraten, H. W., and Viergever, M. A. (**1987**). "Numerical methods for solving one-dimensional cochlear models in the time domain," J. Acoust. Soc. Am. **82**(5), 1655–1666.

Fallah, E., Strimbu, C. E., and Olson, E. S. (**2021**). "Nonlinearity of intracochlear motion and local cochlear microphonic: Comparison between guinea pig and gerbil," Hear. Res. **405**, 108234.

Greenwood, D. D. (**1990**). "A cochlear frequency-position function for several species: 29 years later," J. Acoust. Soc. Am. **87**(6), 2592–2605.

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.

Lyon, R. F., and Mead, C. (**1988**). "An analog electronic cochlea," IEEE Trans. Acoust. Speech Signal Process. **36**(7), 1119–1134.

Mandal, S., Zhak, S. M., and Sarpeshkar, R. (**2009**). "A bio-inspired active radio-frequency silicon cochlea," IEEE J. Solid State Circuits **44**(6), 1814–1828.

Monniaux, D. (**2008**). "The pitfalls of verifying floating-point computations," ACM Trans. Program. Lang. Syst. **30**(3), 12.

Murakami, Y. (**2021**). "Fast time-domain solution of a nonlinear three-dimensional cochlear model using the fast Fourier transform," J. Acoust. Soc. Am. **150**(4), 2589–2599.

Neely, S. T., and Kim, D. O. (**1986**). "A model for active elements in cochlear biomechanics," J. Acoust. Soc. Am. **79**(5), 1472–1480.

Osses Vecchi, A., Varnet, L., Carney, L. H., Dau, T., Bruce, I. C., Verhulst, S., and Majdak, P. (**2022**). "A comparative study of eight human auditory models of monaural processing," Acta Acust. **6**, 17.

Saremi, A., Beutelmann, R., Dietz, M., Ashida, G., Kretzberg, J., and Verhulst, S. (**2016**). "A comparative study of seven human cochlear filter models," J. Acoust. Soc. Am. **140**(3), 1618–1634.

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.

Verhulst, S., Altoè, A., and Vasilkov, V. (**2018**). "Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss," Hear. Res. **360**, 55–75.

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.

Wen, H., and Meaud, J. (**2022**). "Link between stimulus otoacoustic emissions fine structure peaks and standing wave resonances in a cochlear model," J. Acoust. Soc. Am. **151**(3), 1875–1894.

Xu, Y., Thakur, C. S., Singh, R. K., Hamilton, T. J., Wang, R. M., and van Schaik, A. (**2018**). "A FPGA implementation of the CAR-FAC cochlear model," Front. Neurosci. **12**, 198.

Zwicker, E. (**1986**). "A hardware cochlear nonlinear preprocessing model with active feedback," J. Acoust. Soc. Am. **80**(1), 146–153.