Hearing Research 360 (2018) 55–75

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss

Sarah Verhulst [a, *], Alessandro Altoè [b], Viacheslav Vasilkov [c]

[a] WAVES, Dept. of Information Technology, Ghent University, Technologiepark 15, 9052 Zwijnaarde, Belgium
[b] Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University, P.O. Box 13000, FI-00076 Aalto, Finland
[c] Medizinische Physik, Oldenburg University, Carl-von-Ossietzky strasse 9-11, 26120 Oldenburg, Germany

ARTICLE INFO

Article history:
Received 17 July 2017
Received in revised form
17 December 2017
Accepted 23 December 2017
Available online 28 December 2017

ABSTRACT

Models of the human auditory periphery range from very basic functional descriptions of auditory filtering to detailed computational models of cochlear mechanics, inner-hair cell (IHC), auditory-nerve (AN) and brainstem signal processing. It is challenging to include detailed physiological descriptions of cellular components into human auditory models because single-cell data stems from invasive animal recordings while human reference data only exists in the form of population responses (e.g., otoacoustic emissions, auditory evoked potentials). To embed physiological models within a comprehensive human auditory periphery framework, it is important to capitalize on the success of basic functional models of hearing and render their descriptions more biophysical where possible. At the same time, comprehensive models should capture a variety of key auditory features, rather than fitting their parameters to a single reference dataset. In this study, we review and improve existing models of the IHC-AN complex by updating their equations and expressing their fitting parameters into biophysical quantities. The quality of the model framework for human auditory processing is evaluated using recorded auditory brainstem response (ABR) and envelope-following response (EFR) reference data from normal and hearing-impaired listeners. We present a model with 12 fitting parameters from the cochlea to the brainstem that can be rendered hearing impaired to simulate how cochlear gain loss and synaptopathy affect human population responses. The model description forms a compromise between capturing well-described single-unit IHC and AN properties and human population response features.

Contents

* Corresponding author.
E-mail addresses: s.verhulst@ugent.be (S. Verhulst), Alessandro.altoe@aalto.fi (A. Altoè), Viacheslav.vasilkov@uni-oldenburg.de (V. Vasilkov).

https://doi.org/10.1016/j.heares.2017.12.018

1. Introduction

Recent physiological and histological evidence has demonstrated that sensorineural hearing loss not only comprises inner and outer-hair-cell (IHC & OHC) deficits, but also includes auditory-nerve (AN) synapse damage (Kujawa and Liberman, 2009). This latter AN deficit has been called cochlear synaptopathy or "hidden hearing loss" because it does not affect threshold audiometry while resulting in supra-threshold hearing difficulties (Bharadwaj et al., 2014, 2015; Plack et al., 2014; Liberman et al., 2016). Because it is presently unclear how different aspects of sensorineural hearing loss interact to impact sound encoding along the ascending auditory pathway, there is an increased need for computational models of the human auditory periphery that include a representation of OHC functionality as well as that of the IHC-AN complex. These models can be used to study how specific deficits alter human physiological responses such as the auditory brainstem response (ABR) and envelope-following response (EFR) in order to develop sensitive hearing diagnostics or hearing-loss compensation strategies. However, computational modeling efforts face several challenges: (i) models should capture the functional aspects of normal-hearing sound processing faithfully before hearing impairment can be considered. (ii) Physiological/histological correlates of specific hearing deficits do not necessarily translate into a single model parameter, complicating the simulation of frequency-specific combinations of hearing deficits. Lastly, (iii) given species-specific differences in cochlear mechanics and encoded acoustic frequencies, it is unknown whether model stages that are based on invasive animal measurements represent the physiology of the human auditory periphery well. Historical models have compromised between these constraints in the development of human sound perception models (e.g., Dau et al., 1997; Meddis and O'Mard, 1997; Ewert and Dau, 2000; Zilany and Bruce, 2006; Jepsen et al., 2008; Jepsen and Dau, 2011; Jørgensen et al., 2013; Mao and Carney, 2015; Jürgens et al., 2016) and models of auditory-evoked brain potentials (e.g. Dau, 2003; Schaette and McAlpine, 2011; Rønne et al., 2012; Verhulst et al., 2015, 2016). The listed models differ in their description detail of the physiology of hearing, and the following sections provide an overview.

1.1. Cochlear mechanics and filtering

Most human auditory models use a functional, parallel filter bank that captures empirical human cochlear frequency-tuning (Q_{ERB}) properties at low stimulation levels. In an alternative approach, human cochlear filtering can also be simulated using a transmission-line architecture for basilar-membrane (BM) vibration (Verhulst et al., 2015, 2016). The main difference between the different cochlear models relates to how Q_{ERB} is modeled to vary with level, and the way in which longitudinal coupling is introduced. In this context, longitudinal coupling refers to how cochlear models account for phenomena emerging from the coupled architecture of the BM and surrounding fluids resulting in cochlear traveling waves. These phenomena include two-tone suppression (Ruggero et al., 1992), asymmetrical filter shapes (von Békésy, 1970; Rosen and Baker, 1994) and phase changes in BM responses near the best frequency for higher stimulation levels (Ruggero et al., 1997). Parallel filterbanks such as the dual-resonance nonlinear model (Meddis et al., 2001; Sumner et al., 2002) and preprocessors to the Zhang et al. (2001) and Zilany et al. (2009, 2014) models capture these coupling phenomena by summing the output of the cochlear filter response with that of a second, broader filter of the same CF. In contrast, transmission-line front-ends (e.g., Zweig, 1976; Neely and Kim, 1983; de Boer, 1980, Zweig 1991; Talmadge et al., 1998; Meaud and Grosh, 2010; Duifhuis, 2012; Verhulst et al., 2012) and other serially coupled oscillator models (e.g., Lyon, 2011) benefit from their natural cascaded architecture that results in the described coupling phenomena without introducing a second filter stage. A recent study has detailed how specific model architectures influence cochlear filtering by comparing the output of several cochlear models that are commonly used as preprocessors for human auditory periphery and perception models to a fixed set of stimuli (Saremi et al., 2016).

While the computationally heavy transmission-line models are not commonly used as preprocessors for sound perception models (e.g., Takanen et al., 2014; Pieper et al., 2016), they can be adapted to simulate reverse traveling waves (i.e., otoacoustic emissions or OAEs; e.g., Talmadge et al., 1998; Elliott et al., 2007; Moleti et al., 2009; Liu and Neely, 2010; Epp et al., 2010; Verhulst et al., 2012). OAEs form a non-invasive correlate of human cochlear filter tuning and OHC health (Manley and Fay, 2007) and thus offer a means to humanize the fitting parameters of the cochlear model stage. Further, as OAEs are not affected by IHC processing (Trautwein et al., 1996), cochlear filter tuning parameter can be constrained on the basis of OAE recordings, and separately from the parameters relating to IHC, AN and brainstem processing. The use of a transmission-line architecture thus yields a more transparent parameter fitting procedure as the limitations of the cochlear filtering stage are known upfront. Additionally, this process safeguards against overfitting the IHC-AN model parameters for which experimental data mimicking in-vivo environments are scarce. When using other (parallel and cascaded) cochlear models as input

to auditory periphery models, parameters can only be evaluated at the level of the AN (e.g., Zilany et al., 2009, 2014; Sumner et al., 2002) or at the very end for sound perception models (e.g., Dau et al., 1997; Jepsen et al., 2008) necessitating parameter fine-tuning at several processing stages to match the desired AN or perceptional auditory filter tuning data.

1.2. The IHC-AN synapse

While human cochlear frequency-tuning parameters can be validated using OAE recordings, parameters related to the functioning of the IHC-AN synaptic complex rely on in-vitro, whole-cell patch clamp measurements of the cellular structures and channel properties (e.g., Kros and Crawford, 1990; Johnson, 2015). For this reason, there are several controversies regarding the transduction of stereocilia vibration into AN action potentials (e.g., Goutman and Glowatzki, 2007 *vs.* Heil and Neubauer, 2010; Grant et al., 2010 *vs.* Chapochnikov et al., 2014). Furthermore, there is a substantial heterogeneity in the biophysical properties across different synapses connecting to the same IHC (Frank et al., 2009; Ohn et al., 2016), rendering it difficult to develop precise and quantitative models of the synapse.

Fortunately, single-unit recordings from the AN are reliable and have over the years resulted in a large collection of AN responses to basic auditory stimuli in cats and small rodents (e.g., Kiang et al., 1969; Liberman, 1978; Rhode and Smith, 1985; Winter and Palmer, 1991; Sachs and Abbas, 1974; Relkin and Doucet, 1991; Joris and Yin, 1992). There are several models of AN signal processing that describe the instantaneous AN firing rate as resulting from the depletion and replenishment of different neurotransmitter stores (i.e., the three-store diffusion model; Meddis, 1986; Westerman and Smith, 1988). The neurotransmitter release is either driven by the cell's receptor potential V_{IHC} (Zhang et al., 2001; Heinz et al., 2001; Zilany and Bruce, 2006; Zilany et al., 2009, 2014; Verhulst et al., 2015) or by the calcium influx at the synapse (Kidd and Weiss, 1990; Sumner et al., 2002, 2003; Meddis, 2006). All mentioned AN synapse models simulate the instantaneous neurotransmitter release as the product of a driven rate (or permeability) and the number of vesicles in the ready-to-release pool (RRP). In this framework, adaptation of the simulated AN responses results from the depletion and diffusion of neurotransmitter material between the different neurotransmitter stores. The listed models of the AN synapse differ in their detailed implementation, but were shown to share the same functionality (Zhang and Carney, 2005) as they were all designed to match a rapid (~2 ms) and a short-term (60 ms) AN adaptation component to sustained tone stimulation (Westerman and Smith, 1984).

However, modeling AN adaptation properties using a classical three-store diffusion model cannot predict the strong asymmetries observed between adaptation and recovery properties of recorded AN responses (see Zhang and Carney, 2005). A solution that captures both on- and offset AN adaptation properties was found in the introduction of a power-law function that follows the exponential adaptation introduced in the three-store diffusion model (Zilany et al., 2009). While this solution yields satisfactory simulated AN adaptation properties, the functional description does not relate back to specific physiological processes that may cause such power-law like adaptation behavior in recorded AN responses. Another unknown in the IHC-AN complex relates to how fibers of different spontaneous rates (SR) should be modeled. It is known that the morphology (Meyer et al., 2009; Liberman et al., 2011) and the biophysical properties (Frank et al., 2009; Grant et al., 2010; Ohn et al., 2016) of IHC-AN synapses vary across the same IHC and that these differences result in altered functionality across AN fiber types (e.g., peak-to-steady state ratios and maximum sustained firing; Liberman, 1978; Rhode and Smith, 1985; Taberner and Liberman, 2005). Approaches to capture properties of different SR fibers in diffusion-based AN models have increased the threshold relating the electrical excitation of the synapse and firing rate (V_{IHC} in Verhulst et al., 2015, and $[Ca^{2+}]$ threshold in Sumner et al., 2002), while reducing the parameters controlling the maximal excitation of low SR synapses. These latter parameters refer to the permeability in the immediate store (Zilany et al., 2009, 2014; Verhulst et al., 2015 models), the conductance of the synaptic Ca^{2+} channels (Sumner et al., 2002 model), and the time constant of Ca^{2+} buffers controlling the magnitude of $[Ca^{2+}]$ (Meddis, 2006 implementation). While all discussed models qualitatively capture SR-dependent AN adaptation, peak-to-steady-state and sustained firing rate properties, they follow different approaches to reach this result. This is partly due to the relatively large number of five fitting parameters in the IHC-AN synapse model (Westerman and Smith, 1988). It is worthwhile to reduce this number of fitting parameters using recent IHC ribbon synapse data (Meyer et al., 2009; Pangršič et al., 2010) to unify these different ways in which SR-dependence can be introduced. Lastly, AN models require a cochlear mechanics stage to simulate sound-driven AN responses. Because there are numerous differences in cochlear tonotopy (Greenwood 1990), and frequency tuning across species (Joris et al., 2011), it is dangerous to fit human AN models using data from other species. The evaluation of the IHC-AN complex should ideally be conducted for cochlear mechanical models that are tailored to the species from which the AN recordings were taken.

1.3. Phase-locking

Even though advances have been made in the description of the biophysics of IHC transduction (see Altoè et al., 2017 for an overview), almost all AN models predict the driven firing rate as resulting from a cascade of nonlinearities followed by a low-pass filtering action (Sumner et al., 2002; Zhang et al., 2001; Zilany et al., 2009, 2014; Verhulst et al., 2015). The low-pass filter accounts for the observed loss in phase-locking to the temporal fine structure of sound as the characteristic frequency (CF) increases (Sellick and Russell, 1980; Russell and Sellick, 1983; Palmer and Russell, 1986). Auditory perception models simplify the IHC transduction process even further by representing it as a half-wave rectifier followed by a low-pass filter (e.g., Jepsen et al., 2008). Although it has been known for ~30 years that IHC transduction is affected by the dynamics of voltage-dependent basolateral K^+ currents (Hudspeth and Lewis, 1988; Kros and Crawford, 1990), AN modelers have so far refrained from introducing these concepts. IHC basolateral K^+ channel properties are known to affect the V_{IHC} dynamic range, onset/offset response and frequency response (Zeddies and Siegel, 2004; Lopez-Poveda and Eustaquio-Martín, 2006; Altoè et al., 2017, submitted) as well as the statistical properties of AN interspike intervals (Moezzi et al., 2016). Hence, it is worthwhile including the description of IHC basolateral K^+ channels in AN models because a misrepresentation or over-simplification of the IHC transduction process affects simulated phase-locking and AN adaptation properties. Consequently, this may lead to an overfitting of AN synapse parameters that should have been associated with IHC transduction.

1.4. Evaluating human models of the auditory periphery

To evaluate human models of AN processing, a compromise needs to be made between precisely matching AN recordings from a particular CF and species, and qualitatively matching a multitude of AN features across a wide range of CFs. The latter, more generic approach has over the years yielded successful AN models that

capture a large range of across-species functional characteristics such as phase-locking, modulation characteristics, SR-dependent firing thresholds, maximum AN firing rates and peak-to-steady state ratios (Sumner et al., 2002; Zilany et al., 2009, 2014; Verhulst et al., 2015). However, there remain a number of fitting parameters in those models that cannot be traced back to physiological processes, and that should be attributed to quantifiable parameters in order to unify functional and biophysical descriptions of the IHC-AN complex.

The final evaluation of human AN models should be made on the basis of simulated population responses that can be compared to non-invasively recorded human responses. Population responses constitute the compound action potential (CAP) that can be recorded using a round window electrode, or auditory evoked potentials that are recorded using scalp electrodes. Examples include the auditory brainstem response (ABR), frequency-following response (FFR), envelope-following responses (EFRs), auditory-steady-state response (ASSR) and complex ABR to speech stimuli (cABR). Human population response metrics are altered due to OHC (ABR: Picton, 2011; Gorga et al., 1984), IHC and AN deficits (or synaptopathy; ABR: Furman et al., 2013; Lin et al., 2011; Mehraei et al., 2016; Möhrle et al., 2016; EFR: Shaheen et al., 2015), and hence provide a human physiological metric that can validate hearing-impaired models of the human periphery. An additional benefit of using population responses over sound perception characteristics is that the influence of cortical processing can be neglected, thereby avoiding the need for optimal detector or perception schemes at the model back-end to allow for evaluation. However, it must be noted that existing population response models (Dau, 2003; Rønne et al., 2012; Bourien et al., 2014; Verhulst et al., 2015) rely on the summation of activity across all simulated CFs as a proxy to scalp-recorded population responses. The brainstem morphology studies of Melcher and Kiang (1996) justify this approach, even though in reality a population of neurons generate an electrical dipole of which the far field potential is reflected in the scalp recordings. In this framework, the ABR wave-I reflects the summed activity of all AN fibers to transient stimuli and the ABR wave-V reflects the summed activity of inferior colliculus (IC) neurons across the represented cochlear CFs. The limitation of this approach is that the absolute values of the summed responses no longer directly relate to physical and measurable units. However, it remains possible to evaluate how simulated population responses *change* as a function of stimulus characteristic alterations (e.g. level, amplitude modulation, clicks *vs* tone-bursts) or as a result of hearing deficits.

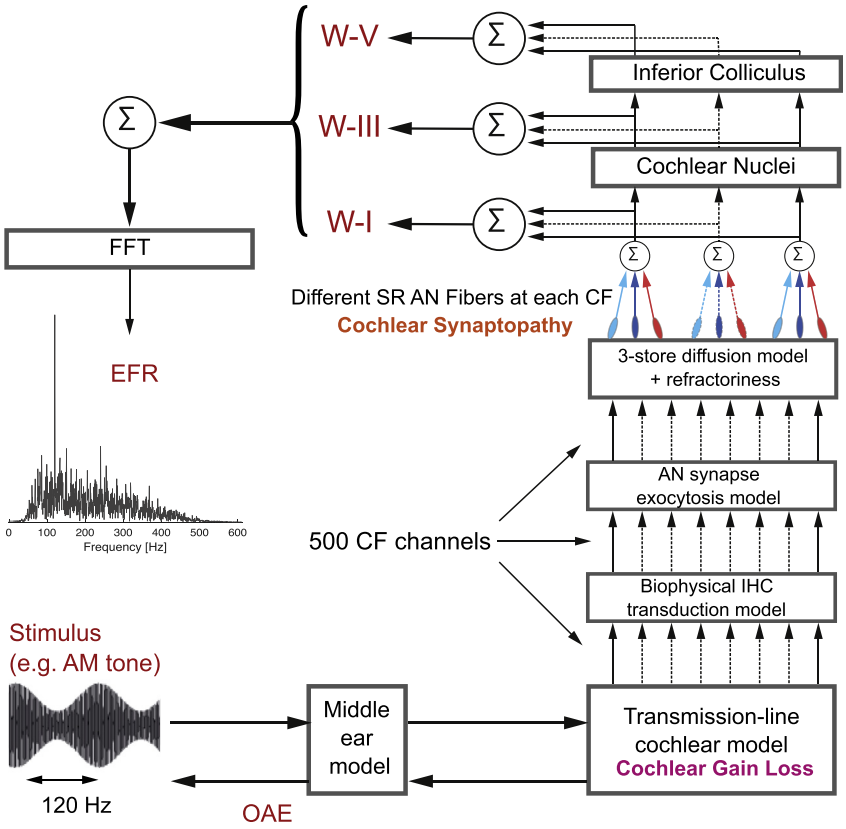

Fig. 1. Overview of the human auditory periphery model. The stimulus pressure passes through a first order middle-ear bandpass filter after which it enters a transmission-line cochlear model that includes human stimulus-frequency OAE estimates of tuning and cochlear compression. The transmission-line model simulates OAEs that can be compared to human data at the middle-ear filter output, as well as v_{BM} waveforms across 1000 Greenwood map spaced cochlear sections. Half of the simulated cochlear sections are passed on to the IHC-AN synaptic complex model that includes a biophysical description of the IHC membrane potential, a synaptic exocytosis model as well as a three-store diffusion and refractoriness model that generates AN firing rates. At each CF, rates of 19 fibers (13 HSR, 3MSR, 3LSR) are summed to yield r_{AN}. r_{AN} waveforms between 112 Hz and 12 kHz are summed to yield the ABR wave-I. r_{AN} is passed through a same-frequency bushy cell model to yield r_{CN} and r_{IC} modeling cochlear nucleus and inferior-colliculus generators to the ABR wave-III and wave-V, respectively. To model the EFR, simulated W-I, W-III and W-V waveforms are weighted according to human ABR wave ratios, after which they are added and a Fourier transform is applied to yield the response component relating to the modulation frequency of the EFR stimulus. The model can be rendered hearing impaired by changing the amount of mechanical gain in the cochlea on a CF-dependent basis yielding wider cochlear filters that simulate the effect of OHC loss. Cochlear synaptopathy is simulated by altering the number and types of auditory-nerve fibers that connect to each IHC.

1.5. Towards a comprehensive model of the human auditory periphery

This study aims to integrate biophysical models of the IHC-AN complex with a transmission-line cochlear model to improve existing functional models of the human auditory periphery in simulating a range of human population response characteristics and their alterations as a function of stimulus characteristics and hearing deficits. In this approach, we build on successes of earlier ABR models (Dau, 2003; Rønne et al., 2012; Verhulst et al., 2015) that can capture a realistic 1–1.2 ms ABR wave-V latency decrease for a 40-dB stimulus level increase (Prosser and Arslan, 1987; Jiang et al., 1991; Serpanos et al., 1997) when using a transmission-line front-end (Verhulst et al., 2015). We improve the latter human ABR model by incorporating a biophysical description of IHC transduction that accounts for single-unit AN adaptation and response properties to amplitude-modulated stimuli. The adopted approach is similar to that of Meddis and colleagues (e.g., Sumner et al., 2002; Meddis, 2006) who constrain the IHC-AN model equations based on IHC physiology. However, we make three important improvements: (i) the transmission-line model of the cochlea is based on the description by Zweig (1991) to offer a physically plausible cochlear mechanical description as input to the IHC model stage. (ii) The model parameters are constrained on the basis of recently quantified IHC properties (e.g., Johnson and Marcotti, 2008; Jia et al., 2007; Johnson et al., 2011; Pangršič et al., 2010) and resulting in fewer fitting parameters. (iii) By establishing a relationship between cochlear gain loss and the pole of the BM admittance function, OHC gain loss is adjusted using a single parameter at each CF. To summarize, we merge the success of the broadband ABR model with improved single-unit AN adaptation and modulation properties to yield a model framework that is well positioned to simulate human sustained population responses for pure-tone (FFR) and amplitude-modulated stimuli (EFR) for normal and hearing-impaired listeners. Along the path to finding this model compromise, we bear in mind that the ideal model describes key auditory processing features and its impairments with the fewest possible number of parameters.

2. The model compromise

We present a model of the human auditory periphery (Fig. 1) that captures a range of key principles in hearing that have been described in a number of independent studies. The model outputs single-unit simulations of AN, cochlear nucleus (CN) and inferior colliculus (IC) neurons for different CFs as well as physiological correlates of human peripheral hearing such as ABRs, EFRs and OAEs. While human OAE recordings were used to set and validate the cochlear mechanics model, simulated ABR and EFR responses were not designed to match the reference data directly. Rather, simulated brainstem responses were used as a metric to evaluate the broadband cochlear mechanical and IHC-AN characteristics of the human auditory periphery model.

2.1. Broadband cochlear mechanics

There are three cochlear mechanical principles that allow for a parameter reduction in the basilar-membrane (BM) equations describing cochlear frequency tuning and level-dependent compressive response growth: (i) assuming that the BM traveling wave wavenumber and envelope are approximately constant across CF (Zweig, 1976) (ii) modeling the cochlea as having a resistive cochlear input impedance Z_0 (i.e., the load impedance seen by the middle ear; Nedzelnitsky, 1980; Lynch et al., 1982; Shera and Zweig, 1991), and lastly (iii), modeling the level-dependence of cochlear frequency tuning and response growth such that the zero-crossings in local BM impulse responses remain intensity invariant (Recio and Rhode, 2000; Shera, 2001; Zweig, 2016). In transmission-line cochlear models, the second model principle governs the relationship between the longitudinal coupling strength and the properties of the CF-dependent BM stiffness, damping and mass parameters. While these latter parameters can be set based on BM impulse-response or frequency selectivity measurements (e.g., Allen and Sondhi, 1979; de Boer, 1980; Neely and Kim, 1983; Meaud and Grosh, 2010), the longitudinal coupling parameters can be chosen to remain constant or to vary as a function of CF to yield a capacitive (complex; e.g., de Boer, 1980; Van Hengel et al., 1996; Epp et al., 2010) or resistive (real; Shera and Zweig, 1991) Z_0, respectively. To model a resistive Z_0, Shera and Zweig (1991) introduced the concept of cochlear tapering that maintains a proportionality between the cochlear fluids (i.e. the serial coupling mass) and the compliance of the BM. In a resistive Z_0 model, cochlear traveling waves undergo little reflections in either direction, thereby limiting the generation of low-frequency standing waves that are known to exist for capacitive Z_0 models (Puria and Allen, 1991). Even though the tapering concept was also described in the earlier cochlear transmission-line models of Verhulst et al. (2012, 2015), these models suffered from a numerical error (going from Eq. B21 and B.22 in Verhulst, 2010) that resulted in a reactive (imaginary) component in Z_0.[1] The present model follows the implementation outlined in Altoè et al. (2014) to yield a resistive Z_0 and describes the BM as cascade of tapered series impedances (Z_{s_n}) and shunt admittances (Y_{p_n}) (Zweig, 1991; Verhulst et al., 2012, 2015):

$$Z_{s_n}(s_n) = \omega_n M_{s_0} s_n, \tag{1}$$

$$Y_{p_n}(s_n) = s_n \left[\omega_n M_{p_0} \left(s_n^2 + \delta_n s_n + 1 + \rho_n e^{-\mu_n s_n} \right) \right]^{-1}. \tag{2}$$

Subscripts s and p represent serial and shunt parameters respectively, and n refers to one of the 1000 discretized cochlear sections that span the human range of hearing according to the tonotopic Greenwood (1990) function. The scaling symmetric variable $s = i\omega/\omega_c$ is given by the frequency normalized by the location-dependent characteristic frequency of the cochlear oscillators. δ, ρ and μ describe the damping, stiffness strength and feedback delay respectively and are determined by the principal "double" pole α^* of the BM admittance. α^* is responsible for the tuning characteristics of the cochlear filters. The parameters M_{s_0} and M_{p_0} represent the longitudinal acoustic mass of the cochlear fluids and the phenomenological mass of the BM admittance respectively (see Table I of Verhulst et al., 2015).

The third cochlear mechanics principle that yields a simplified description of how cochlear frequency tuning changes as a function of level is that of intensity-invariant BM impulse response zero-crossings. The linear cochlear impedance equations of Zweig (1991) can be modified by moving the double-pole of the BM admittance α^* horizontally in the s-plane yielding broader cochlear filters as the poles are moved further away from the imaginary axis in response to stimulus intensity changes, while the BM zero-crossings remain in the same position (Shera, 2001; Verhulst et al., 2012). α^* determines the δ, ρ and μ parameters in Eq. (2)

[1] Consequently, the cochlear model implementations of Verhulst et al. (2012, 2015) showed (i) linear basal cochlear mechanics up to relatively high SPLs even though the nonlinearity threshold was set to 30 dB SPL, (ii) a reduced mechanical excitation in the cochlear base yielding high AN fiber thresholds for high CF units, and (iii) a distorted shape of high-CF tuning curves caused by the low-pass characteristics of the non-tapered cochlea.

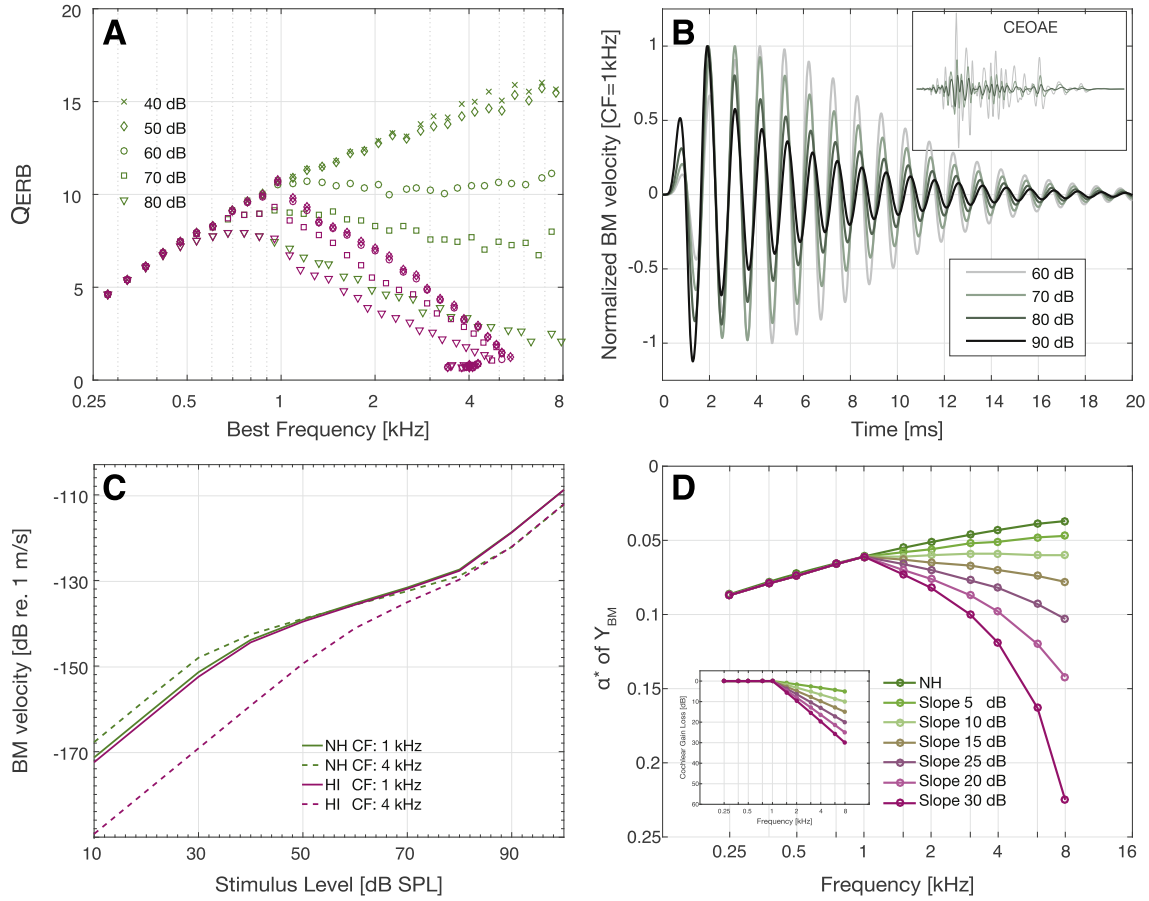

Fig. 2. Cochlear mechanical features of the human auditory periphery model. **A**. Q_{ERB} values computed from the energy underneath the power spectrum of local BM impulse responses to 80-μs clicks of different intensities. Simulations are shown for the normal-hearing model (green) and for a hearing-impaired model that had a sloping high-frequency audiogram (red: corresponding to the slope 30 dB profile in panel D) **B**. Simulated and normalized v_{BM} impulse response waveforms (CF = 1 kHz) and reflection-source CEOAE waveforms (inset) for different intensities of a 80-μs click. Waveform zero-crossings are well preserved for the BM IRs across the intensity range whereas the CEOAEs show a small latency increase of zero-crossings for the highest stimulus levels that corroborates observations from human CEOAE recordings. **C**. Simulated v_{BM} input/output functions (rms of the v_{BM}) for CFs of 1 (solid) and 4 kHz (dashed) in response to pure-tone stimuli of the same CF. Simulations are shown for both the normal-hearing (NH; green) and hearing-impaired (HI; red) model and shows that the 4-kHz CF yields a more linearized input-output function due to the cochlear gain loss at that CF. **D**. CF-dependent profiles of α^*_A corresponding to the sloping high-frequency audiograms in the inset. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

using a derivation from Shera (2001):

$$\delta = 2(\alpha^* - a),$$

$$\mu = (2\pi a)^{-1},$$

$$\rho = 2a\sqrt{1-(\delta/2)^2}e^{(-\alpha^*/a)}, \tag{3}$$

$$a = c^{-1}\left(\alpha^* + \sqrt{\alpha^{*2} + c(1-\alpha^{*2})}\right),$$

$$c = 120,8998691636393.$$

The α^*-trajectory can be determined as a function of level to model compressive BM-response growth, and as a function of CF to account for human cochlear frequency tuning (Shera et al., 2010). At each CF, the fitting parameters in this process are: (i) the active pole α^*_A corresponding to low-level stimulation at which cochlear filters are at their sharpest, (ii) the passive pole α^*_P corresponding to broader cochlear filters for high-level stimulation or post-mortem conditions. (iii) The intensity-dependent growth of v_{BM} was described using a compression slope (C) of 0.31 dB/dB. (iv) The compression threshold ($v_{BM,30dB}$) describes the v_{BM} above which v_{BM} grows compressively with increasing stimulus levels. $v_{BM,30dB}$ was determined from the peak v_{BM} in response to a 30-dB SPL 1-kHz pure tone evaluated at CF. The instantaneous nonlinearity function that describes the α^*-trajectory as a function of v_{BM} follows a hyperbolic interpolation method (Verhulst et al., 2012). In particular, when v_{BM} increases from $v_{BM,30dB}$ to $v_{BM,P}$, the pole α^* increases between α^*_A and α^*_P through the following equation:

$$\alpha^* = \alpha^*_A + \frac{x_p}{A}\sin\theta + \frac{y_p}{A}\cos\theta, \tag{4}$$

where θ is given by:

$$\theta = 0.5\cdot\tan^{-1}\left(\frac{A(\alpha^*_P - \alpha^*_A)}{\frac{v_{BM,P}}{v_{BM,30dB}} - 1}\right), \tag{5}$$

using a smoothing factor of $A = 100$. $v_{BM,P}$ is calculated from the

fitting parameter $v_{BM,30dB}$ to obtain the desired v_{BM} compression slope C. x_p normalizes and transforms the input v_{BM}, while y_p is given by the conjugate hyperbola function:

$$\left(\frac{x_p}{a_p}\right)^2 - \left(\frac{y_p}{b_p}\right)^2 = -1, \tag{6}$$

$$x_p = \left(\frac{|v_{BM}|}{v_{BM,30dB}} - 1\right)\frac{\cos\theta}{\cos 2\theta}. \tag{7}$$

a_p and b_p describe the slope of the asymptote of the hyperbola based on the focal point position (F_p):

$$F_p = \frac{A\,\alpha_A^*}{v_{BM,P}/v_{BM,30dB}},$$

$$a_p = F_p \cos\theta, \tag{8}$$

$$b_p = F_p \sin\theta.$$

Note that the α^*-trajectory function increases as the input v_{BM} increases, to yield a progressively more damped Y_p.

The CF-dependence of α^*_A was described by fitting a power-law function to the human cochlear frequency-tuning curve as a function of CF (Shera et al., 2010), yielding:

$$Q_{ERB,n} = 11.46(CF_n/1000)^{0.25}, \tag{9}$$

where α^*_A for the 1-kHz CF ($\alpha^*_A = 0.062$) corresponds to a Q_{ERB} value of 11.5 at low stimulus levels (Fig. 2A). In contrast to earlier implementations, the passive pole α^*_P was set to 0.305 instead of 0.7 (Verhulst et al., 2012) or 0.35 (Verhulst et al., 2015) because for very passive poles (i.e., large α^*_P) the feedback delay in Eq. (2) becomes too long and causes a violation of the intensity-invariant zero-crossings. This troubled the model simulations in Verhulst et al. (2015) to some degree, causing simulated BM impulse responses and click-evoked OAE (CEOAE) waveforms that did not preserve the zero-crossing invariance at the highest stimulation levels. As Fig. 2B shows, the adopted α^*_P yields simulated BM impulse responses and CEOAE waveforms that preserve their zero-crossings for stimulus level increases, in agreement with experimental observations (e.g., BM: Recio and Rhode, 2000; CEOAE: Kapadia and Lutman, 2000; Verhulst et al., 2011). However, as α^*_A is derived from OAE-based Q_{ERB} estimates and hence constrained, the total dynamic range of cochlear mechanical gain that can be represented in the model depends on α^*_P as well. A compromise between zero-crossing preservation and overall dynamic range was reached in the choice of the 0.305 α^*_P value. Fig. 2C depicts simulated v_{BM} input-output functions in response to 1-kHz and 4-kHz pure-tones evaluated at CF and shows that using our compromise, compressive growth can be achieved for stimulation levels up to 80 dB SPL, while preserving CEOAE zero-crossings throughout the level range. The cochlea of small rodents behaves compressively over wider dB ranges than predicted by the present model (see e.g. Rhode, 2007). It is presently not clear whether this is due to species-specific differences in dynamic range, or whether it is a limitation of the adopted model architecture, where the maximal achievable compression is determined by α^*_A and α^*_P.

2.2. Simulating outer-hair-cell deficits

α^*_A can be adjusted in a CF-dependent manner to simulate cochlear gain loss. The hearing deficit that is modeled this way, is associated with stereocilia damage, the actual loss of OHC bodies, or a metabolic reduction of the gain properties of OHCs (e.g., due to presbycusis) that are all known to reduce cochlear gain. Because cochlear gain and Q_{ERB} are tightly connected in the adopted transmission-line model (Eqs. (1) and (2)), α^*_A can be set to reflect a BM gain reduction corresponding to a specific hearing sensitivity loss [in dB HL]. To find the trajectories that relate α^*_A to different degrees of cochlear gain loss at each CF, frequency responses of simulated v_{BM} impulse responses were computed for a large range of α^*_A values. The power spectra describe local BM filter tuning, and their maxima were tracked to describe the difference in dB between the maximum of the considered α^*_A power spectrum and that of the corresponding normal-hearing model at the same CF. A table was constructed that relates filter gain reductions for each considered α^*_A and CF, such that any desired dB HL loss within the audiometric frequency range could be translated into an associated hearing-impaired α^*_A (Fig. 2D). Because filter gain and bandwidth are related through α^*, the BM filters of the cochlear gain loss models also show smaller Q_{ERB}'s. The hearing-impaired models adopt α^*_A profiles that correspond to specific CF-dependent dB HL profiles (or audiograms) but keep all other parameters controlling the BM nonlinearity (Eqs. (3)–(9)) fixed. Examples of α^*_A-functions for several high-frequency sloping hearing loss profiles are shown in Fig. 2D for losses above 1 kHz and slopes corresponding to the labeled dB HL value at 8 kHz. Note that it is not possible to simulate cochlear gain losses greater than the amount of v_{BM} filter gain that was present in the normal-hearing model (i.e., about 35 dB gain for the sharpest high-CF filters). Fig. 2A and C (red traces) show simulated Q_{ERB}'s and cochlear input-output functions for a hearing-impaired model in which the α^*_A values corresponded to a high-frequency sloping hearing loss above 1 kHz (30 dB HL at 8 kHz; slope 30 dB in Fig. 2D).

2.3. Inner-hair-cell transduction

We employ the biophysical IHC model presented in Altoè et al. (2017), that includes basolateral outward K^+ currents which were introduced in earlier circuit-models of the IHC (Zeddies and Siegel, 2004; Lopez-Poveda and Eustaquio-Martín, 2006). Appendix A details the IHC model description and the adopted parameters were based on recent IHC conductance measurements (e.g., Jia et al., 2007; Johnson et al., 2011) to improve on these earlier IHC models. Functionally, the adopted circuit model is different from the prevailing electrical low-pass filter description of IHC membrane characteristics that is commonly employed in auditory models (e.g., Jepsen et al., 2008; Zilany et al., 2014; Rønne et al., 2012; Verhulst et al., 2015). In fact, including V_{IHC}-dependent outward K^+ currents turns the IHC basolateral membrane into a nonlinear system rather than into a linear electrical low-pass filter (Davis, 1965; Shamma et al., 1986; Zhang et al., 2001; Sumner et al., 2002).

The dynamics of outward IHC K^+ currents have an important role in shaping AN response properties (Altoè et al., submitted). For example, they compress IHC responses dynamically and thereby extend the range of v_{BM} amplitudes that can be encoded by the IHC, as postulated in experimental and theoretical studies (Kros and Crawford, 1990; Lopez-Poveda and Eustaquio-Martín, 2006). To illustrate this, Fig. 3A shows how the simulated half-wave rectified

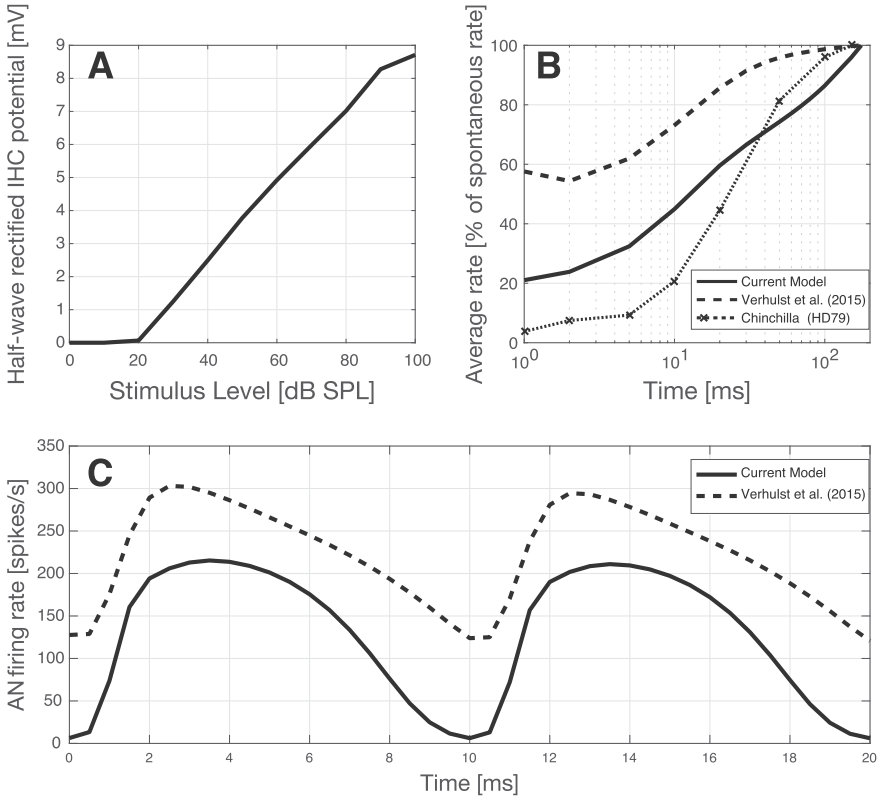

Fig. 3. IHC transduction aspects of the human auditory periphery model. A. Half-wave rectified IHC receptor potential (V_{IHC}) in response to a 4-kHz pure-tone plotted as function of sound level. **B**. Normalized firing rate over time (bin size of 15 ms) upon the offset of a pure-tone stimulus (time 0) of 80 dB SPL (CF = 8 kHz). Simulations were compared to chinchilla data collected 30 dB above the fibers' threshold (CF = 2.75 kHz; Harris and Dallos, 1979). **C.** Comparison of AN firing rate time-courses in response to a 60-dB 4-kHz amplitude modulated pure tone (f_m = 100 Hz) between the present and earlier Verhulst et al. (2015) model. The inclusion of basolateral outward K^+ currents in the IHC transduction model yields sharper offset responses and AM firing rate profiles that return to zero in the stimulus envelope minima, thus yielding a modulation gain.

receptor potential (a physiologically relevant measure alternative to the V_{IHC} DC component) grows roughly linearly with SPL (in dB) for stimulus levels up to 90 dB SPL. As the corresponding ν_{BM} responses only showed compressive growth up to levels of 80 dB SPL (at 1 kHz; Fig. 2C), IHC transduction extended the compressive growth range by 10 dB. This feature of mechanical-to-electrical transduction corroborates the guinea-pig IHC transduction measured *in vivo* by Cheatham and Dallos (1999) who reported an unsaturated, almost linear growth of the half-wave rectified IHC receptor potential (in dB) for stimulation levels up to 90 dB (in the second, third, as well as fourth cochlear turn of the guinea pig).

A second characteristic ascribed to outward IHC K^+ channels relates to their delayed activation that creates small, but fast, hyperpolarization sags at the stimulus offset (Zeddies and Siegel, 2004). Consequently, the V_{IHC} decay time upon stimulus offset is not limited by the (slower) IHC membrane time constant and hence causes sharp offset responses at the level of the AN. Previous functional models of the auditory periphery (both AN and perception models) have not accounted for the dynamic properties of IHC K^+ channels and hence overestimate recorded AN offset responses if the (slower) IHC time constant is not compensated for by introducing additional dynamics at the AN synapse. The IHC model description adopted here is based on *in vitro* recordings (e.g., Jia et al., 2007; Johnson et al., 2011) whereas previous models of the auditory periphery modeled IHC membrane characteristics based on intracellular, *in vivo* recordings (e.g. Zhang et al., 2001; Sumner et al., 2002). These *in vivo* intracellular recordings are known to be compromised by the leakage created by the perforation of the basolateral membrane (Zeddies and Siegel, 2004; Kros and Crawford, 1990) and might therefore not represent the functioning of the IHC membrane under physiological conditions.

Thirdly, we have shown that basolateral voltage-dependent K^+ conductances contribute to the asymmetry between AN adaptation and recovery properties (Altoè et al., submitted). Other IHC-AN complex models that include a low-pass filter model for the IHC membrane can only account for these asymmetries by including a power-law adaptation process in the AN synapse model (Zilany et al., 2009, 2014). As shown in Fig. 3B, including the dynamics of outward K^+ channels yields more realistic AN recovery properties to pure-tone stimulation when comparing the performance to an AN model in which the IHC membrane is modeled as a low-pass filter (Verhulst et al., 2015). Note that unlike the chinchilla data (Harris and Dallos, 1979), simulated human adaptation does not reach zero upon stimulation offset (i.e., t = 0) because human cochlear Q_{ERB}'s are much sharper than chinchilla Q_{ERB}'s. This difference in cochlear tuning yields longer ringing ν_{BM} responses for humans after stimulation offset (Raufer and Verhulst, 2016), and causes the observed offset in AN recovery functions between species at t = 0.

A last important difference between the functionality of traditional low-pass filter IHC models and the present circuit-model is related to the relative weight of the DC and AC components of the IHC receptor potential. The activation of the basolateral K^+ channels in the adopted nonlinear IHC circuit-model compresses the DC component of the IHC response more so than its AC components (Altoè et al., submitted). Consequently, the transmitted DC

component is much smaller in the presented IHC model than in traditional low-pass filter models. This has important consequences for the representation of amplitude-modulated (AM) stimuli. Fig. 3C shows how simulated AN firing rates to a 4-kHz AM pure-tone yield realistic AN responses that return to SR in the dips of the stimulus envelope (see Fig. 1 in Joris and Yin, 1992 for an experimental reference). Simulated AN responses for models that include a low-pass filter model (e.g., Zhang et al., 2001; Verhulst et al., 2015) show an unrealistically high DC component in the high-frequency AM firing patterns (Fig. 3C). The Zilany et al. (2009, 2014) models counteracted the large DC component resulting from the IHC low-pass filtering by a including a functional power-law adaptation model after the AN synapse (Zilany et al., 2009) to match the reference AN data. This example thus illustrates that there are multiple ways in which realistic AN responses to stimulus envelope modulations can be simulated. In fact, based on the simulations in Fig. 3 and the additional evidence Altoè et al. (submitted), we note that multiple AN properties requiring power-law adaptation in the model by Zilany et al., (2009) (e.g., modulation gain, recovery properties), can be explained by the voltage-dependent activation of the basolateral K^+ channels in the present model.

2.4. Auditory-nerve synapse

The auditory-nerve synapse model consists of two stages: (i) a description of the driven neurotransmitter exocytosis rate and (ii) a three-store diffusion model that simulates the kinetics of neurotransmitter at the synapse. The IHC receptor potential controls the exocytosis rate (k) through a non-instantaneous nonlinearity which activates with a second-order kinetics (time constant τ_{Ca} of 0.2 ms):

$$n_\infty = \left(1 + e^{-\frac{V_{\mathrm{IHC}} - V_{0.5}}{s}}\right)^{-1/2}, \tag{10}$$

$$n + \tau_{\mathrm{Ca}} \frac{dn}{dt} = n_\infty, \tag{11}$$

$$k(t) = k_{max} n^2, \tag{12}$$

where k_{max} is the maximal neurotransmitter release rate. The equations simulate the activation of the synaptic Ca^{2+} currents that drive exocytosis (Beutner et al., 2001) and qualitatively match the experimental relationship between AN firing rate and V_{IHC} *in vivo* (Zagaeski et al., 1994). The voltage sensitivity of the exocytosis rate (s in Eq. (10)) is a fitting parameter (1.5 mV) and the value of the time constant τ_{Ca} in Eq. (11) was chosen based on macroscopic IHC Ca^{2+} current recordings in the mature gerbil (Johnson and Marcotti, 2008). The half-activation potential ($V_{0.5}$) and peak exocytosis discharge rate (k_{max}) parameters were varied among SR synapse types to match measured rate-level curves for different SR fibers qualitatively. In particular, the half-activation potential ($V_{0.5}$) results from matching the desired SR given k_{max}. $V_{0.5}$ was determined from k_{max} and spontaneous exocytosis rate (k_{SR}) using the following description:

$$V_{0.5,\mathrm{SR}} = \log\left(\frac{k_{max} - k_{\mathrm{SR}}}{k_{\mathrm{SR}}}\right) s + V_{\mathrm{rest}}. \tag{13}$$

k_{max} was selected to match experimental AN peak-to-steady-state ratios (Rhode and Smith, 1985; Taberner and Liberman, 2005) and resulted in values of 800 (LSR), 1000 (MSR) and 3000 (HSR) spikes/s. k_{max} was kept CF-independent based on the Taberner and Liberman (2005) observations for mice. Even though cat data show a CF-dependence of k_{max} (Liberman, 1978), we refrained from including additional fitting parameters. The variability of the $V_{0.5}$ and k_{max} parameters across the simulated SR fibers mimics the heterogeneity in Ca^{2+} signaling at the synapse reported for a range of SR fibers (Frank et al., 2009; Ohn et al., 2016).

Once the driven exocytosis rate k is computed, the model implements a classic three-store diffusion model of AN adaptation (Westerman and Smith, 1988), that is described using a notation similar to that of Meddis (1986):

$$\begin{aligned} y(t) &= q(t)k(t) \\ \frac{dq}{dt} &= r(t) - y(t) \\ r(t) &= \alpha_{\mathrm{q}}\left(\frac{l(t)}{L} - \frac{q(t)}{M}\right), \\ \frac{dl}{dt} &= \alpha_{\mathrm{l}}\left(1 - \frac{l(t)}{L}\right) \end{aligned} \tag{14}$$

where the release rate $y(t)$ is the effective exocytosis rate. The probability that one vesicle is released is thus $P_{\mathrm{Rel}}(t) = y(t)\Delta t$, with Δt the simulation time step. $q(t)$ and $l(t)$ model the vesicles populating the ready-releasable pool (RRP) and the reserve pool (RP) of neurotransmitter, which can contain a maximum of M and L vesicles, respectively. Finally, α_{q} and α_{l} represent the maximum possible replenishment rates for the RRP and RP.

The difference between our model description and that of Westerman and Smith (1988), is that we assigned a physical meaning to the phenomenological parameters of the original model, and thereby constrain two of the four fitting parameters to recently published physical properties of the IHC. In particular, the variables M (i.e., the vesicle size of the RRP), and α_{q} (its maximal replenishment rate) are constrained to experimental data (Meyer et al., 2009; Pangršič et al., 2010). The replenishment rate of the RP (α_{l}) controls the steady-state AN firing rate and was set to 300 vescicles/s to match experimentally observed steady-state AN firing rates of 220 spikes/s (Zagaeski et al., 1994). The size of the reserve pool L is a fitting parameter and was set to match the experimental 60-ms short-term adaptation time constant reported by Westerman and Smith (1984).

The model computes the probability of releasing a vesicle at each simulation time step, and this probability drives the AN firing probability (P_{Firing}). Refractoriness was included and based on the work of Peterson et al. (2014):

$$P_{\mathrm{Firing}}(t) = P_{\mathrm{Rel}}(t) - \int_{t-t_a}^{t} P_{\mathrm{Firing}}(\tau)d\tau - \int_{-\infty}^{t-t_a} P_{\mathrm{Firing}}(\tau)e^{-\tau/t_a}\, d\tau, \tag{15}$$

where t_a (0.6 ms) controls both the absolute and relative refractoriness time constant. Even though Miller et al. (2001) showed that AN refractory periods can vary among AN fibers, we opted to keep this parameter constant across different AN fiber types. Finally, the firing rate for a single AN nerve fiber is given by:

$$r_{\mathrm{Firing}}(t) = P_{\mathrm{Firing}}(t)/\Delta t, \tag{16}$$

where Δt denotes the bin width of the model which was $1/f_s$ for the adopted f_s of 20 kHz. Note that the developed IHC-AN model is purely deterministic as it aims to predict responses of large populations of auditory neurons efficiently. This approach is different from other functional and biophysical IHC-AN model descriptions

Table 1
Model parameters.

Cochlear Mechanics			
Number of cycles of the traveling waves that is traversed before its maximum is reached. Fitting parameter in the Zweig (1991) model for the cochlear impedance (see Eqs. (1) and (2)).	N_W	1.5 [-]	Fitting parameter
Pole of the BM admittance in Eqs. (1) and (2) to yield realistic human Q_{ERB} values for low stimulus levels. The CF-dependence of this function is given by the normal-hearing data in Fig. 2A	$\alpha*_A$	0.062 at 1 kHz [-]	Fitting parameter
Pole of the BM admittance in Eqs. (1) and (2) to corresponding to minimal Q_{ERB} values corresponding to high stimulation levels. This value is independent of CF.	$\alpha*_P$	0.305 [-]	Fitting parameter
Compression slope of the instantaneous nonlinearity function describing the relation between α^* and v_{BM} in Eq. (5).	C	0.31 dB/dB	Fitting parameter
Kneepoint of the instantaneous nonlinearity function describing the relation between α^* and v_{BM} in Eq. (5). v_{BM} grows linearly below the kneepoint, and grows compressively above.	$v_{BM,30dB}$	$1.6e^{-7}$ m/s	Fitting parameter
IHC transduction parameters			
Ratio that brings the IHC cilia displacement within a realistic range as input to the biophysical IHC transduction model.	$d_{IHCcilia}/v_{BM}$	0.118 s	Fitting parameter
IHC membrane capacitance	EP	90 mV	
MET channel	C_m	12.5 pF	Johnson et al. (2011) and
Max. conductance	G_{max}	30 nS	Jia et al. (2007)
Displacement offset	x_0	20 nm	Jia et al. (2007)*
Displacement sensitivity	s_0	16 nm	Jia et al. (2007)*
Displacement sensitivity	s_1	48 nm	Jia et al. (2007)*
Activation time constant	τ_{MET}	50 µs	Kennedy et al. (2003)
K^+channels			
Max. conductance	$G_{K,f/s}$	230 nS	Johnson et al. (2011)
Half-activation potential	$V_{0.5}$	−31 mV	Johnson et al. (2011)
Voltage sensitivity	s_k	10.5 mV	Johnson et al. (2011)
Reversal potential, fast	$E_{K,f}$	71 mV	Kros and Crawford (1990)
Reversal potential, slow	$E_{K,s}$	78 mV	Kros and Crawford (1990)
Activation time constant, fast	$\tau_{K,f}$	0.3 ms	Kros and Crawford (1990)
Activation time constant, slow	$\tau_{K,s}$	8 ms	Kros and Crawford (1990)
AN synapse			
RP maximum replenishment rate	α_l	300 spikes/s	Fitting parameter
RRP maximum replenishment rate	α_q	700 spikes/s	Pangršič et al. (2010)
RP size	L	60 vesicles	Fitting parameter
RRP size	M	14 vesicles	Meyer et al., 2009; Pangršič et al. (2010)
Sensitivity of the Boltzmann function relating V_{IHC} and driven exocytosis rate	s	1.5 mV	Fitting parameter
Peak exocytosis rate	k_{max}	800 (LSR), 1000 (MSR), 3000 (HSR) spikes/s	Fitting parameter
Spontaneous exocytosis rate	k_{SR}	1 (LSR), 10 (MSR), 70 (HSR) spikes/s	Fitting parameter
Time constant of the absolute and relative refractoriness	t_a	0.6 ms	Fitting parameter
Brainstem model			
Scalar to achieve realistic CN firing rates	A_{CN}	1.5 [-]	Nelson and Carney (2004)
Relative CN inhibition/excitation strength	S_{CN}	0.6 [-]	Nelson and Carney (2004)
CN inhibition delay	D_{CN}	1 ms	Oertel (1983)
Excitation time constant	τ_{exc}	0.5 ms	Oertel (1983)
Inhibition time constant	τ_{inh}	2 ms	Oertel (1983)
Scalar to achieve realistic IC firing rates	A_{IC}	1 [-]	Nelson and Carney (2004)
Relative IC inhibition/excitation strength	S_{IC}	1.5 [-]	Nelson and Carney (2004)
IC inhibition delay	D_{IC}	2 ms	Nelson and Carney (2004)
Population response parameters			
Scalar for realistic human ABR W-I peak amplitude	A_{W-I}	$6.28e^{-14}$ [-]	Scaling parameter
Scalar for human ABR W-III peak amplitude	A_{W-III}	$7.22e^{-14}$ [-]	Scaling parameter
Scalar for human ABR W-V peak amplitude	A_{W-V}	$3.52e^{-20}$ [-]	Scaling parameter

See also Table I of Verhulst et al. (2012, 2015) for additional parameters relating to the middle-ear and cochlear transformation. * see additional explanation in Appendix A.

that include the stochastics of AN spiking and focus more on simulating single-unit AN spike-time histogram properties (Sumner et al., 2002; Zilany et al., 2009; Peterson et al., 2014; Moezzi et al., 2016).

2.5. Population responses

Even though simulated single-unit properties of the human AN model can be compared to physiological recordings in animals, there is no guarantee that the human periphery works similarly to that of the cat. To address this, population responses were simulated to offer a direct comparison between model predictions and human physiological responses. As the schematic in Fig. 1 depicts, population responses can be written out on the level of the auditory nerve, cochlear nucleus and inferior colliculus. The population AN response sums 13 HSR (70 spikes/s), 3 MSR (10 spikes/s) and 3 LSR

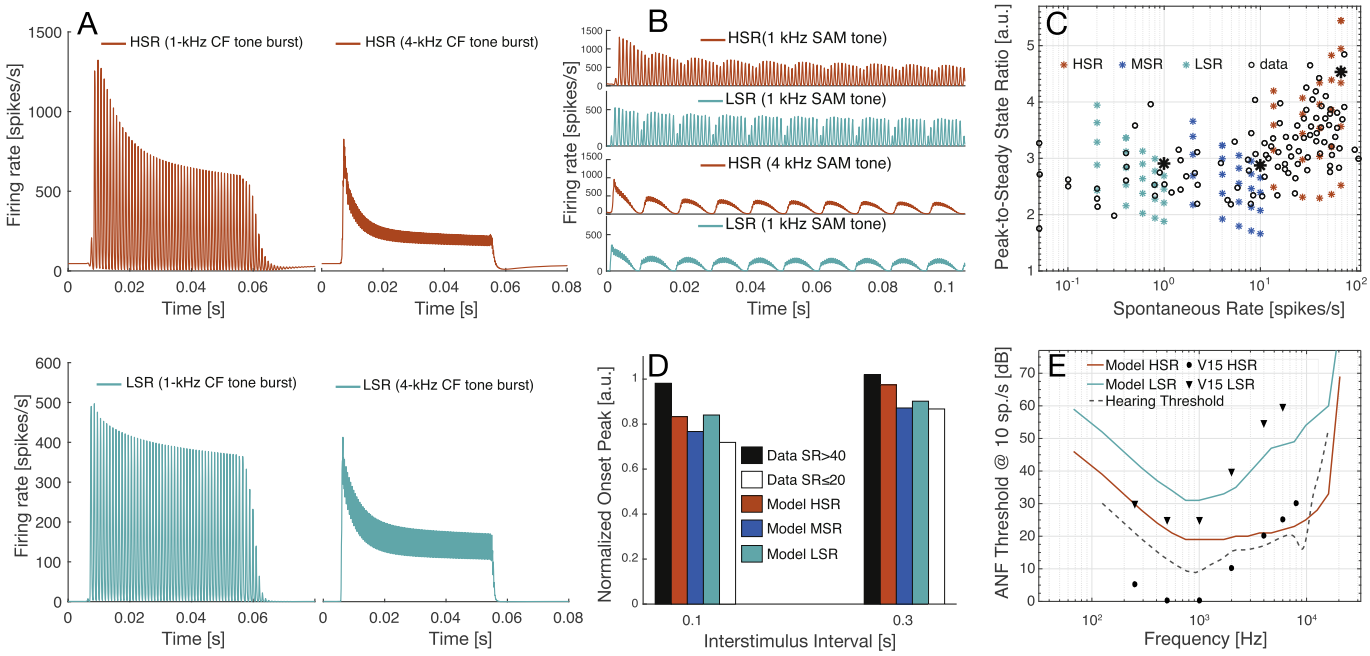

Fig. 4. Simulated AN response properties. Responses of HSR (70 spikes/s), MSR (10 spikes/s) and LSR (1 spike/s) fibers are marked in red, blue and cyan, respectively. **A**. Firing rate (r_{firing}) of HSR and LSR fibers in response to 50-ms long 1 and 4-kHz pure tones presented at CF. Stimuli were presented 30 and 50 dB above threshold for HSR and LSR fibers, respectively. **B**. HSR and LSR fiber responses to continuous 100% amplitude-modulated tones ($f_m = 100$ Hz) evaluated at CF for a 1 and 4 kHz carrier frequency. Stimulus intensities matched the pure tone conditions in A. **C.** Effect of spontaneous rate (k_{SR}) on simulated peak-to-steady-state firing rate (Taberner and Liberman, 2005) for different k_{max} values. The black * values correspond to the adopted k_{max} values and the vertically stacked *-traces for each SR represent a parameter sweep between 0.2 (bottom) and 1 k_{max} (top * at each CF). Stimulus intensities of 40, 70 and 90 dB SPL were chosen to present supra-threshold stimuli (>30 dB above threshold) for the distribution of HSR, MSR, LSR fiber types, respectively. The CF was 4 kHz and the reference data were mice recordings (Taberner and Liberman, 2005). **D**. Onset-peak rate of HSR and LSR fibers to the second pure tone presented in a stimulus pair with a 0.1 or 0.3 s interval, normalized by the onset-peak response of the second tone in a 1.9 s inter-stimulus interval pair. The reference data and analysis method stemmed from a chinchilla AN study (Relkin and Doucet, 1991). **E**. AN threshold curves for LSR and HSR fibers across CF. Thresholds were defined as the minimum intensity required to increase the average firing rate by 10 spikes/s above SR; a criterion similar to that of Liberman (1978). Simulated fiber thresholds were compared to the human threshold of hearing for insert earphones (Han and Poulsen, 1998) and the simulated AN thresholds in Verhulst et al. (2015). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(1 spikes/s) fiber responses (i.e., r_{AN}) across each CF and for CFs between 112 Hz and 12 kHz. This population response corresponds to the compound action potential (CAP) and is translated to a human ABR wave-I (W-I) by multiplying the CAP response with a scaling factor $A_{W\text{-}I}$ ($6.28e^{-14}$) to yield W-I amplitudes that match human observations (Table VIII- 1 in Picton, 2011). We thus represent the electrical dipole at the heart of auditory evoked potentials in a simplified form that assumes that changes in neural population responses correlate with changes in the electrical dipole that generates the far-field potential recorded from the human scalp. This simplification follows the work of Melcher and Kiang (1996) who suggested that the ABR response can be described as the convolution of the instantaneous firing rate of a group of neurons with their unitary response. The unitary response describes the relationship between the individual neurons' discharge rate and the corresponding potential recorded between two scalp electrodes. The unitary response, or dipole-generation mechanism, is assumed to be stimulus and neuron type independent which justifies our approach of relating *changes* in summed neuronal populations to *changes* in stimulus characteristics or hearing loss profiles. All existing human ABR/EFR models that we are aware of follow the approach outlined by Melcher and Kiang (1996) by either summing firing rates of populations of neurons that were convolved with a unitary response (Dau, 2003; Rønne et al., 2012) or by summing spike waveforms convolved with action potentials (Schaette and McAlpine, 2011) to simulate individual ABR waves.

Even though ABR wave-I characteristics most closely represent the summed activity of the auditory-nerve population (and its deficits), it is hard to measure this response reliably in humans. To ease comparison between simulated and recorded human population response data the ABR W-III and W-V source generators were included in the model . Additionally, the dominant EFR generators have been identified to originate from the upper brainstem (IC; Bidelman, 2015), which further justifies the simulation of CN and IC source generators when comparing simulated EFRs with recorded responses. We assume that the functionality of bushy cells in the CN and IC are the dominant source generators of the ABR W-III and W-V, respectively (Melcher and Kiang, 1996) and include the same sources for the EFR simulations.[2] Even though there are other neuron types that may differentially contribute to the ABR or EFR (e.g. onset-sensitive octopus cells), we employ a simplified and functional brainstem model that captures both the onset and modulation sensitivity of CN and IC bushy cells using a same-

[2] The finding that a convolution of recorded ABRs with a pulse train did not predict the frequency-following waveform characteristics to click train stimuli in Bidelman (2015), does not preclude the use of the same CN and IC neuron model for both evoked potential types in the present study. The presented model of the auditory periphery (before brainstem processing) already predicts a breakdown of this linear convolution concept because (i) AN responses to sustained stimuli and pulses are not linearly related through the model equations, and (ii), not all AN fiber types contribute equally to simulated AN population responses in response to onset (more HSR driven) or sustained stimuli (more relative contributions of LSR fibers).

frequency inhibition-excitation model (Nelson and Carney, 2004). We reformulate the original model equations as follows[3]:

$$r_{CN}(t) = A_{CN}\left[\frac{t}{\tau_{exc}^2}e^{-t/\tau_{exc}} * r_{AN}(t) - S_{CN,inh}\frac{t}{\tau_{inh}^2}e^{-t/\tau_{inh}} * r_{AN}(t - D_{CN})\right], \quad (17)$$

where r_{AN} represents the sum of the different included AN fiber responses innervating at each CF. r_{IC} is obtained by replacing all CN subscripts with IC and all AN subscripts with CN.

The population response W-III was obtained by summing all r_{CN} responses from CF channels between 112 Hz and 12 kHz, and by applying scaling factor of $A_{W\text{-}III}$ ($7.22e^{-14}$) to match the experimental human ABR W-III peak amplitude (Table VIII- 1 in Picton, 2011). Similarly, the r_{IC} amplitude was scaled ($A_{W\text{-}V} = 3.52e^{-20}$) to match the human ABR W-V peak amplitude. These scaling factors do not change the behavior of the model, but ensure that simulated W-I, III, and V peak amplitudes fall within the experimental range. Without these scaling factors, the population responses would have arbitrary units due to a summation of large amount of firing rates across the simulated CFs and fiber types. The characteristics of each simulated ABR wave were evaluated individually as they occur separately in time due to the synaptic delays at ascending processing stages (Table 1 and Eq. (17)). On the other hand, simulated EFRs sum up the scaled population responses at the level of the AN, CN and IC (i.e., W-I, III and V) to capture the contribution of sources from different ascending processing stages that were shown to contribute to the recorded EFR (Bharadwaj and Shinn-Cunningham, 2014).

2.6. Simulating synaptopathy and IHC deficits

The model can simulate frequency-specific degrees of cochlear synaptopathy by varying the number and types of AN fibers that synapse onto each IHC. The summed $r_{AN}(t)$ response at each CF is given by:

$$r_{AN}(t) = N_{LSR}r_{Firing,\ LSR}(t) + N_{MSR}r_{Firing,\ MSR}(t) + N_{HSR}r_{Firing,\ HSR}(t), \quad (18)$$

where $N_{LSR}(3)$, $N_{MSR}(3)$ and N_{HSR} (13) represent the number of LSR, MSR and HSR fiber synapses with rates of 1, 10 and 70 spikes/s, respectively at each CF for the normal-hearing model. These numbers were roughly based on distributions of AN fiber types reported in cat (Liberman, 1978) as well as the total number of IHC synapses reported in several synaptopathy studies in guinea pig, mice and monkey (Kujawa and Liberman, 2009; Furman et al., 2013; Valero et al., 2017). Selective cochlear synaptopathy (Furman et al., 2013) is simulated by removing the contribution of LSR and MSR fibers, whereas more severe degrees of synaptopathy can be simulated by lowering N_{HSR} as well. For the simulations presented here, the number of AN fibers synapsing onto each IHC was kept constant across CF. Experimental synapse counts from rhesus monkeys suggest that fewer numbers of AN fibers synapse to each IHC near the cochlear base and apex (Fig. 4 in Valero et al., 2017), and it is also known that noise exposure may selectively remove AN fiber synapses at CFs near and above the noise exposure band (Kujawa and Liberman, 2009; Valero et al., 2017). As the N parameters can be inserted as a CF-dependent vector, it is possible to simulate specific CF-dependent synaptopathy profiles (see also Verhulst et al., 2016) eventhough this was not further pursued in the present study.

IHC lesions can be simulated by setting all N values in Eq. (18) to zero for a particular CF. Other IHC deficits, such as those associated with metabolic presbycusis which can reduce the endocochlear potential (EP) down to 40 mV (Schmiedt, 2010) were not specifically considered in this work. However, because the IHC-AN model is biophysical, the EP can be reduced in Eq. (A4) resulting in a reduced MET current and V_{IHC} . As presbycusis may also affect the reversal potential of the K^+ currents ($V_{K,F/S}$) in Eq. (A7), the corresponding change in $V_{K,F/S}$ should be calculated from ion concentration differences between the intra and extracellular environment (Zeddies and Siegel, 2004).

2.7. Fitting parameters

The total number of fitting parameters is 12 (see Table 1) and most of these parameters cannot be quantified directly using physiological measurements. The fitting parameters are not independent and keep the multitude of model responses within a realistic operating range. They were determined to match physiological responses at the level of processing they referred to, and were not specifically calibrated to improve ABR or EFR simulations. The population responses served as an evaluation tool for the functioning of the human auditory periphery model and were thus never the target for parameter optimization. A good match between experimental and model data at the level of the population response would thus only reflect a realistic representation of cochlear mechanics and the IHC-AN complex for humans. Consequently, the A_W parameters were not considered as fitting parameters as they just scale the population responses to match human ABR peak amplitudes.

The number of fitting parameters related to cochlear and IHC processing amounts to six. By calibrating BM frequency tuning using the double pole α^* of the BM admittance, the number of parameters related to BM frequency tuning and nonlinearity are minimal (α^*_A, α^*_P, C and $v_{BM,30dB}$). The Zweig (1991) equations for the cochlear impedance (Eqs. (1) and (2)) have one fitting parameter ($N_W = 1.5$) that describes the number of cycles the traveling wave traverses before it reaches its maximum. The cochlear model parameters were kept CF-invariant such that the observed CF-dependence of the kneepoint of compressive v_{BM} growth (see Fig. 2C) solely results from the BM frequency tuning and the shape of the middle-ear filter that attenuates the input to the cochlear model for high stimulus frequencies. We adopt a 1st order bandpass filter with cut-off frequencies of 600 and 4000 Hz, using a forward pressure gain of 18 dB (Puria, 2003). The sixth fitting parameter ($d_{IHCcilia}/v_{BM}$) connects v_{BM} [in m/s] to deflection of the IHC cilia [in m]. A factor of 0.118 is multiplied with v_{BM} to bring IHC cilia deflection into realistic biophysical operating ranges for the IHC transduction stage. As the parameters in the IHC transduction model were based on physical values, this model stage does not include fitting parameters.

The included biophysical IHC-AN model has an additional six fitting parameters (s, L, α_l, t_a, k_{max}, k_{SR}). Even though the spontaneous exocytosis rate (k_{SR}) is technically a fitting parameter, its choice only matters in determining the correct k_{max} from recorded

[3] Both the original model description in Nelson and Carney (2004), and the description of the Nelson and Carney (2004) model in Verhulst et al. (2015) have an error in the description of the CN/IC model equation that was present in the paper but not in the model code (Eq. (2) of Nelson and Carney (2004) and Eq. (12) and (13) of Verhulst et al., 2015). We clarified this error in a personal communication with Prof. Carney, and agree that Eq. (17) describes the intended CN/IC model which functions as a bandpass modulation filter that is most sensitive to modulations in the 100 Hz range.

peak-to-steady state ratios. Our model description allowed for a parameter reduction in comparison to earlier AN synapse models by deriving the size (M) and replenishment rate (α_q) of the ready-releasable pool from measured values. Additionally, by describing the model variables in units of neurotransmitter vesicle numbers, the maximum replenishment rate of the reserve pool (α_l) could be derived from AN recordings. In comparison, the original AN synapse model by Westerman and Smith (1988) has four dependent fitting parameters aside from a fifth parameter that describes the level-dependence of k at its input. As k depends on the details of the IHC transduction process (that were not described by Westerman and Smith, 1988) and their synapse model does not include refractoriness, additional fitting parameters would be required to yield a realistic description of the entire IHC-AN complex using the Westerman and Smith (1988) synapse model (e.g., Zilany et al., 2009, 2014).

The entire model of the human cochlear periphery has 12 fitting parameters of which only one is frequency-dependent, namely α^*_A. This number of fitting parameters is minimal and smaller than in other AN-models that often introduce complex BM processing descriptions (e.g., Zhang et al., 2001; Sumner et al., 2002) and CF-dependencies in several (including the AN) model stages (e.g., Zilany et al., 2009, 2014) to account for the variety of AN reference data properties. The cochlear transmission-line model as a descriptor for human cochlear mechanics is a big advantage for this process, as well as the efforts in rendering several functional equations of the IHC-AN complex biophysical without adding complexity. The following sections demonstrate that this mixed functional/biophysical model can unify both biophysical modeling efforts of the IHC-AN complex (Sumner et al., 2002; Zeddies and Siegel, 2004; Lopez-Poveda and Eustaquio-Martín, 2006) and more functional descriptions of the human auditory periphery (e.g., Rønne et al., 2012; Zilany et al., 2014; Verhulst et al., 2015) to yield a model of the human auditory periphery that accounts for normal and hearing-impaired human physiological response features (i.e., otoacoustic emissions, ABRs and EFRs).

3. Model performance

First, we evaluate key properties of simulated AN responses to amplitude-modulated and pure tones for which single-unit reference AN recordings are available. Afterwards we evaluate the model performance in capturing signatures of normal and hearing-impaired human ABRs and EFRs.

3.1. Single-unit AN response properties

3.1.1. Responses to amplitude-modulated and pure tones

Fig. 4A shows simulated firing rates (r_{Firing}) for low- and high SR fibers to 1 and 4-kHz tone-bursts presented 30 dB above the fibers' threshold. Phase-locking to the stimulus fine structure was present for the 1-kHz fiber response and absent for the 4-kHz fiber. The 4-kHz fiber predominantly fired to the stimulus envelope after its onset response; generally consistent with physiological studies that describe phase-locking properties of the AN (see Joris et al., 2004 for a review). Phase-locking differences between the 1 and 4-kHz CF fiber responses are also evident from the amplitude-modulation simulations in Fig. 4B.

3.1.2. AN response adaptation

Simulated pure-tone responses showed an onset peak that decayed to a steady-state firing rate over time (Fig. 4A). The ratio between the onset peak and the steady-state firing rate can be used as measure of response adaptation (Taberner and Liberman, 2005). Simulated and experimental peak-to-steady-state firing rates are

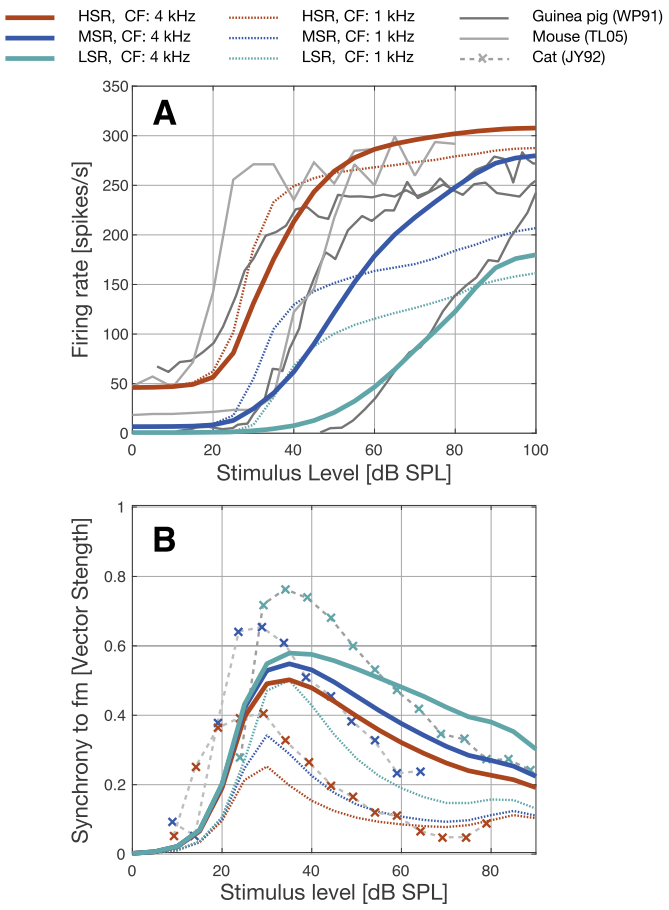

Fig. 5. Level-dependent properties of AN fibers A. AN rate-level curves were computed for LSR, MSR and HSR fibers with CFs of 1 (dashed colored) and 4 kHz (solid colored) using 50-ms tone-bursts that were presented on-CF and that had a 0.5-ms ramp in agreement with experimental procedures (Winter and Palmer, 1991; Taberner and Liberman, 2005). The reference data stemmed from mouse (Fig. 6 in Taberner and Liberman, 2005; 18.8 kHz CF for SR of 47.6 spikes/s and CF of 23.7 kHz for SR of 0.1 spikes/s) and guinea pig recordings (Fig. 1 of Winter and Palmer, 1991; CF ~ 1.5 kHz for 65 spikes/s, 10 spikes/s and 0 spikes/sec fibers). **B**. Synchrony-level functions were calculated for 600-ms long pure tone stimuli that were 100% modulated using a 100-Hz pure tone. Vector strength (see text) to f_m was plotted against the stimulus intensity for three fiber types, and CFs of 1 and 4 kHz. The reference data came from cat AN recordings (Figs. 5 and 8 in Joris and Yin, 1992; 8.1 kHz CF for 2.6 spikes/s, 1.14 kHz CF for 6.3 spikes/s and 2.83 kHz for 73 spikes/s fibers, respectively). In agreement with the reference data, the present model was able to simulate modulation gain (i.e., vector strength > 0.5) for MSR and LSR fibers.

depicted in Fig. 4C as a function of SR, where the black *-symbols correspond to the ratios chosen to represent the LSR, MSR and HSR fibers in the model. In general, the ratio was higher for HSR fibers, and by modifying the k_{max} parameter (Eq. (13)), the model replicated measured peak-to-steady-state ratios within the entire physiological range for low CF fibers (i.e., CFs < 20 kHz). The vertically stacked *-symbols in Fig. 4C show stimulated ratios to a parameter sweep between $1 \cdot k_{max}$ (top) and $0.2 \cdot k_{max}$ (bottom) for each considered SR. For HSR fibers, the peak-to-steady-state ratio was mostly determined by an onset peak increment as k_{max} increased, while peak-to-steady-state ratio increments resulted from reduced steady-state responses for LSR fibers.

To further evaluate single-unit adaptation properties, AN recovery to prior stimulation was simulated. 100-ms, 2-kHz tones were presented 40 dB above the fibers' threshold and inter-stimulus intervals were 0.1, 0.3, and 1.9 s. In accordance with the experimental procedures by Relkin and Doucet (1991), onset-peak

firing rates to the second tone were normalized by the onset-peak rate to the second tone in the longest (1.9s) inter-stimulus interval pair. Fig. 4D depicts normalized onset-peak rates for LSR, MSR and HSR fibers as a function of inter-stimulus interval and shows that it took longer for LSR and MSR fibers to recover from prior stimulation than for the HSR fiber in the 0.3 s interval. This trend is consistent with the experimental observations and reflects how the recovery properties of AN fibers are affected by the AN peak-to-steady state ratio to which the k_{max} parameter was fit in the model. For the 0.1 s interval, the model captured recovery differences between MSR and HSR fibers well, while it predicted that LSR and HSR fibers types exhibit similar recovery properties.

It is important to note that the adaptation time constant (i.e., through L) was kept constant across the simulated AN fiber types, in agreement with the time constant observations reported for gerbil (Westerman and Smith, 1984). The SR-differences in onset-peak recovery in Fig. 4D thus have another origin in the model. The longer recovery for MSR and LSR fibers over HSR fibers can be explained by the combination of (i) the smaller onset-peak exocytosis rate for LSR *vs* HSR fibers, and (ii) the AN refractoriness which produces a nonlinear relationship between the neurotransmitter release rate (y) and discharge rate (r_{Firing}). In fact, the effect of an absolute refractory period (τ) on the relationship between y and r_{Firing} is approximately (Vannucci and Teich, 1978)[4]:

$$r_{\text{Firing}}(t) = \frac{y(t)}{1+\tau y(t)}. \tag{19}$$

It can be seen from Eq. (19) that when y decreases from y_1 to $y_2 = ay_1$, for $0 < a < 1$, r decreases from r_1 to $r_2 = br_1$, with $b = a(1+y_1)/(1+ay_1)$. Larger y_1 values thus result in smaller changes in r. We therefore conclude that the data by Relkin and Doucet (1991) cannot directly be interpreted as reflecting different adaptation properties between LSR and HSR synapses. Rather, their results can be explained by the (well-documented) stronger onset responses for HSR than LSR fibers (even when their adaptation time constants are identical).

3.1.3. AN firing thresholds

Even though there were no CF-dependencies introduced in the IHC transduction and AN fiber descriptions, the simulated threshold of hearing (here defined as the minimum intensity that increased the average firing rate by 10 spikes/s above SR; Liberman, 1978) showed a marked frequency dependence (Fig. 4E). The $d_{\text{IHCcilia}}/v_{\text{BM}}$ parameter describes the relation between v_{BM} and the inner-hair-cell cilia deflection at 1 kHz, and can be used to increase the HSR (and consequently the LSR) firing threshold. This offset does not affect the frequency dependence of the threshold and was set to yield a 20 dB SPL firing threshold at 1 kHz. The simulated threshold is somewhat higher than the human threshold of hearing for insert earphones (Han and Poulsen, 1998, Fig. 4E) because the adopted $d_{\text{IHCcilia}}/v_{\text{BM}}$ formed a compromise between simultaneously capturing near-threshold and supra-threshold AN data (i.e., the shape of the AN rate level curves). A recent study by Huet et al. (2016) shows a good agreement between gerbil AN thresholds (10 spikes/s increment criterion) and corresponding behavioral thresholds across the CF range, supporting that this direct comparison between the AN firing threshold and behavioural threshold is meaningful.

The simulated AN threshold shape resulted from the combined effect of the middle-ear filter characteristics and the implemented CF-dependence of Q_{ERB} that yielded overall higher cochlear gain and v_{BM} amplitudes for high-CFs when the stimulus level was kept fixed (see Fig. 2A). The middle-ear filter order (1^{st}) and cut-off frequencies (0.6–4 kHz) were chosen to comply with human middle-ear efficiency characteristics (1^{st} order, passband: 0.25–5 kHz; Lewis and Neely, 2015) and human temporal bone middle-ear impedance measurements (~10 dB filter bandwidth: 0.4–3.5 kHz; Puria, 2003). Even though middle-ear filter characteristics have a strong influence on shaping the AN firing threshold, the threshold curve in Fig. 4E was also influenced by the CF-dependence of cochlear frequency tuning. In particular, the threshold showed a shallower roll-off towards high CFs than predicted by the 6 dB/octave middle-ear filter. Additionally, the present model showed a more realistic low-frequency AN firing threshold roll-off than the earlier transmission-line cochlear model that implemented a 2^{nd} order HP and 1^{st} order LP middle-ear filter (black dots in Fig. 4E; Verhulst et al., 2015). This difference was attributed to cochlear input impedance differences between the two models (i.e., resistive Z_0 in the present model and reactive Z_0 in Verhulst et al., 2015). Lastly, simulated LSR firing thresholds were approximately 12 dB higher than HSR thresholds for CFs below 1 kHz, and increased up to 30 dB for higher CFs. These values corroborate the observations in cat, where LSR fibers had overall 15–40 dB higher AN firing thresholds (Fig. 11 in Liberman, 1978). The larger separation between LSR and HSR firing thresholds towards high CFs was caused by the increasingly greater amounts of BM compression seen by the LSR fiber as the cochlear base was approached, while HSR fibers always operated in the linear regime and showed more homogeneous thresholds across CF.

3.1.4. Effect of stimulus level

Rate-level (Fig. 5A) curves were computed to evaluate AN fiber responses to stimulus level changes of 1 or 4 kHz tone-bursts (50-ms duration, 0.5-ms ramp) and were derived from the average discharge rate of an on-CF fiber (Winter and Palmer, 1991; Taberner and Liberman, 2005). Data from experimental studies are plotted alongside our simulations in Fig. 5A and reflect a variety of experimental AN rate-level curves from different species and CFs. Despite the differences in the reference data, the 4-kHz simulations captured the qualitative differences between LSR, MSR and HSR guinea pig fibers well (CF~1.5 kHz; Winter and Palmer, 1991). The mouse rate-level curves show somewhat steeper growth than our simulations predicted (CF: 18.8 kHz (HSR) and 23.7 kHz (LSR); Taberner and Liberman, 2005), but we refrained from overfitting the IHC-AN parameters to specific experiments given that the cochlear mechanics are also fundamentally different across species. Fig. 5A shows that the shape of the simulated rate-level curves was different for the different CF fibers (colored 1-kHz dotted lines *vs* 4-kHz solid lines) despite the CF-independent IHC-AN parameters. Differences in BM processing across CF are thus reflected in the shape of AN rate-level curves. For example, as the AN firing threshold was lower for 1-kHz than for 4-kHz fibers (Fig. 4E), the steeper part of the 1-kHz rate-level curves occurred at lower stimulus levels. Secondly, the smaller dynamic range of levels encoded by the BM for the 1-kHz than 4-kHz CF (see I/O function differences in Fig. 2) yielded AN level-curve compression at lower stimulus levels for the 1-kHz CF.

Synchrony-level functions were simulated for 600-ms long 1 or 4 kHz pure tones that were 100% modulated with a modulation frequency (f_{m}) of 100 Hz (Joris and Yin, 1992). Synchrony to the stimulus envelope was quantified using vector strength (Goldberg and Brown, 1969) and was calculated by extracting the magnitude of the f_{m} component from the Fourier spectrum of the fibers'

[4] The solution of Vannucci and Teich (1978) which describes the rate of a Poisson process with a dead-time is only exact for steady driven rates over time windows that exceed the dead-time. Even though the validity of these equations for AN transient responses are thus only approximate, they do confirm our simulations using the biophysical IHC-AN complex model.

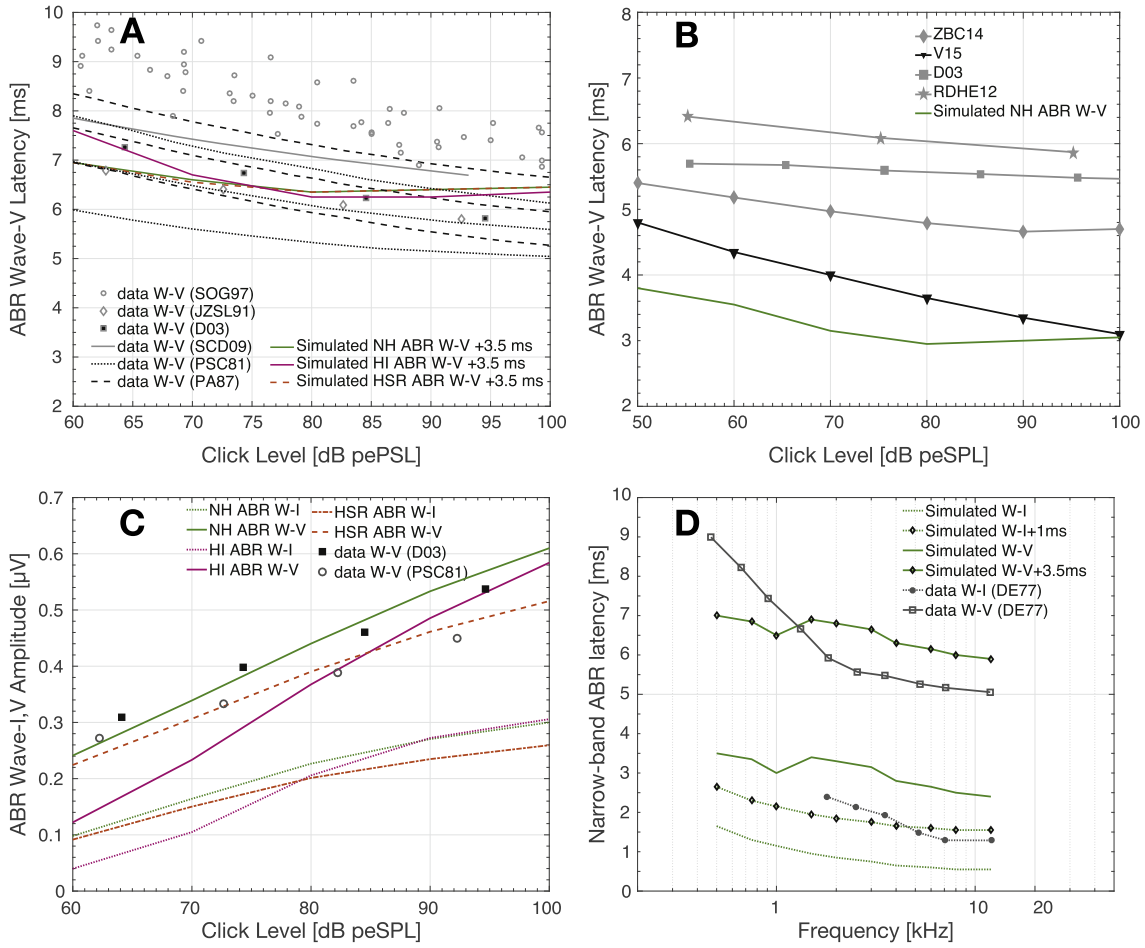

Fig. 6. Simulated ABR wave-I and wave-V characteristics. **A**. Simulated ABR W-V latencies as a function of the 80-μs click intensity for a normal-hearing (NH; green), a sloping high-frequency hearing loss (red; HI) and a synaptopathy model in which only HSR fibers contributed to the population response while LSR and MSR fibers were removed (HSR; dashed red). Simulated wave-V latencies were offset by 3.5 ms to ease comparison to the human reference data that were offset by 30 dB to translate reported dB HL values to our dB SPL scale (Strelcyk et al., 2009). Synaptopathy did not affect the ABR W-V latency curve, whereas cochlear gain loss steepened the latency curve. Human reference data came from Picton et al. (1981; PSC81 mean ±2 stds), Prosser and Arslan (1987; PA87 mean and confidence interval), Serpanos et al. (1997; SOG97), Jiang et al. (1991; JZSL91), Dau (2003; D03) and Strelcyk et al. (2009, SCD09). **B**. Simulated ABR W-V latencies compared across available models: Verhulst et al. (2015; V15), Zilany et al. (2014) with Nelson and Carney (2004; ZBC14), Rønne et al. (2012; RDHE12) and Dau (2003; D03). **C.** Simulated ABR wave-I and wave-V amplitude as a function of stimulus level for a normal-hearing (green; NH), a sloping high-frequency hearing loss (red; HI) and a selective synaptopathy model (dashed; HSR). Cochlear gain loss steepened the ABR amplitude growth curve, whereas synaptopathy resulted in shallower supra-threshold slopes. Human ABR reference data came from Picton et al. (1981; PSC81) and Dau (2003; D03) and was offset by 30 dB to translate dB HL values into dB peSPL. **D**. Simulated latencies of narrow-band ABR W-I and W-V derived from responses to clicks masked by a high-pass noise with increasing cut-off frequencies (Don and Eggermont, 1978). Simulated ABR W-I and W-V latencies were offset by 1 and 3.5 ms respectively to ease comparison between the reference data and simulations. The simulations were conducted for clicks of 90 dB peSPL to match the 60 dB SL levels adopted experimentally. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

firing rate (r_{Firing} in Eq. (16)). The f_{m} magnitude was normalized to the average firing rate of the fiber: the DC component (0 Hz) of the Fourier spectrum (Joris et al., 2004). Experimental synchrony-level functions show a non-monotonic relation to the stimulus level and exhibit maxima that occur near the steepest part of AN rate-level curves (e.g., Joris and Yin, 1992). In qualitative agreement with the reference data, the maxima of the simulated synchrony-level curves shifted towards higher levels as the fibers' threshold (and steepest rate-level curve slope) increased. However, the simulated synchrony-level curves overlapped towards the lower levels for the different fiber types where the reference data showed horizontal offsets. Because there is only a very small sound-driven AN activity at these low stimulus levels, the recording accuracy may be less good. Additionally, the model only considered simulations from a single CF whereas the reference data covers fibers from different CFs (and perhaps even different cochleae). These complications render a comparison between model and data inconclusive for very low stimulus levels. A last signature of experimental AN synchrony is that LSR fibers exhibit higher synchrony than HSR fibers for medium-to-high stimulus levels; a feature that was successfully captured. The difference between LSR and HSR synchrony was most pronounced for the 1-kHz simulations (colored dotted curves) that also showed a small “synchrony bump” near 80 dB that coincided with the level at which the BM response returns to linear growth (see Fig. 2C).

Overall, our simulations confirm that modeling the SR-dependence of AN fibers by setting the spontaneous and maximum exocytosis rates (k_{SR} and k_{max}) to match the dynamic range of experimental AN discharge rates, yields realistic AN level and synchrony properties using a minimal number of fitting parameters.

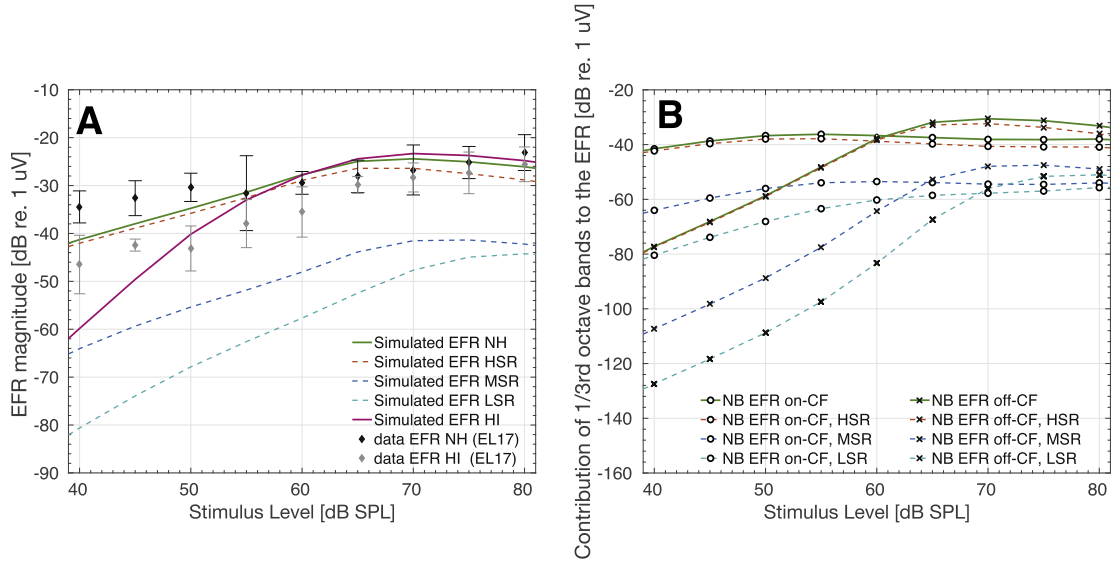

Fig. 7. Simulated EFR characteristics. EFRs were calculated by summing up weighted r_{AN}, r_{CN} and r_{IC} waveforms across the 112 Hz to 12 kHz CF channels, and by calculating the magnitude of the f_m component of the frequency spectrum of the response. **A**. Simulated EFR amplitudes to a 4-kHz 85% amplitude-modulated pure tone (f_m = 98 Hz) as a function of stimulus level. A normal (green) and a sloping high-frequency hearing loss model (red) are considered as well as models that only had contributions of a specific SR fiber type (i.e., 13 HSR, 3 MSR or 3 LSR fibers). The human reference data were recorded to 85% modulated 4-kHz AM tones (f_m = 98 Hz) in normal and mildly sloping high-frequency hearing loss participants (Encina-Llamas et al., 2018; EL17). **B**. Narrow-band (1/3 octave) contribution to the EFR amplitude computed by summing up CF contributions surrounding the 4-kHz (on-CF) or 8-kHz (off-CF) frequency regions in response to the same 4-kHz AM pure-tone stimulus (85% modulated, f_m of 98 Hz). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.2. Population response properties

3.2.1. Auditory brainstem responses (ABRs)

Fig. 6A&B show simulated ABR wave-V peak latencies to 80-μs condensation clicks of different intensities and results are compared to human data (Fig. 6A) and simulations (Fig. 6B) from other ABR (Dau, 2003; Rønne et al., 2012; Verhulst et al., 2015) and AN preprocessing models (Zilany et al., 2014 combined with Nelson and Carney, 2004). Simulated ABR wave-V latencies were overall shorter than found experimentally as the present model did not include all existing synaptic delays in the human brainstem. The ABR models by Dau (2003) and Rønne et al. (2012) compensate for these synaptic delays by adopting a unitary-response approach (Melcher and Kiang, 1996) to translate summed AN responses to different ABR waves (Fig. 6B). We decided not to compensate for this latency discrepancy, but to offset the simulated ABR W-V latency by 3.5 ms in Fig. 6A to overlay the simulations with experimental data and enable an evaluation of the slope characteristics. Note that this offset was arbitrary and not applied when adding the AN (W-I), CN (W-III) and IC (W-V) contributions to the simulated EFR. Simulated ABR W-V latencies showed a 0.9 ms decrease between 50 and 90 dB peSPL that reduced to 0.7 ms between 60 and 100 dB peSPL (Fig. 6A). The simulated W-V latency did not decrease beyond 90 dB peSPL because simulated cochlear click excitation patterns did not widen much beyond this level and the ABR thus stopped recruiting from short onset-latency, high-CF fibers. In comparison, the simulated W-I exhibited a 1.2 ms latency decrease between 50 and 90 dB (results not shown) that is more in line with human W-V latency decrease observations reflecting a 1.2–2 ms reduction over a 40-dB stimulus level increase (Prosser and Arslan, 1987; Serpanos et al., 1997; Dau, 2003; Strelcyk et al., 2009; Elberling et al., 2010). In this model, the difference between W-I and W-V latency-intensity slopes stemmed from how the interplay of excitation/inhibition time-constants in the CN and IC neurons (i.e., the 100-Hz modulation bandpass filter in Eq. (17)) ended up smoothing the W-I response peak, and consequently washing out W-III and W-V latency differences in ascending processing stages.

The simulated W-V latency slope was less shallow (and therefore in lesser agreement with experimental slopes) than for the Verhulst et al. (2015) model (Fig. 6B) which resulted from differences in the CF-dependence of the simulated AN firing thresholds (Fig. 4E). The present AN firing thresholds were more realistic than those of the 2015 model which showed a steep high-frequency threshold roll-off. This steeper roll-off maximized the additional recruitment of short-latency, high-CF fibers as the stimulus level increased and thereby resulted in larger latency changes across level. In contrast, the present model had a shallower (and more realistic) AN firing threshold curve in which high-CF fiber recruitment was dominated by how cochlear click excitation patterns changed as a function of level. Because the model parameters were never adjusted to fit ABR latency characteristics directly, the improvement in modeling the frequency-dependence of the threshold of hearing resulted in a worsened intensity dependence of simulated ABR W-V latency slopes. With the given model stages included, it is necessary to compromise between matching ABR latency differences or the CF-dependence of the hearing threshold. The need for this compromise might point to the importance of including additional ABR brainstem pathways or nonlinearities in the auditory neurons to the model, which are not presently captured by the basic CN/IC model. Discrepancies between recordings and simulations could be further investigated to improve future modeling efforts. Because the present model was evaluated for EFRs, which often have 4-kHz carrier frequencies, we found it more important to model a realistic hearing threshold at 4 kHz than to account for steeper ABR latency-intensity characteristics. Further, given that experimental data show considerable variability in the ABR latency slope estimates (Fig. 6A), which partly reflect spectral differences between the adopted transient stimuli, the simulated 0.7 ms shift for a 40-dB stimulus level increase still falls within the bounds of the experimental slope estimates reported in Picton et al. (1981).

Simulated wave-V amplitude growth corroborated human ABR

recordings (Fig. 6C), and both synaptopathy and cochlear gain loss had an impact on the amplitude-intensity slope. Selective synaptopathy of the LSR and MSR fibers (dashed lines; HSR contributions only) resulted in gradually shallower ABR W-I and W-V amplitude growth than observed for the NH model, in line with observations in several synaptopathy studies (Lin et al., 2011; Furman et al., 2013; Möhrle et al., 2016). At the same time, selective synaptopathy did not influence the ABR latency-intensity slope (Fig. 6A), which reflects the minimal contribution of the delayed onset-peaks of LSR fibers to the population response in the model and observed experimentally (Bourien et al., 2014). The HI simulations (i.e., cochlear gain loss without synaptopathy) for a high-frequency sloping hearing loss resulted in a steeper latency-intensity curve (Fig. 6A). This latency-intensity curve steepening is characteristic for listeners with a high-frequency sloping hearing loss (Gorga et al., 1984; Strelcyk et al., 2009; Verhulst et al., 2016) and also yields steeper ABR amplitude growth (Fig. 6C; Verhulst et al., 2016) caused by elevated hearing and ABR thresholds and more linear v_{BM} response growth.

3.2.2. Masked ABRs

Another approach to uncover the mechanisms of ABR generation that can be used to evaluate the quality of the model, is an experiment that determines how narrow-band frequency regions contribute to the ABR latency (Don and Eggermont, 1978; Strelcyk et al., 2009). 170-μsec clicks were presented at 90 dB peSPL and were masked with high-pass pink noise that had cut-off frequencies between 0.75 and 8 kHz in 1/2-octave steps. Narrow-band ABRs were derived by subtracting response waveforms to clicks that had half an octave difference in masker noise cut-off frequency. Fig. 6D depicts derived narrow-band ABR wave-I and wave-V latencies as a function of CF. The simulated ABR wave-I and V latency showed a decreasing trend as the cut-off frequency increased, in line with experimental observations (Don and Eggermont, 1978; Strelcyk et al., 2009). The W-I and W-V simulations were offset by 1 and 3.5 ms to ease comparison with the experimental data and show that the model simulated the latency-shift as a function of noise cut-off frequency well above 1 kHz. However, the model underestimated the experimental W-V latency shift for frequencies below 1 kHz. Because this was only observed for the W-V simulations, there is likely no cochlear origin for this discrepancy, but rather a brainstem synaptic or neural mechanism that was not fully captured by the basic CN bushy cell and IC model employed. On the basis of the ABR simulations, we conclude that the present model captures level-dependent and broadband characteristics of W-I across the frequency range, and of W-V for frequencies beyond 1 kHz. Lastly, the model captures how cochlear gain loss and synaptopathy are known to affect ABR waveform characteristics.

3.2.3. Envelope-following responses (EFRs)

EFRs were simulated by summing up the signals $A_{W-I}\, r_{AN}$, $A_{W-III}\, r_{CN}$ and $A_{W-V}\, r_{IC}$ across CFs between 112 Hz and 12 kHz to capture the different generators to the population response. The stimulus was a 85% AM tone (CF = 4 kHz) that had a modulation frequency of 98 Hz. Both the intensity dependence of EFRs as well as the contribution of different frequency regions to the response were studied. The simulations were compared to reference data from normal and hearing-impaired listeners in Encina-Llamas et al. (CF = 4 kHz, 85% f_m = 98 Hz; 2018). EFR magnitudes were calculated from the f_m component of the Fourier spectrum of the simulated time-domain signal and plotted in Fig. 7A as a function of stimulus level. Simulated EFRs grew in magnitude as the stimulus level increased up to approximately 80 dB SPL, in agreement with various experimental observations (Kuwada et al., 1986; Dolphin and Mountain, 1992; Dolphin, 1996; Encina-Llamas et al., 2018). This growth can be explained by the rise of AN synchrony with increasing SPL when different fiber types are available (Fig. 5B). At the same time, progressively more off-CF fibers can contribute to the EFR generation as cochlear excitation patterns broaden due to intensity increases (Encina-Llamas et al., 2018). This off-CF recruitment is further illustrated in Fig. 7B that shows how two 1/3 octave band cochlear regions (i.e., 4 kHz; on-CF, and 8 kHz; off-CF) contributed to the simulated EFR magnitude in response to the 4-kHz AM tone of Fig. 7A. HSR fibers in the 8-kHz region contributed more to the EFR generation than the fibers in the on-CF region for stimulus levels above 60 dB (Fig. 7B, red dashed traces with crosses and circles for the 1/3 octave band HSR contribution). This resulted from a saturation of on-CF fibers (Fig. 5A) at higher stimulus levels, while the broader BM excitation patterns yielded a progressive recruitment of off-CF fibers to the EFR. The magnitude decrease in the broadband EFR for stimulus levels above 70 dB (Fig. 7A) resulted from this interplay of effects that entails saturation of on-CF fibers and the recruitment of off-CF fibers.

Cochlear gain loss and synaptopathy differentially impacted the simulated EFRs. A loss of LSR/MSR fibers (Fig. 7A, red dashed HSR trace) reduced the overall EFR magnitude in agreement with synaptopathy observations in mice (Shaheen et al., 2015), and caused a high-level roll-off of EFR magnitude (Fig. 7A). Cochlear gain loss (red solid curve) elevated the EFR magnitude above 65 dB SPL and yielded steeper magnitude growth below 60 dB SPL in agreement with recorded EFRs for levels above 50 dB SPL (diamonds in Fig. 7A; Encina-Llamas et al., 2018) and observations in chinchilla (Zhong et al., 2014). The steeper EFR growth at low levels resulted from the hearing-threshold shift that caused a recruitment effect of the EFR. The stronger high-level EFRs for the HI model resulted from a lower effective drive to the AN fibers that yielded more unsaturated AN fiber responses (for on-CF channels) compared to the NH model for the same stimulus level (i.e., as in subtracting the dB HL loss amount from the x-axis values in Fig. 5).

3.3. Model compromises

During the development of this model, we compromised in determining the value of several fitting parameters corresponding to different levels of processing. These model compromises either yield open questions that can direct future experimental studies, or motivate the need to add complexity to specific model stages.

(i) Using a 1-D transmission-line cochlear model in which α^*_A was set on the basis of human OAE-derived cochlear tuning, it was not possible to achieve intensity-invariant zero-crossings while maintaining a strong compression of v_{BM} up to high stimulus levels. However, small rodents do show v_{BM} compression characteristics upto high SPLs (see Robles and Ruggero, 2001). The limited compression range in the model results from the maximal amount of cochlear mechanical gain (35 dB) that can be applied, and that underestimates the larger gain amounts reported for small rodents (Rhode, 2007). However, as cochlear gain and tuning are a correlated phenomenon in transmission-line models (different from filter-bank architectures), our model simulations predict that the encoded gain range is not the same in humans as for laboratory animals given their cochlear tuning differences (see Shera et al., 2010; Joris et al., 2011; Raufer and Verhulst, 2016 for species comparisons of Q_{ERB}). To further explore whether human encoded gain ranges and v_{BM} growth are indeed different than reported for rodents, human OAE paradigms that estimate cochlear compression (e.g. Neely et al., 2009; Gorga et al., 2011a,b) could be simulated while relating their characteristics to the v_{BM}

simulations in the same model. This method could further validate whether the transmission-line architecture is a good basis to collectively capture the functional tuning, compression and gain aspects of human BM processing. If this analysis shows that human compression is indeed different than reported for rodents, it is possible to increase α^*_P for the human constrained α^*_A values to model a larger dynamic range, but at the cost of misaligned BM impulse response zero-crossing for the highest stimulus levels. Given that human CEOAE waveforms only show mild drifts in their waveform zero-crossings (e.g., Verhulst et al., 2011), we opted to maintain the BM zero-crossing invariance criterion across the entire intensity range and ended up with a model compromise in which cochlear processing shows linear growth for stimulus levels above 80 dB SPL (at 1 kHz).

(ii) At the level of the AN, it was not possible to obtain perfect near *and* supra-threshold response characteristics. The cochlear compression threshold of 30 dB at the 1-kHz CF ($\nu_{BM,30dB}$) influences how AN fibers operate near their firing threshold. When matching the HSR firing threshold perfectly to the human threshold of hearing (i.e., 10 dB at 1 kHz; Fig. 4E), it was not possible to capture reference AN firing threshold differences between LSR and HSR fibers (results not shown). The failure to capture both the HSR threshold and LSR/HRS threshold differences simultaneously results from that *both* fiber types see linear processing until their respective thresholds are reached, hence the difference between the AN thresholds across fiber types are minimal. We compromised by setting the $d_{IHCcilia}/\nu_{BM}$ parameter such that the HSR fiber threshold occurred at 20 dB and closer to the $\nu_{BM,30dB}$ compression threshold to yield more realistic AN threshold differences across the different fiber types (see Fig. 4E).

(iii) The brainstem pathways responsible for ABR and EFR generation are very basic in the model as it is presently not clear which brainstem pathways and cell-types are crucial contributors. We opted to include a bushy cell model based on the lesion studies performed by Melcher and Kiang (1996). The functionality of the adopted bushy cell model results in a 100-Hz modulation filtering and onset sharpening process (Nelson and Carney, 2004). Modulation filtering and onset sharpening have both experimentally been observed in CN and IC neurons (Delgutte et al., 1998; Frisina et al., 1990) and are important for peripheral auditory processing (e.g., Dau et al., 1997). However, it is likely that bushy cells are not the only contributors to ABR and EFR generation (e.g., Schaette and McAlpine, 2011; Bidelman, 2015) and that across-CF projections of neurons might be important as well. We opted to keep the model back-end simple, as it serves to evaluate the broadband and level-dependent characteristics of human AN processing. Future brainstem functionality and morphology studies can be directed to study the importance of different neuronal populations (and their CF projections) for ABR and EFR generation. To support these studies, the model can identify which stimulus paradigms cause mismatches between simulated and recorded population responses and thereby improve the functional descriptions of the relevant auditory brainstem structures.

4. Conclusion

The presented model for the human auditory periphery yields satisfactory predictions of key AN properties such as adaptation, offset recovery time-constants, rate- and synchrony-level functions. This result was achieved using 12 fitting parameters, and without including power-law adaptation or CF-dependencies in the IHC-AN complex parameters. The only parameter that was rendered CF-dependent related to the cochlear frequency tuning. When connecting the developed broadband AN model to a functional model of CN and IC neurons, population responses can be simulated to evaluate the broadband and level-dependent model characteristics using human reference data. The simulated population responses captured several key aspects of human ABR and EFR recordings in normal and hearing-impaired listeners well. Examples are: (i) a 0.7-ms ABR wave-V latency decrease following a 40-dB stimulus level increase, (ii) compressive ABR amplitude growth and the ways in which OHC loss steepens the amplitude growth and renders it shallower for simulated synaptopathy, and (iii), realistic EFR magnitude growth that shows a high-intensity roll-off when selective LSR/MSR synaptopathy occurs and a steeper EFR growth for lower intensities when OHC loss is simulated.

We chose to keep the model as simple as possible to capture a large range of human auditory processing aspects while avoiding the overfitting of specific physiological datasets. Given that the best model would be one that describes key auditory processing features and its impairments with the fewest possible parameters, we learned from the success of basic functional auditory models and converged their description with detailed physiological processes into one model of human peripheral hearing that has 12 fitting parameters. In our opinion, this cautious approach of preferring models based on biophysical properties over empirical functions fitted to specific (and therefore limited) responses of laboratory animals, will continue to be successful when over time the complex and elusive aspects of the inner-ear physiology will be clarified experimentally. In time, this will hopefully lead to a concise and elegant theory of the functioning of peripheral hearing. *The model code is available on:* http://www.github.com/HearingTechnology/Verhulstetal2018Model *(email the corresponding author to for access) or* http://www.waves.intec.ugent.be/hearing-technology.

Acknowledgements

This work was supported by Academy of Finland, the Aalto ELEC doctoral school (AA) the German Research Foundation (DFG Priority Program 1608 ″Ultrafast and temporally precise information processing: Normal and dysfunctional hearing" VE924/1-1 (VV) and European Research Counsil (ERC) under the Horizon 2020 Research and Innovation Programme (grant agreement No 678120 RobSpear; SV). We thank Gerard Encina-Llamas for providing the human reference EFR data, Laurel Carney and Christopher Shera for numerous in-depth discussions on cochlear and AN modeling.

Appendix A. the IHC transduction model

The IHC model follows the description in Altoè et al. (2017), with the difference that the activation of the MET channels is described by a double Boltzmann function here rather than by a single Boltzmann function. This choice was motivated by the different application ranges in the two studies: Altoè et al. (2017) focused on small levels of stereocilia deflection for which single or double Boltzmann functions would produce similar results. Here, we focussed on modeling neural responses up to high SPLs, where the employment of a single Boltzmann function would - even at moderate SPLs - produce unrealistically strong V_{IHC} saturation. A summary of the model equations is given here, and the reader is referred to Altoè et al. (2017) for additional details on the model principles and the parameter selection. The IHC receptor potential (V_{IHC}) is given by the interplay between K^+ and mechanoelectrical transduction (MET) currents:

$$-C_m\frac{dV_{IHC}}{dt} = I_{MET} + I_{K,f} + I_{K,s}, \quad \text{(A1)}$$

where C_m is the IHC membrane capacitance and I_{MET}, $I_{K,f}$ and $I_{K,s}$ correspond to the inward MET and outward fast and slow-activating K^+ currents, respectively. The activation of the MET current is described using a standard channel equation with fast-activation kinetics:

$$n_{MET,\infty} = \left[1 + e^{-\frac{u-x_0}{s_0}}\left(1 + e^{-\frac{u-x_0}{s_1}}\right)\right]^{-1}, \quad \text{(A2)}$$

$$n_{MET} + \tau_{MET}\frac{dn_{MET}}{dt} = n_{\infty,MET}, \quad \text{(A3)}$$

$$I_{MET} = n_{MET}G_{MET,max}(V_{IHC} - EP), \quad \text{(A4)}$$

where u corresponds to the cilia displacement and $\tau_{MET} = 50$ μs to the activation time (Kennedy et al., 2003). $G_{MET,max}$ (30 nS) is the maximal conductance of the MET channels (Jia et al., 2007; Johnson et al., 2011) and EP the endocochlear potential. To our knowledge, there are no published double Boltzmann fits to the steady-state activation of the MET channels for the mature gerbil IHCs in near-physiological levels of Ca^{2+} surrounding of the stereocilia. The parameters were therefore constrained to physiological data in the following way. First, s_1 was set to 16 nm based on the single Boltzmann fit to a basal IHC in Jia et al. (2007). Second, s_0 was set to be three times larger than s_1 to obtain a ratio between these two parameters that reasonably matches recordings in immature hair cells of the guinea-pig (Russell et al., 1986; Corns et al., 2014) and mouse (Kros et al., 1992). Finally, once s_0 and s_1 were determined, x_0 was set to 20 nm to obtain about 13% of the MET channels to be open at rest (Jia et al., 2007).

The equations of the kinetics of the fast and slow IHC basolateral K^+ channels were based on experimental data and entail a number of simplifications (Altoè et al., 2017). First, the voltage dependency of the channel-activation time constants was omitted as the time constants should not vary dramatically for V_{IHC} excursions in the physiological range. Second, the second-order equations describing the activation of K^+ channels were approximated by single-order equations. This approximation was justified theoretically in Altoè et al. (2017), and shown to approximate the experimental process reasonably (Marcotti et al., 2004). Finally, to employ recordings from mature gerbil cells (Johnson et al., 2011) in near-physiological levels of Ca^{2+} and mobile buffers (which affect the functioning of the K^+ channels), the fast and slow channels were assumed to have similar activation kinetics and conductances. The resulting equations are:

$$n_{K,F/S,\infty} = \left(1 + e^{-\frac{V_{IHC}-V_{0.5}}{s_k}}\right)^{-1}, \quad \text{(A5)}$$

$$n_{K,F/S} + \tau_{K,F/S}\frac{dn_{K,F/S}}{dt} = n_{K,F/S,\infty}, \quad \text{(A6)}$$

$$I_{K,F/S} = n_{K,F/S}\,G_{K,F/S}\left(V_{IHC} - V_{K,F/S}\right), \quad \text{(A7)}$$

where the subscripts F/S refer to the fast or the slow channels properties respectively. In these equations, $\tau_{K,F/S}$ is the principal activation time constant of the channels, and $G_{K,F/S}$ and $V_{K,F/S}$ refer to the maximal conductance and reversal potential of the channels, respectively.

References

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

Altoè, A., Pulkki, V., Verhulst, S., 2014. Transmission line cochlear models: improved accuracy and efficiency. J. Acoust. Soc. Am. 136 (4), EL302–EL308.

Altoè, A., Pulkki, V., Verhulst, S., 2017. Model-based estimation of the frequency-tuning of the inner hair cell stereocilia from neural tuning curves. J. Acoust. Soc. Am. 141.

Altoè, A., Pulkki, V., Verhulst, S. (submitted). The effects of the activation of the inner-hair-cell basolateral K^+ channels on auditory nerve responses. Hear. Res

Beutner, D., Voets, T., Neher, E., Moser, T., 2001. Calcium dependence of exocytosis and endocytosis at the cochlear inner hair cell afferent synapse. Neuron 29 (3), 681–690.

Bharadwaj, H.M., Shinn-Cunningham, B.G., 2014. Rapid acquisition of auditory subcortical steady state responses using multichannel recordings. Clin. Neurophysiol. 125 (9), 1878–1888.

Bharadwaj, H.M., Verhulst, S., Shaheen, L., Liberman, M.C., Shinn-Cunningham, B.G., 2014. Cochlear neuropathy and the coding of supra-threshold sound. Front. Syst. Neurosci. 8.

Bharadwaj, H.M., Masud, S., Mehraei, G., Verhulst, S., Shinn-Cunningham, B.G., 2015. Individual differences reveal correlates of hidden hearing deficits. J. Neurosci. 35 (5), 2161–2172.

Bidelman, G.M., 2015. Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient ABR. Hear. Res. 323, 68–80.

Bourien, J., Tang, Y., Batrel, C., Huet, A., Lenoir, M., Ladrech, S., et al., 2014. Contribution of auditory nerve fibers to compound action potential of the auditory nerve. J. Neurophysiol. 112 (5), 1025–1039.

Chapochnikov, N.M., Takago, H., Huang, C.H., Pangršič, T., Khimich, D., Neef, J., et al., 2014. Uniquantal release through a dynamic fusion pore is a candidate mechanism of hair cell exocytosis. Neuron 83 (6), 1389–1403.

Cheatham, M.A., Dallos, P., 1999. Response phase: a view from the inner hair cell. J. Acoust. Soc. Am. 105 (2), 799–810.

Corns, L.F., Johnson, S.L., Kros, C.J., Marcotti, W., 2014. Calcium entry into stereocilia drives adaptation of the mechanoelectrical transducer current of mammalian cochlear hair cells. Proceed. Natl Acad. Sci. U. S. A. 111 (41), 14918–14923.

Dau, T., 2003. The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. J. Acoust. Soc. Am. 113 (2), 936–950.

Dau, T., Kollmeier, B., Kohlrausch, A., 1997. Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. J. Acoust. Soc. Am. 102 (5), 2892–2905.

Davis, H., 1965. A model for transducer action in the cochlea. In: Cold Spring Harbor Symposia on Quantitative Biology, Vol. 30. Cold Spring Harbor Laboratory Press, pp. 181–190.

de Boer, E., 1980. Auditory physics. Physical principles in hearing theory. I. Phys. Rep. 62 (2), 87–174.

Delgutte, B., Hammond, B.M., Cariani, P.A., 1998. Neural coding of the temporal envelope of speech: relation to modulation transfer functions. In: Palmer, A.R., Rees, A., Summerfield, A.Q., Meddis, R. (Eds.), Psychophysical and Physiological Advances in Hearing. Whurr, London, pp. 595–603.

Dolphin, W.F., 1996. Auditory evoked responses to amplitude modulated stimuli consisting of multiple envelope components. J. Comp. Physiol. A 179 (1), 113–121.

Dolphin, W.F., Mountain, D.C., 1992. The envelope following response: scalp potentials elicited in the Mongolian gerbil using sinusoidally AM acoustic signals. Hear. Res. 58 (1), 70–78.

Don, M., Eggermont, J.J., 1978. Analysis of the click-evoked brainstem potentials in man using high-pass noise masking. J. Acoust. Soc. Am. 63 (4), 1084–1092.

Duifhuis, H., 2012. Cochlear Mechanics: Introduction to a Time Domain Analysis of the Nonlinear Cochlea. Springer Science & Business Media.

Elberling, C., Callø, J., Don, M., 2010. Evaluating auditory brainstem responses to different chirp stimuli at three levels of stimulation. J. Acoust. Soc. Am. 128 (1), 215–223.

Elliott, S.J., Ku, E.M., Lineton, B., 2007. A state space model for cochlear mechanics. J. Acoust. Soc. Am. 122 (5), 2759–2771.

Encina-Llamas, G., Dau, T., Epp, B., 2007. Estimates of peripheral compression using envelope following responses. J. Assoc. Res. Otolaryngol.

Epp, B., Verhey, J.L., Mauermann, M., 2010. Modeling cochlear dynamics: interrelation between cochlea mechanics and psychoacoustics a. J. Acoust. Soc. Am. 128 (4), 1870–1883.

Ewert, S.D., Dau, T., 2000. Characterizing frequency selectivity for envelope fluctuations. J. Acoust. Soc. Am. 108 (3), 1181–1196.

Frank, T., Khimich, D., Neef, A., Moser, T., 2009. Mechanisms contributing to synaptic Ca2+ signals and their heterogeneity in hair cells. Proc Natl Acad. Sci. 106 (11), 4483–4488.

Frisina, R.D., Smith, R.L., Chamberlain, S.C., 1990. Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. Hear. Res. 44 (2),

99–122.
Furman, A.C., Kujawa, S.G., Liberman, M.C., 2013. Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. J. Neurophysiol. 110 (3), 577–586.
Goldberg, J.M., Brown, P.B., 1969. Response of binaural neurons of dog. J. Neurophysiol. 32, 613–636.
Gorga, M.P., Worthington, D.W., Reiland, J.K., Beauchaine, K.A., Goldgar, D.E., 1984. Some comparisons between auditory brain stem response thresholds, latencies, and the pure-tone audiogram. Ear Hear. 6 (2), 105–112.
Gorga, M.P., Neely, S.T., Kopun, J., Tan, H., 2011a. Growth of suppression in humans based on distortion-product otoacoustic emission measurements. J. Acoust. Soc. Am. 129 (2), 801–816.
Gorga, M.P., Neely, S.T., Kopun, J., Tan, H., 2011b. Distortion-product otoacoustic emission suppression tuning curves in humans. J. Acoust. Soc. Am. 129 (2), 817–827.
Goutman, J.D., Glowatzki, E., 2007. Time course and calcium dependence of transmitter release at a single ribbon synapse. Proc Natl Acad. Sci. 104 (41), 16341–16346.
Grant, L., Yi, E., Glowatzki, E., 2010. Two modes of release shape the postsynaptic response at the inner hair cell ribbon synapse. J. Neurosci. 30 (12), 4210–4220.
Greenwood, D.D., 1990. A cochlear frequency-position function for several species—29 years later. J. Acoust. Soc. Am. 87 (6), 2592–2605.
Han, L., Poulsen, T., 1998. Equivalent threshold sound pressure levels for Sennheiser HDA 200 earphone and Etymotic Research ER-2 insert earphone in the frequency range 125 Hz to 16 kHz. Scand. Audiol. 27 (2), 105–112.
Harris, D.M., Dallos, P., 1979. Forward masking of auditory nerve fiber responses. J. Neurophysiol. 42 (4), 1083–1107.
Heil, P., Neubauer, H., 2010. Summing across different active zones can explain the quasi-linear Ca2+-dependencies of exocytosis by receptor cells. Front. Synaptic neurosci. 2.
Heinz, M.G., Zhang, X., Bruce, I.C., Carney, L.H., 2001. Auditory nerve model for predicting performance limits of normal and impaired listeners. Acoust Res. Lett. Online 2 (3), 91–96.
Hudspeth, A.J., Lewis, R.S., 1988. A model for electrical resonance and frequency tuning in saccular hair cells of the bull-frog, Rana catesbeiana. J. Physiol. 400 (1), 275–297.
Huet, A., Batrel, C., Tang, Y., Desmadryl, G., Wang, J., Puel, J.L., Bourien, J., 2016. Sound coding in the auditory nerve of gerbils. Hear. Res. 338, 32–39.
Jepsen, M.L., Dau, T., 2011. Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss. J. Acoust. Soc. Am. 129 (1), 262–281.
Jepsen, M.L., Ewert, S.D., Dau, T., 2008. A computational model of human auditory signal processing and perception. J. Acoust. Soc. Am. 124 (1), 422–438.
Jia, S., Dallos, P., He, D.Z., 2007. Mechanoelectric transduction of adult inner hair cells. J. Neurosci. 27 (5), 1006–1014.
Jiang, Z.D., Zheng, M.S., Sun, D.K., Liu, X.Y., 1991. Brainstem auditory evoked responses from birth to adulthood: normative data of latency and interval. Hear. Res. 54 (1), 67–74.
Johnson, S.L., Marcotti, W., 2008. Biophysical properties of CaV1. 3 calcium channels in gerbil inner hair cells. J. Physiol. 586 (4), 1029–1042.
Johnson, S.L., Beurg, M., Marcotti, W., Fettiplace, R., 2011. Prestin-driven cochlear amplification is not limited by the outer hair cell membrane time constant. Neuron 70 (6), 1143–1154.
Jørgensen, S., Ewert, S.D., Dau, T., 2013. A multi-resolution envelope-power based model for speech intelligibility. J. Acoust. Soc. Am. 134 (1), 436–446.
Johnson, S.L., 2015. Membrane properties specialize mammalian inner hair cells for frequency or intensity encoding. Elife 4, e08177.
Joris, P.X., Yin, T.C., 1992. Responses to amplitude-modulated tones in the auditory nerve of the cat. J. Acoust. Soc. Am. 91 (1), 215–232.
Joris, P.X., Schreiner, C.E., Rees, A., 2004. Neural processing of amplitude-modulated sounds. Physiol. Rev. 84 (2), 541–577.
Joris, P.X., Bergevin, C., Kalluri, R., Mc Laughlin, M., Michelet, P., van der Heijden, M., Shera, C.A., 2011. Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proc Natl Acad. Sci. 108 (42), 17516–17520.
Jürgens, T., Clark, N.R., Lecluyse, W., Meddis, R., 2016. Exploration of a physiologically-inspired hearing-aid algorithm using a computer model mimicking impaired hearing. Int. J. Audiol. 55 (6), 346–357.
Kapadia, S., Lutman, M.E., 2000. Nonlinear temporal interactions in click-evoked otoacoustic emissions. II. Experimental data. Hear. Res. 146 (1), 101–120.
Kennedy, H.J., Evans, M.G., Crawford, A.C., Fettiplace, R., 2003. Fast adaptation of mechanoelectrical transducer channels in mammalian cochlear hair cells. Nat. Neurosci. 6 (8), 832–836.
Kiang, N.Y.S., Baer, T., Marr, E.M., Demont, D., 1969. Discharge rates of single auditory nerve fibers as functions of tone level. J. Acoust. Soc. Am. 46 (1A), 106–106.
Kidd, R.C., Weiss, T.F., 1990. Mechanisms that degrade timing information in the cochlea. Hear. Res. 49 (1), 181–207.
Kros, C.J., Crawford, A.C., 1990. Potassium currents in inner hair cells isolated from the guinea-pig cochlea. J. Physiol. 421 (1), 263–291.
Kros, C.J., Rusch, A., Richardson, G.P., 1992. Mechano-electrical transducer currents in hair cells of the cultured neonatal mouse cochlea. Proceed. Roy. Soc. Lon. B Biol. Sci. 249 (1325), 185–193.
Kujawa, S.G., Liberman, M.C., 2009. Adding insult to injury: cochlear nerve degeneration after “temporary” noise-induced hearing loss. J. Neurosci. 29 (45), 14077–14085.
Kuwada, S., Batra, R., Maher, V.L., 1986. Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones. Hear. Res. 21 (2), 179–192.
Lewis, J.D., Neely, S.T., 2015. Non-invasive estimation of middle-ear input impedance and efficiency a. J. Acoust. Soc. Am. 138 (2), 977–993.
Liberman, M.C., 1978. Auditory-nerve response from cats raised in a low-noise chamber. J. Acoust. Soc. Am. 63 (2), 442–455.
Liberman, M.C., Epstein, M.J., Cleveland, S.S., Wang, H., Maison, S.F., 2016. Toward a differential diagnosis of hidden hearing loss in humans. PLoS One 11 (9), e0162726.
Liberman, L.D., Wang, H., Liberman, M.C., 2011. Opposing gradients of ribbon size and AMPA receptor expression underlie sensitivity differences among cochlear-nerve/hair-cell synapses. J. Neurosci. 31 (3), 801–808.
Lin, H.W., Furman, A.C., Kujawa, S.G., Liberman, M.C., 2011. Primary neural degeneration in the Guinea pig cochlea after reversible noise-induced threshold shift. J. Assoc. Res. Otolaryngol. 12 (5), 605–616.
Liu, Y.W., 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.
Lopez-Poveda, E.A., Eustaquio-Martín, A., 2006. A biophysical model of the inner hair cell: the contribution of potassium currents to peripheral auditory compression. J. Assoc. Res. Otolaryngol. 7 (3), 218–235.
Lynch III, T.J., Nedzelnitsky, V., Peake, W.T., 1982. Input impedance of the cochlea in cat. J. Acoust. Soc. Am. 72 (1), 108–130.
Lyon, R.F., 2011. Cascades of two-pole–two-zero asymmetric resonators are good models of peripheral auditory function. J. Acoust. Soc. Am. 130 (6), 3893–3904.
Manley, G.A., Fay, R.R., 2007. Active Processes and Otoacoustic Emissions in Hearing, vol. 30. Springer Science & Business Media.
Mao, J., Carney, L.H., 2015. Tone-in-noise detection using envelope cues: comparison of signal-processing-based and physiological models. J. Assoc. Res. Otolaryngol. 16 (1), 121–133.
Marcotti, W., Johnson, S.L., Kros, C.J., 2004. A transiently expressed SK current sustains and modulates action potential activity in immature mouse inner hair cells. J. Physiol. 560 (3), 691–708.
Meaud, J., Grosh, K., 2010. The effect of tectorial membrane and basilar membrane longitudinal coupling in cochlear mechanics. J. Acoust. Soc. Am. 127 (3), 1411–1421.
Meddis, R., 1986. Simulation of mechanical to neural transduction in the auditory receptor. J. Acoust. Soc. Am. 79 (3), 702–711.
Meddis, R., 2006. Auditory-nerve first-spike latency and auditory absolute threshold: a computer model. J. Acoust. Soc. Am. 119 (1), 406–417.
Meddis, R., O'Mard, L., 1997. A unitary model of pitch perception. J. Acoust. Soc. Am. 102 (3), 1811–1820.
Meddis, R., O'Mard, L.P., Lopez-Poveda, E.A., 2001. A computational algorithm for computing nonlinear auditory frequency selectivity. J. Acoust. Soc. Am. 109 (6), 2852–2861.
Mehraei, G., Hickox, A.E., Bharadwaj, H.M., Goldberg, H., Verhulst, S., Liberman, M.C., Shinn-Cunningham, B.G., 2016. Auditory brainstem response latency in noise as a marker of cochlear synaptopathy. J. Neurosci. 36 (13), 3755–3764.
Melcher, J.R., Kiang, N.Y., 1996. Generators of the brainstem auditory evoked potential in cat. III: identified cell populations. Hear. Res. 93 (1–2), 52–71.
Meyer, A.C., Frank, T., Khimich, D., Hoch, G., Riedel, D., Chapochnikov, N.M., Yarin, Y.M., Harke, B., Hell, S.W., Egner, A., Moser, T., 2009. Tuning of synapse number, structure and function in the cochlea. Nat. Neurosci. 12 (4), 444–453.
Miller, C.A., Abbas, P.J., Robinson, B.K., 2001. Response properties of the refractory auditory nerve fiber. J. Assoc. Res. Otolaryngol. 2 (3), 216–232.
Moezzi, B., Iannella, N., McDonnell, M.D., 2016. Ion channel noise can explain firing correlation in auditory nerves. J. Comput. Neurosci. 41 (2), 193–206.
Möhrle, D., Ni, K., Varakina, K., Bing, D., Lee, S.C., Zimmermann, U., et al., 2016. Loss of auditory sensitivity from inner hair cell synaptopathy can be centrally compensated in the young but not old brain. Neurobiol. Aging 44, 173–184.
Moleti, A., Paternoster, N., Bertaccini, D., Sisto, R., Sanjust, F., 2009. Otoacoustic emissions in time-domain solutions of nonlinear non-local cochlear models. J. Acoust. Soc. Am. 126 (5), 2425–2436.
Nedzelnitsky, V., 1980. Sound pressures in the basal turn of the cat cochlea. J. Acoust. Soc. Am. 68 (6), 1676–1689.
Neely, S.T., Kim, D.O., 1983. An active cochlear model showing sharp tuning and high sensitivity. Hear. Res. 9 (2), 123–130.
Neely, S.T., Johnson, T.A., Kopun, J., Dierking, D.M., Gorga, M.P., 2009. Distortion-product otoacoustic emission input/output characteristics in normal-hearing and hearing-impaired human ears. J. Acoust. Soc. Am. 126 (2), 728–738.
Nelson, P.C., Carney, L.H., 2004. A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. J. Acoust. Soc. Am. 116 (4), 2173–2186.
Oertel, D., 1983. Synaptic responses and electrical properties of cells in brain slices of the mouse anteroventral cochlear nucleus. J. Neurosci. 3 (10), 2043–2053.
Ohn, T.L., Rutherford, M.A., Jing, Z., Jung, S., Duque-Afonso, C.J., Hoch, G., 2016. Hair cells use active zones with different voltage dependence of Ca2+ influx to decompose sounds into complementary neural codes. Proc Natl Acad. Sci., 201605737
Palmer, A.R., Russell, I.J., 1986. Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. Hear. Res. 24 (1), 1–15.
Pangršič, T., Lasarow, L., Reuter, K., Takago, H., Schwander, M., Riedel, D., et al., 2010. Hearing requires otoferlin-dependent efficient replenishment of synaptic vesicles in hair cells. Nat. Neurosci. 13 (7), 869–876.

Peterson, A.J., Irvine, D.R., Heil, P., 2014. A model of synaptic vesicle-pool depletion and replenishment can account for the interspike interval distributions and nonrenewal properties of spontaneous spike trains of auditory-nerve fibers. J. Neurosci. 34 (45), 15097–15109.
Picton, T.W., 2011. Human auditory evoked potentials. Chapter 8: Auditory brainstem responses: peaks along the way. Plural Publishing, pp. 213–245.
Picton, T.W., Stapells, D.R., Campbell, K.B., 1981. Auditory evoked potentials from the human cochlea and brainstem. J. Otolaryngol. 9, 1–41.
Pieper, I., Mauermann, M., Kollmeier, B., Ewert, S.D., 2016. Physiological motivated transmission-lines as front end for loudness models. J. Acoust. Soc. Am. 139 (5), 2896–2910.
Plack, C.J., Barker, D., Prendergast, G., 2014. Perceptual consequences of "hidden" hearing loss. Trends Hear. 18, 2331216514550621.
Prosser, S., Arslan, E., 1987. Prediction of auditory brainstem wave V latency as a diagnostic tool of sensorineural hearing loss. Audiology 26 (3), 179–187.
Puria, S., Allen, J.B., 1991. A parametric study of cochlear input impedance. J. Acoust. Soc. Am. 89 (1), 287–309.
Puria, S., 2003. Measurements of human middle ear forward and reverse acoustics: implications for otoacoustic emissions. J. Acoust. Soc. Am. 113 (5), 2773–2789.
Raufer, S., Verhulst, S., 2016. Otoacoustic emission estimates of human basilar membrane impulse response duration and cochlear filter tuning. Hear. Res. 342, 150–160.
Recio, A., Rhode, W.S., 2000. Basilar membrane responses to broadband stimuli. J. Acoust. Soc. Am. 108 (5), 2281–2298.
Relkin, E.M., Doucet, J.R., 1991. Recovery from prior stimulation. I: relationship to spontaneous firing rates of primary auditory neurons. Hear. Res. 55 (2), 215–222.
Rhode, W.S., 2007. Basilar membrane mechanics in the 6–9 kHz region of sensitive chinchilla cochleae. J. Acoust. Soc. Am. 121 (5), 2792–2804.
Rhode, W.S., Smith, P.H., 1985. Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory nerve fibers. Hear. Res. 18 (2), 159–168.
Robles, L., Ruggero, M.A., 2001. Mechanics of the mammalian cochlea. Physiol. Rev. 81 (3), 1305–1352.
Rønne, F.M., Dau, T., Harte, J., Elberling, C., 2012. Modeling auditory evoked brainstem responses to transient stimuli. J. Acoust. Soc. Am. 131 (5), 3903–3913.
Rosen, S., Baker, R.J., 1994. Characterising auditory filter nonlinearity. Hear. Res. 73 (2), 231–243.
Ruggero, M.A., Robles, L., Rich, N.C., 1992. Two-tone suppression in the basilar membrane of the cochlea: mechanical basis of auditory-nerve rate suppression. J. Neurophysiol. 68 (4), 1087–1099.
Ruggero, M.A., Rich, N.C., Recio, A., Narayan, S.S., Robles, L., 1997. Basilar-membrane responses to tones at the base of the chinchilla cochlea. J. Acoust. Soc. Am. 101 (4), 2151–2163.
Russell, I.J., Sellick, P.M., 1983. Low-frequency characteristics of intracellularly recorded receptor potentials in guinea-pig cochlear hair cells. J. Physiol. 338 (1), 179–206.
Russell, I.J., Cody, A.R., Richardson, G.P., 1986. The responses of inner and outer hair cells in the basal turn of the guinea-pig cochlea and in the mouse cochlea grown in vitro. Hear. Res. 22 (1–3), 199–216.
Sachs, M.B., Abbas, P.J., 1974. Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli. J. Acoust. Soc. Am. 56 (6), 1835–1847.
Saremi, A., Beutelmann, R., Dietz, M., Ashida, G., Kretzberg, J., Verhulst, S., 2016. A comparative study of seven human cochlear filter models. J. Acoust. Soc. Am. 140 (3), 1618–1634.
Schaette, R., McAlpine, D., 2011. Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model. J. Neurosci. 31 (38), 13452–13457.
Schmiedt, R.A., 2010. The physiology of cochlear presbycusis. In: The Aging Auditory System. Springer, New York, pp. 9–38.
Sellick, P.M., Russell, I.J., 1980. The responses of inner hair cells to basilar membrane velocity during low frequency auditory stimulation in the guinea pig cochlea. Hear. Res. 2 (3), 439–445.
Serpanos, Y.C., O'malley, H., Gravel, J.S., 1997. The relationship between loudness intensity functions and the click-ABR wave V latency. Ear Hear. 18 (5), 409–419.
Shaheen, L.A., Valero, M.D., Liberman, M.C., 2015. Towards a diagnosis of cochlear neuropathy with envelope following responses. J. Assoc. Res. Otolaryngol. 16 (6), 727–745.
Shamma, S.A., Chadwick, R.S., Wilbur, W.J., Morrish, K.A., Rinzel, J., 1986. A biophysical model of cochlear processing: intensity dependence of pure tone responses. J. Acoust. Soc. Am. 80 (1), 133–145.
Shera, C.A., 2001. Frequency glides in click responses of the basilar membrane and auditory nerve: their scaling behavior and origin in traveling-wave dispersion. J. Acoust. Soc. Am. 109 (5), 2023–2034.
Shera, C.A., Zweig, G., 1991. A symmetry suppresses the cochlear catastrophe. J. Acoust. Soc. Am. 89 (3), 1276–1289.
Shera, C.A., Guinan, J.J., Oxenham, A.J., 2010. Otoacoustic estimation of cochlear tuning: validation in the chinchilla. J. Assoc. Res. Otolaryngol. 11 (3), 343–365.
Strelcyk, O., Christoforidis, D., Dau, T., 2009. Relation between derived-band auditory brainstem response latencies and behavioral frequency selectivity. J. Acoust. Soc. Am. 126 (4), 1878–1888.
Sumner, C.J., Lopez-Poveda, E.A., O'Mard, L.P., Meddis, R., 2002. A revised model of the inner-hair cell and auditory-nerve complex. J. Acoust. Soc. Am. 111 (5), 2178–2188.
Sumner, C.J., Lopez-Poveda, E.A., O'Mard, L.P., Meddis, R., 2003. Adaptation in a revised inner-hair cell model. J. Acoust. Soc. Am. 113 (2), 893–901.
Taberner, A.M., Liberman, M.C., 2005. Response properties of single auditory nerve fibers in the mouse. J. Neurophysiol. 93 (1), 557–569.
Takanen, M., Santala, O., Pulkki, V., 2014. Visualization of functional count-comparison-based binaural auditory model output. Hear. Res. 309, 147–163.
Talmadge, C.L., Tubis, A., Long, G.R., Piskorski, P., 1998. Modeling otoacoustic emission and hearing threshold fine structures. J. Acoust. Soc. Am. 104 (3), 1517–1543.
Trautwein, P., Hofstetter, P., Wang, J., Salvi, R., Nostrant, A., 1996. Selective inner hair cell loss does not alter distortion product otoacoustic emissions. Hear. Res. 96 (1), 71–82.
Valero, M.D., Burton, J.A., Hauser, S.N., Hackett, T.A., Ramachandran, R., Liberman, M.C., 2017. Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta). Hear. Res. 353, 213–223.
Van Hengel, P.W.J., Duifhuis, H., Van den Raadt, M.P.M.G., 1996. Spatial periodicity in the cochlea: the result of interaction of spontaneous emissions? J. Acoust. Soc. Am. 99 (6), 3566–3571.
Vannucci, G., Teich, M.C., 1978. Effects of rate variation on the counting statistics of dead-time-modified Poisson processes. Optic Commun. 25 (2), 267–272.
Verhulst, S., 2010. Characterizing and Modeling Dynamic Processes in the Cochlea Using Otoacoustic Emissions. Ph.D. thesis. Technical University of Denmark, Kongens Lyngby, Denmark.
Verhulst, S., Harte, J.M., Dau, T., 2011. Temporal suppression of the click-evoked otoacoustic emission level-curve. J. Acoust. Soc. Am. 129 (3), 1452–1463.
Verhulst, S., Dau, T., Shera, C.A., 2012. Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission. J. Acoust. Soc. Am. 132 (6), 3842–3848.
Verhulst, S., Bharadwaj, H.M., Mehraei, G., Shera, C.A., Shinn-Cunningham, B.G., 2015. Functional modeling of the human auditory brainstem response to broadband stimulation a. J. Acoust. Soc. Am. 138 (3), 1637–1659.
Verhulst, S., Jagadeesh, A., Mauermann, M., Ernst, F., 2016. Individual differences in auditory brainstem response wave characteristics: relations to different aspects of peripheral hearing loss. Trends Hear. 20, 2331216516672186.
von Békésy, G., 1970. Travelling waves as frequency analysers in the cochlea. Nature 225 (5239), 1207–1209.
Westerman, L.A., Smith, R.L., 1984. Rapid and short-term adaptation in auditory nerve responses. Hear. Res. 15 (3), 249–260.
Westerman, L.A., Smith, R.L., 1988. A diffusion model of the transient response of the cochlear inner hair cell synapse. J. Acoust. Soc. Am. 83 (6), 2266–2276.
Winter, I.M., Palmer, A.R., 1991. Intensity coding in low-frequency auditory-nerve fibers of the guinea pig. J. Acoust. Soc. Am. 90 (4), 1958–1967.
Zagaeski, M., Cody, A.R., Russell, I.J., Mountain, D.C., 1994. Transfer characteristic of the inner hair cell synapse: steady-state analysis. J. Acoust. Soc. Am. 95 (6), 3430–3434.
Zeddies, D.G., Siegel, J.H., 2004. A biophysical model of an inner hair cell. J. Acoust. Soc. Am. 116 (1), 426–441.
Zhang, X., Carney, L.H., 2005. Analysis of models for the synapse between the inner hair cell and the auditory nerve. J. Acoust. Soc. Am. 118 (3), 1540–1553.
Zhang, X., Heinz, M.G., Bruce, I.C., Carney, L.H., 2001. A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. J. Acoust. Soc. Am. 109 (2), 648–670.
Zhong, Z., Henry, K.S., Heinz, M.G., 2014. Sensorineural hearing loss amplifies neural coding of envelope information in the central auditory system of chinchillas. Hear. Res. 309, 55–62.
Zilany, M.S., Bruce, I.C., 2006. Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery. J. Acoust. Soc. Am. 120 (3), 1446–1466.
Zilany, M.S., Bruce, I.C., Nelson, P.C., Carney, L.H., 2009. A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. J. Acoust. Soc. Am. 126 (5), 2390–2412.
Zilany, M.S., Bruce, I.C., Carney, L.H., 2014. Updated parameters and expanded simulation options for a model of the auditory periphery. J. Acoust. Soc. Am. 135 (1), 283–286.
Zweig, G., 1976. Basilar membrane motion. In: Cold Spring Harbor Symposia on Quantitative Biology, vol. 40. Cold Spring Harbor Laboratory Press, pp. 619–633.
Zweig, G., 1991. Finding the impedance of the organ of Corti. J. Acoust. Soc. Am. 89 (3), 1229–1254.
Zweig, G., 2016. Nonlinear cochlear mechanics. J. Acoust. Soc. Am. 139 (5), 2561–2578.