Calibration and reference simulations for the auditory periphery model of Verhulst et al. 2018 version 1.2

Alejandro Osses Vecchi and Sarah Verhulst*

Hearing Technology Lab @ WAVES,

Department of Information Technology, Ghent University

Abbreviations			
ABR	Auditory Brainstem Response	IC	Inferior Colliculus
AM	Amplitude Modulation	IIR	Infinite Impulse Response
AN	Auditory Nerve	MTF	Modulation Transfer Function
CF	Characteristic Frequency	NH	Normal hearing
CN	Cochlear Nucleus	HI	Hearing impaired
EFR	Envelope Following Response	pe SPL	peak-equivalent Sound Pressure Level
FFT	Fast Fourier Transform	W-I,-III,-V	Waves I, III, V

The code of the auditory periphery model can be retrieved from: https://github.com/HearingTechnology/Verhulstetal2018Model/. The current model version can be cited as follows: "We adopted the model of the human auditory periphery described by Verhulst, Altoè, and Vasilkov (2018) in the v1.2 implementation (Osses Vecchi & Verhulst, 2019)."

Abstract

This document describes a comprehensive procedure of how the biophysical model published by Verhulst et al. (2018) [1] can be calibrated on the basis of reference auditory brainstem responses (ABRs). Additionally, the filter design used in two of the model stages, cochlear nucleus (CN) and inferior colliculus (IC), is described in detail. These descriptions are valid for a new release of the Verhulst et al. model, version 1.2, as well as for previous versions of the model (version 1.1 or earlier), whose block diagram is shown in Fig. 1 [see also, 1, 2, their Fig. 1]. The differences between the model versions are explicitly mentioned and simulations to basic auditory stimuli are shown for model versions 1.1 and 1.2. Version 1.2 of the model includes a new implementation of the CN and IC stages (Stages 5 and 6). All previous model stages (Stages 1-4 in Fig. 1: outer and middle ear, transmission-line cochlear filter bank, inner hair cell model, and auditory nerve model) remained unchanged.

In the new release (model version 1.2), in addition to the updated CN and IC stages, we employed a different calibration procedure to match human reference ABR amplitudes of waves I, III, and V more faithfully. This release note shows the implications of these model adjustments on the simulations presented in the original 2018 model paper. For this purpose, results from two model versions are reported:

- New model release (version 1.2), labelled as **'model v1.2'** (GitHub commit f59309b).

- Previous model release (version 1.1) as used in [1], labelled as **'model v1.1'** (GitHub commit 9a7a7ef).

The main difference between IC model stages relates to the degree of IC inhibition that was applied, with more inhibition in **v1.2** than implemented in **v1.1**. The time domain simulations presented in this document show that this change in inhibition strength does not drastically change the results presented in the original paper. However, **v1.2** more correctly captures the physiologically derived CN and IC inhibition/excitation strengths [3].

1 Model description

The model of the auditory periphery, whose framework is described in [1, 4, 5], offers a quantitative description of different functional components of the hearing system, providing simulated neural representations in the ascending auditory pathway up to the inferior colliculus. Its functioning was evaluated on the basis of human otoacoustic emissions [5] and auditory evoked potentials [1, 4]. The block diagram of the model is shown in Fig. 1. The first processing stage of the model is the outer- and middle-ear models. The outer ear model [6] simulates the presentation of sounds through circumaural headphones.

---

*Contact: Alejandro Osses Vecchi (alejandro.osses@ugent.be), prof. Sarah Verhulst (s.verhulst@ugent.be).

If the outer ear is omitted, then the sounds are assumed to be presented via in-ear earphones. After the outer- and middle-ear filters (Stage 1), the signal is fed into a transmission-line cochlear filter bank (Stage 2) where basilar membrane vibrations are simulated at 401 equidistant cochlear locations. A non-linear compression is applied to the cochlear vibrations, followed by a conversion into inner-hair-cell receptor potentials (Stage 3). These potentials are used to generate spike-rate patterns at the level of the auditory nerve (AN, Stage 4), which are related to fibres that have high- (HSR, 70 spikes/s), medium- (MSR, 10 spikes/s), and low-spontaneous rates (LSR, 1 spikes/s). These patterns are finally combined (13 HSR, 3 MSR, and 3 LSR fibres, i.e., 13-3-3) to form the input to a model of the spherical bushy cell in the cochlear nucleus (Stage 5) and inferior colliculus (Stage 6). The information across simulated characteristic frequencies available at the outputs of Stages 4, 5, and 6 offer a correlate of human brainstem responses recorded from scalp electrodes. The "central processor stage" (Stage 7) represents a way to integrate the simulated information in the ascending auditory pathway that is assumed to represent the way in which human listeners code and compare sounds among each other [2, 7].

Fig. 1: Block diagram of the biophysical model of the auditory periphery by Verhulst et al. [1, 2].

Model update in v1.2

In **model v1.2** two changes were introduced to the CN/IC model description (script ic_cn2018.py, from the GitHub repository). One of the changes (correction of a bug) is relevant and affects the IC stage implementation while the other is a minor change that was introduced to the lowpass filter design of alpha functions (**v1.2** using Eq. (6), **v1.1** using Eq. (7), see Section 2).

A bug in the **v1.1** IC implementation resulted in a dominance of the excitatory neural responses while the effect of the inhibitory responses was nearly marginal. This situation was a consequence of using $1/\tau_{\text{exc}}^4$ and $1/\tau_{\text{inh}}^4$ as scaling constants for the excitatory and inhibitory signals, respectively, instead of $1/\tau_{\text{exc}}^2$ and $1/\tau_{\text{inh}}^2$ as shown later in Eq. (2). The excitation-dominated IC outputs were due to the much larger $1/\tau_{\text{exc}}^4$ value than $1/\tau_{\text{inh}}^4$.

As a consequence of correcting this bug, the original M5 scaling factor (referred to as $A_{\text{W-V}}$ in the model paper [1]) had to be updated to maintain human-like ABR wave-V amplitudes. Because we needed to re-calibrate, we decided to adopt a more accurate procedure to calibrate the model constants to match the normative data by Picton [8] than originally presented. This resulted in updated M1 ($A_{\text{W-I}}$ in the paper) and M3 ($A_{\text{W-III}}$ in the paper) scaling factors as well. The new calibration procedure is explained in Section 3.1 and the obtained scaling factors are given in Table 1.

Table 1: Scaling constants to yield ABR Waves I, III, and V to a realistic peak amplitude before (v1.1) and after (v1.2) correcting the bug and adopting a new calibration procedure. These constants are specified in the Python script ic_cn2018.py, lines 5-7.

| Constant label | | Value (dimensionless) | | ABR amplitudes from | |
Paper	Model	Model v1.2 (New)	Model v1.1	Model v1.2	Model v1.1
$A_{\text{W-I}}$	M1	$4.2767 \cdot 10^{-14}$	$6.2755 \cdot 10^{-14}$	$0.15\ [\mu V_p]$*	$0.30\ [\mu V_{pp}]$
$A_{\text{W-III}}$	M3	$5.1435 \cdot 10^{-14}$	$7.2161 \cdot 10^{-14}$	$0.17\ [\mu V_p]$*	$0.34\ [\mu V_{pp}]$
$A_{\text{W-V}}$	M5	$13.3093 \cdot 10^{-14}$	$3.5200 \cdot 10^{-20}$	$0.61\ [\mu V_{pp}]$	$0.61\ [\mu V_{pp}]$

* Picton's normative data [8] are reported to be 0.30, 0.34, 0.61 [μV] peak-to-peak for waves I, III, and V, respectively. We assumed that the corresponding baseline-to-peak voltages are approximately half of these values.

2 CN and IC model implementation within model v1.2

2.1 Analytical formulation
(Same formulation as in model v1.1 [1])

As stated in [3], ventral CN postsynaptic potentials can be approximated by using alpha functions of the form $P(t) = t \cdot \exp(-t/\tau)$. In our model these functions are of the form $P(t) = t/\tau^2 \cdot \exp(-t/\tau)$. The area under the impulse response of the alpha functions is equal to 1. Given that alpha functions act as lowpass filters in the frequency domain, unit energy in the time domain corresponds to a 0-dB gain in the passband of the filter.

The originally described analytical CN model equation is [1, Eq. 17] (refer to the paper for a detailed description of each of the equation parameters):

$$r_{\mathrm{CN}}(t) = A_{\mathrm{CN}} \cdot \left[\frac{t}{\tau_{\mathrm{exc}}^2} \cdot \exp\left(-\frac{t}{\tau_{\mathrm{exc}}} \right) * r_{\mathrm{AN}}(t) - S_{\mathrm{CN,inh}} \cdot \frac{t}{\tau_{\mathrm{inh}}^2} \cdot \exp\left(-\frac{t}{\tau_{\mathrm{inh}}} \right) * r_{\mathrm{AN}}(t - D_{\mathrm{CN}}) \right] \quad (1)$$

The IC model equation is:

$$r_{\mathrm{IC}}(t) = A_{\mathrm{IC}} \cdot \left[\frac{t}{\tau_{\mathrm{exc}}^2} \cdot \exp\left(-\frac{t}{\tau_{\mathrm{exc}}} \right) * r_{\mathrm{IC}}(t) - S_{\mathrm{IC,inh}} \cdot \frac{t}{\tau_{\mathrm{inh}}^2} \cdot \exp\left(-\frac{t}{\tau_{\mathrm{inh}}} \right) * r_{\mathrm{IC}}(t - D_{\mathrm{IC}}) \right] \quad (2)$$

Next, we provide a detailed analysis of how the function $P(t)$ can be implemented in the discrete domain.

2.2 Conversion to discrete domain
(Minor change with respect to model v1.1)

In our model (both, v1.1 and v1.2) alpha functions are implemented as IIR filters in the digital domain, with filter coefficients which are obtained from the corresponding z-transfer function $H(z^{-1})$. The continuous time domain function is $P(t) = t/\tau^2 \cdot \exp(-t/\tau)$. The frequency-domain implementation, from which $H(z^{-1})$ can be obtained, is given by the Laplace transform $H(s)$. For $P(t) = K \cdot t^n \cdot \exp(a \cdot t)$, with constants K (here $K = 1/\tau^2$), a (here $a = -1/\tau$), and n (here $n=1$):

$$H(s) = \frac{K \cdot n!}{(s-a)^{n+1}} \qquad \Longrightarrow \qquad H(s) = \frac{K}{(s-a)^2} \quad (3)$$

where $s = j\omega$ and ω is the natural frequency. To convert the Laplace transform into digital domain, a bilinear transformation can be applied (replacement of each s by $\frac{2}{T} \cdot \left(\frac{1-z^{-1}}{1+z^{-1}} \right)$, with $T = 1/f_s$, [e.g., 9, Chapter 7]). The transfer function of the alpha function in the discrete domain thus becomes:

$$H(z^{-1}) = \frac{K}{\left(\frac{2}{T} \cdot \left(\frac{1-z^{-1}}{1+z^{-1}} \right) - a \right)^2} \quad (4)$$

The coefficients of the corresponding IIR filter can be obtained by developing this expression:

$$H(z^{-1}) = \frac{K \cdot T^2}{4} \cdot \frac{1}{(1 - T \cdot a/2)^2} \cdot \frac{1 + 2z^{-1} + z^{-2}}{1 - 2 \cdot \left(\frac{1+T \cdot a/2}{1-T \cdot a/2} \right) z^{-1} + \left(\frac{1+T \cdot a/2}{1-T \cdot a/2} \right)^2 z^{-2}}$$

Replacing K by $1/\tau^2$, T by $1/f_s$ and a by $-1/\tau$ and reordering yields:

$$H(z^{-1}) = \frac{1}{(2 \cdot f_s \cdot \tau)^2} \cdot \frac{(2 \cdot f_s \cdot \tau)^2}{(2 \cdot f_s \cdot \tau + 1)^2} \cdot \frac{1 + 2z^{-1} + z^{-2}}{1 - 2 \cdot \left(\frac{2 \cdot f_s \cdot \tau - 1}{2 \cdot f_s \cdot \tau + 1} \right) z^{-1} + \left(\frac{2 \cdot f_s \cdot \tau - 1}{2 \cdot f_s \cdot \tau + 1} \right)^2 z^{-2}} \quad (5)$$

The coloured terms are scaling factors. The factor highlighted in red was not considered in the CN/IC implementation in **v1.1**. This factor is, however, almost unity. For instance, if $f_s = 20$ kHz and $\tau = 2$ ms, then $(2 \cdot f_s \cdot \tau)^2 / (2 \cdot f_s \cdot \tau + 1)^2 = 0.9755$, or -0.22 dB.

Hence, the transfer function of the alpha function as implemented in **model v1.2** is (after simplification):

$$H(z^{-1}) = \frac{1}{(2 \cdot f_s \cdot \tau + 1)^2} \cdot \frac{1 + 2z^{-1} + z^{-2}}{1 - 2 \cdot \left(\frac{2 \cdot f_s \cdot \tau - 1}{2 \cdot f_s \cdot \tau + 1} \right) z^{-1} + \left(\frac{2 \cdot f_s \cdot \tau - 1}{2 \cdot f_s \cdot \tau + 1} \right)^2 z^{-2}} \quad (6)$$

In comparison, the implementation in **model v1.1** was:

$$H(z^{-1}) = \frac{1}{(2 \cdot f_s \cdot \tau)^2} \cdot \frac{1 + 2z^{-1} + z^{-2}}{1 - 2 \cdot \left(\frac{2 \cdot f_s \cdot \tau - 1}{2 \cdot f_s \cdot \tau + 1} \right) z^{-1} + \left(\frac{2 \cdot f_s \cdot \tau - 1}{2 \cdot f_s \cdot \tau + 1} \right)^2 z^{-2}} \quad (7)$$

The coefficients of the transfer function (and of the IIR filter) for model versions **v1.2** and **v1.1** are shown in Table 2 and correspond to the constants in the numerator and denominator of Eqs. 6 and 7, respectively. For comparison, the IIR filter coefficients from the CN/IC implementation within the UR EAR toolbox v2.1[1] (IC/CN model described in [3]) are also shown in the table.

Table 2: Coefficients of the transfer function (IIR filter) used within the CN and IC models.

Model	Numerator			Denominator				Scaling	Passband
version	b_0	b_1	b_2	a_0	a_1	a_2	m	factor C	gain G [dB]
v1.2 (Eq. 6)	1	2	1	1	$-2 \cdot m$	m^2	$\frac{2 \cdot f_s \cdot \tau - 1}{2 \cdot f_s \cdot \tau + 1}$	$\frac{1}{(2 \cdot f_s \cdot \tau + 1)^2}$	0
v1.1 and earlier (Eq. 7)	1	2	1	1	$-2 \cdot m$	m^2	$\frac{2 \cdot f_s \cdot \tau - 1}{2 \cdot f_s \cdot \tau + 1}$	$\frac{1}{(2 \cdot f_s \cdot \tau)^2}$	0.22 ($\tau = 2$ ms)* 0.85 ($\tau = 0.5$ ms)*
UR EAR	0	m	0	1	$-2 \cdot m$	m^2	$\exp\left(-\frac{1}{f_s \cdot \tau}\right)$	$\frac{1}{f_s^2 \cdot \tau^2 \cdot [1 - (\tau + 1) \cdot \exp(-1/\tau)]}$	0

*Gains assessed using the default sampling frequency $f_s = 20$ kHz that is used for simulated neural responses.

2.3 Modulation transfer function (MTF) for AM tones

To illustrate how the modifications to the CN/IC model description affect the filter tuning of the CN/IC models (**model v1.2**, **model v1.1**) the MTF for 4-kHz, 100% modulated amplitude-modulated (AM) tones, was assessed. Our implementation is compared with the CN/IC implementation from Laurel Carney's group available within the UR EAR toolbox, which was used to process the AN responses obtained from our model. This model variant is labelled as **model2018+UR EAR**.

Stimuli: AM tones with carrier frequency $f_c = 4$ kHz, duration of 200 ms, level of 70 dB SPL, gated with 5-ms up-down ramps were used. The tones were 100% modulated with modulation rates of f_{mod} between 5 and 250 Hz in steps of 5 Hz.

MTF assessment: On-frequency MTFs for CN and IC outputs in channel 112 (i.e., at CF= 4013 Hz) were assessed. The MTFs were obtained by reading the maximum amplitude of the corresponding 4000-point FFT, after DC removal. The DC removal was needed to obtain a maximum in the FFT response elicited by the sensitivity to stimulus modulations and not by the DC.

Results: The obtained MTFs are shown in Fig. 2 for the different model implementations. The obtained on-frequency MTFs are shown in two forms: (1) As relative modulation strength (left panels), i.e., as simulated (maximum) voltages, normalised to the maximum modulation strength (at around f_{mod}= 100 Hz);
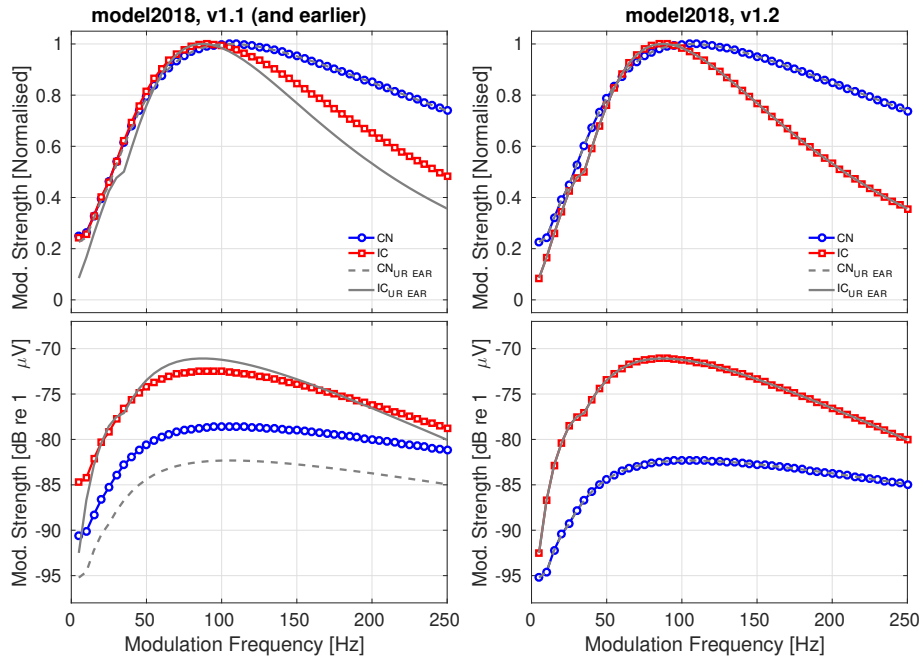
Fig. 2: MTFs using model **v1.1** (left panels) and **v1.2** (right panels). The CN and IC outputs using model2018+UR EAR are indicated by the dashed and continuous grey lines, respectively.

(2) As maximum voltages in dB referenced to 1 μV. The bandpass characteristic of the CN and IC models is similar in the three evaluated models (model **v1.1**, **v1.2**, and **UR EAR**). The bandpass sensitivity (or modulation strength) is, however, slightly sharper for **v1.2** and **UR EAR** in comparison with **v1.1**. The MTFs between model **v1.2** and **model2018+UR EAR** show a sensitivity in the same magnitude range (superimposed grey and coloured curves in right panels).

3 Model validation

To validate **model v1.2**, we simulated responses to the following stimuli:

- Clicks of alternating polarity (for calibration of scaling factors in **v1.2**)
- Clicks of positive polarity to reproduce Fig. 6 in [1]
- AM tones with carrier frequency of 4-kHz and 85% modulation depth to reproduce Fig. 7 in [1]

All model simulations were performed in MATLAB (tested in various MATLAB releases between R2012b and R2018a, on Windows and Unix-based machines) and Python (v2.7) using the following configuration:

Listing 1: Configuration in model versions **v1.1** and **v1.2**. The basilar-membrane pole admittances used as input to the model (sheraP) corresponded to a normal hearing cochlear gain profile ("Flat00") unless otherwise stated.

```
1  flag_cf='abr'; % 401 CF channels
2  flag_proc='cfvihmlbw';
3  non_linear_type='vel';
4  numH=13; numM=3; numL=3; % 19 neurones
5  dir_out=[pwd(),filesep]; % directory were variables will be stored
6  output = model2018(insig,fs,flag_cf,1,flag_proc,1,sheraP,0.05,
       non_linear_type,numH,numM,numL,1,dir_out);
```

*See also Example_Simulation.m from the GitHub repository.

3.1 Clicks: Calibration of the scaling factors

Calibration stimuli: A click train with repetition rate of 11.1 Hz and duration of 6 s (e.g., 66 clicks) was generated. The clicks had alternating polarity (starting with positive) and each click had a duration of 80μs. The clicks with positive and negative polarities had amplitudes of A and $-A$, respectively, which were matched to have a level of 100 dB peSPL, that is, having a baseline-to-peak amplitude equal to the peak-to-peak amplitude of a sinusoid of 100 dB SPL. The resulting clicks had an amplitude A of 5.66 [Pa].

Simulated of ABR amplitudes: As shown in Fig. 3, broadband W-I and W-III are mainly excitatory responses while the simulated W-V has a combined excitatory/inhibitory pattern (with both positive and negative voltage deflections). To compensate for the lack of inhibition in simulated W-I and W-III, the peak-to-peak W-I and W-III amplitudes are assumed to be double as large as the baseline-to-peak amplitudes and this implies that the through-voltage is assumed to have the same amplitude (with opposite sign) than the corresponding peak value. For this reason, **model 2018 v1.2** is suitable to simulate baseline-to-peak W-I and W-III voltages and a peak-to-peak W-V voltage. **In order to express all simulated amplitudes that are obtained with model v1.2 in peak-to-peak values, W-I and W-III [V_p] have to be multiplied by a factor of 2.** This scaling factor was not needed for estimated amplitudes using **model v1.1** because in that model version all W-I, W-III, and W-V showed excitatory responses such that baseline-to-peak and peak-to-peak amplitudes corresponded with each other.

Calibration procedure: The scaling factors M1, M3, and M5 were adjusted to provide calibrated amplitudes of simulated waveforms (wave-I, -III, and -V, respectively) in volts (V) according to the normative data published in [8, Table 8–1] for clicks at a level of 70 dB nHL (the approximation used in the calibration procedure is that 100 dB peSPL≈70 dB nHL). In addition to the multiplication by a factor of 2 to estimate V_{pp} values for W-I and W-III when using **model v1.2** with respect to **v1.1**, the calibration procedure in **v1.2** is based on later clicks along the click train (illustrated here using clicks #59,60), while previous model versions were calibrated using the response of the first click only (click #1).

Results: The model response to the click train using the obtained scaling factors (Table 1) is illustrated in Fig. 3 for clicks #1, 59, and 60. Clicks #59 and 60 have average W-I, W-III, and W-V amplitudes of 0.15 [μV_p], 0.17 [μV_p], and 0.61 [μV_{pp}], which represent estimated peak-to-peak values (two times the simulated W-I and W-III amplitudes) of 0.30, 0.34, and 0.61 [μV_{pp}].
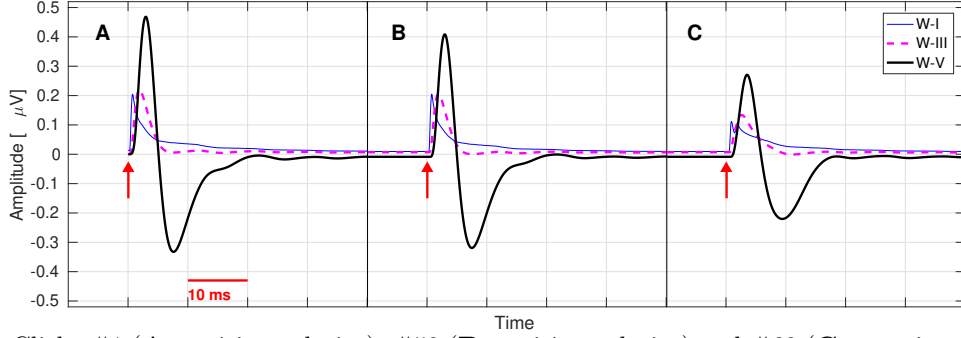
Fig. 3: Clicks #1 (**A**, positive polarity), #59 (**B**, positive polarity) and #60 (**C**, negative polarity) with an interstimulus time of about 90 ms (11.1 Hz presentation rate, not schematised in this figure). Clicks #59 and 60 (used for the model calibration) have baseline-to-peak amplitudes of 0.20 and 0.10μV for W-I (average of 0.15μV, in blue), 0.21 and 0.13μV for W-III (average of 0.17μV, in magenta), and peak-to-peak amplitudes of 0.73 and 0.49μV for W-V (average of 0.61μV, in black), respectively. As a consequence of the calibration process, the average peak amplitudes of the clicks match the amplitudes of the normative data [8] (see Table 1). The red arrows indicate the click onset time.

3.2 Clicks: ABR wave-I and wave-V characteristics

Evaluation stimuli: A click train with a repetition rate of 20 Hz and duration of 0.5 s (10 clicks) was generated [10]. The clicks had a positive polarity and each click had a duration of 80 μs. The first click was set to start after 10 μs. The level of the click was varied between 60 and 100 dB peSPL.

Hearing profiles: Three hearing profiles were considered: NH (Flat00, 13-3-3), HI (Slope35, 13-3-3), and HSR (Flat00, 13-0-0). (1) The **NH** profile uses a normal-hearing cochlear model with no synaptopathy. (2) The **HI** profile simulates a high-frequency sloping audiogram with 0 dB HL at 1 kHz and 35 dB HL at 8 kHz and no synaptopathy. (3) the **HSR** profile uses a normal hearing cochlear model and simulates a loss of medium- and low-spontaneous rate neurones. A graphical representation of the cochlear-gain-loss profiles Flat00 and Slope35 can be found in [11].

Results: The ABRs presented in this section were used to reproduce the simulation results shown in Fig. 6 of the original model paper, where ABR latencies and amplitudes derived from a single click were reported ("Click #1"). For ease of comparison with the current results, the panels in Fig. 4 are labelled as A, B, and C (here from top to bottom), respectively. In Fig. 4, the left-most panels (column 1) reproduce the ABR results from the original model paper (using **model v1.1**), while the middle panels (column 2) show the ABRs to Click #1 but then using the **model v1.2** implementation. However, given that the original aim of the model was to mimic experimental findings to reference ABRs recorded to click trains, the model paper should have rather used click trains in its evaluation/calibration than using Click #1. The experimental data labelled in the model paper as 'D03' (reported in [12] but originally collected in [10]) used a repetition rate of 20 Hz. To provide better experimental consistency, new simulation results were generated with ABRs derived from the 10th click in the train using model v1.1 and v1.2 in panels 3 and 4, respectively. This new approach ensures that all AN time constants have reached steady-state as can be expected experimentally. The ABRs in the right-most panels (column 4) should hence be more directly comparable to the experimental results in [12] and other experimental references reported in the model paper.

Discussion: Simulations with Click #1: Shorter latencies and larger W-I and W-V amplitudes were obtained for **model v1.2** compared to **model v1.1**. This was an expected consequence of the adopted calibration procedure, which was based on later clicks rather than to the first click of the click train, which was not affected by AN adaptation. Due to the larger W-I and W-V amplitudes in **v1.1**, the derived latencies were shorter in **v1.2** than originally, and this effect was more pronounced in the HI profile (purple line) and almost negligible in the NH (green lines) and HSR profiles (red dashed lines).

Discussion: Simulations with Click #10: For the same model version, i.e., panel 3 compared to panel 1 (**model v1.1**), and panel 4 with panel 2 (**model v1.2**), the obtained ABR amplitudes (see Fig. 4C) using Click 10th were lower, which was an expected result. The latencies were also shorter, especially for the lower level clicks (60 and 70 dB peSPL).

Discussion: Simulations with Click #1 (model v1.1, panel 1) and Click #10 (model 1.2, panel 4): Both latencies and amplitudes had a similar range. The results with Click #10 (panel 4) reflect more closely the experimental data [10] which were collected to click trains with a repetition rate of 20 Hz. Note that the peak amplitudes at 100 dB peSPL in panel 1 coincide with the normative data (0.61μV for W-V; 0.30μV for W-I), whereas the amplitudes are slightly higher in panel 4. This is a consequence of the lower repetition rate of the clicks (20 Hz) and, hence, the longer neuronal firing recovery compared to a click presentation at 11.1 Hz, which was used in the model calibration.
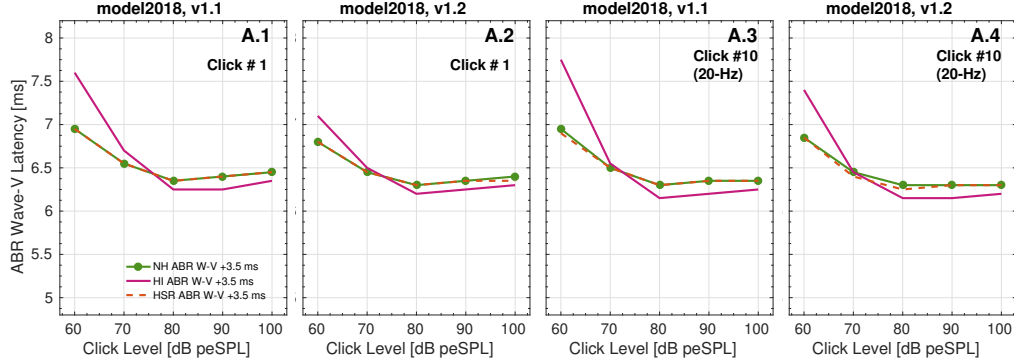
(a) **Simulated ABR W-V latencies** as a function of click level for models using a normal-hearing profile (NH: Flat00, 13-3-3, green line), a sloping high-frequency hearing-loss profile (HI: Slope35, 13-3-3, purple line), and a synaptopathy profile has contributions from only HSR neurones (HSR: Flat00, 13-0-0, red dashed line). The simulated latencies were incremented by 3.5 ms to compare the simulations with the corresponding reference data (which are not shown here), as it was also performed in the model paper. These panels are similar to Fig.6A from [1].
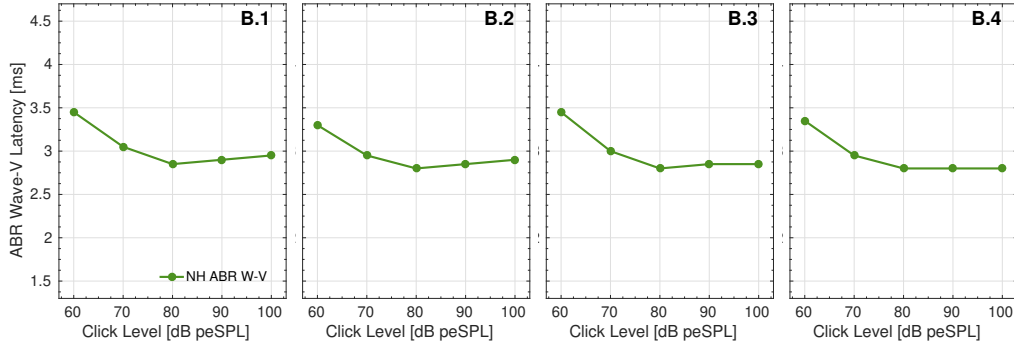
(b) Same simulated latencies as reported in panel A, but here only the results for the NH profile are shown, without the 3.5-ms increment. These panels are similar to Fig.6B from [1].
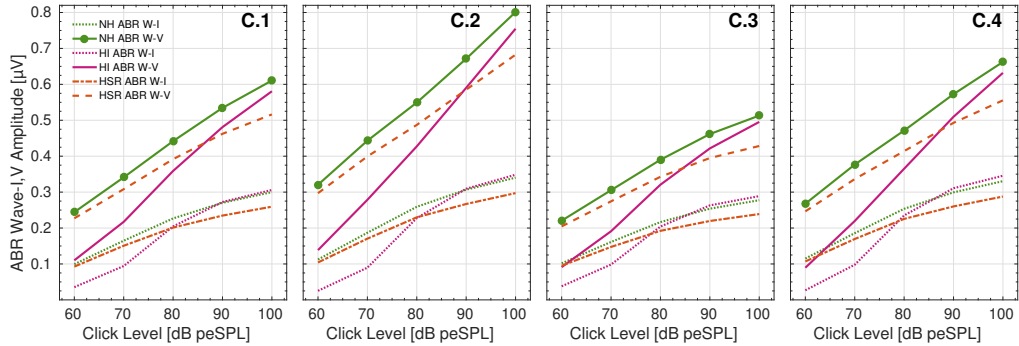
(c) **Simulated ABR W-I and W-V amplitudes** as a function of click levels for models using the NH profile (green lines), the HI profile (purple lines), and a selective synaptopathy profile HSR (red dashed lines). In general, cochlear gain loss (HI profile) steepened the ABR amplitude growth curve and synaptopathy (HSR profile) resulted in shallower curve slopes compared to the NH profile. These panels are similar to Fig. 6C from [1].

Fig. 4: Simulations using **model v1.1** (panel columns 1 and 3) and **model v1.2** (panel columns 2 and 4). The plots in columns 1 and 2 were obtained from the responses to the first click in the click train (Click #1), while the plots in columns 3 and 4 were obtained from the responses to the 10th click in the click train.

3.3 AM tones: Simulating EFRs

Stimuli: A 100-ms pure tone with a carrier frequency of 4 kHz was modulated at a rate of 98 Hz using 85% modulation depth. An up-down cosine ramp of 2.5 ms was applied. The sound was preceded by 20 ms of silence before the stimulus was fed into the model.

EFR assessment: Envelope-following responses (EFRs) were obtained by adding the simulated AN responses, CN, and IC responses as follows:

$$r_{\text{EFR}}(t) = A_{\text{W-I}} \cdot r_{\text{AN}}(t) + A_{\text{W-III}} \cdot r_{\text{CN}}(t) + A_{\text{W-V}} \cdot r_{\text{IC}}(t) \tag{8}$$

Broadband population responses (Fig. 5, top panels) were obtained by adding the CF contributions from
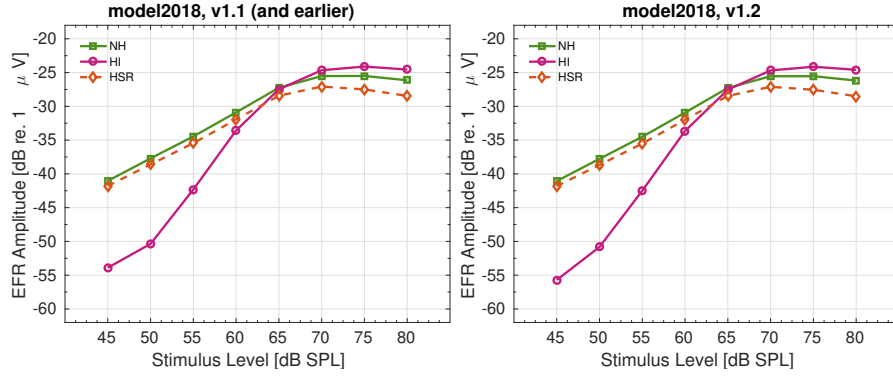
(a) **Simulated EFR amplitudes** obtained using Eq. (9) for a 4-kHz AM tone with f_{mod}=98 Hz (85% modulation depth) presented to the model at different levels. Three hearing profiles are shown: NH (green squares), HI (purple circles), and HSR (red dashed diamonds). These panels are similar to Fig. 7A from [1].
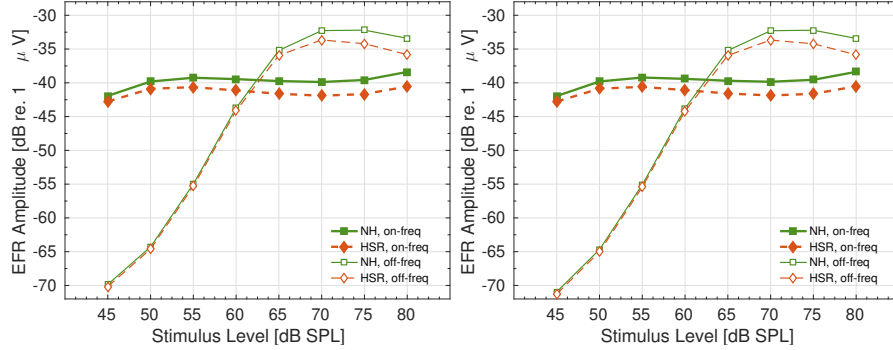
(b) **Simulated EFR amplitudes** for the same 4-kHz tones of panel (a) obtained with Eq. (9) but using r_{EFR} (Eq. 8) with only on-frequency (filled markers) or off-frequency contributions (open markers) for the NH (green squares) and HSR profiles (red dashed diamonds). The on-frequency and off-frequency contributions were obtained from the CF channels around 4 kHz (channels 100 to 123) and 8 kHz (channels 30 to 54), respectively. See the text for further details. These panels are similar to Fig. 7B from [1].
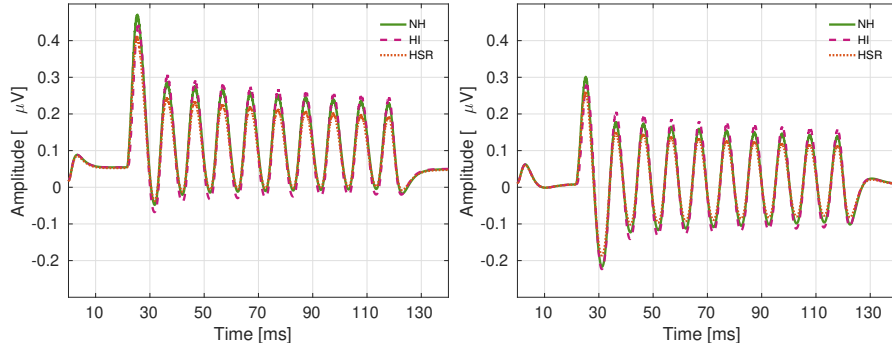
(c) Simulated r_{EFR} (Eq. 8) in time domain using the 70-dB tone.

Fig. 5: Simulations using **model v1.1** (left panels) and **model v1.2** (right panels).

all simulated frequency bins (401 channels), between 112 Hz (channel 401) and 12 kHz (channel 1). On-frequency and off-frequency EFRs (Fig. 5, middle panels) were assessed by summing up the CF contributions within one-third octave-band around 4 kHz (channel 112, channels 100 to 123) and 8 kHz (channel 42, channels 30 to 54), respectively. The simulated population responses (4000-samples long: 200 ms at $f_{s,ABR} = 20$ kHz) were converted to the frequency domain by applying a 4000-point FFT (Fig. 5, top and middle panels) and were expressed in dB re. 1 μV:

$$\text{EFR}_{\text{magnitude}} = 20 \cdot \log_{10}\left(\frac{\max\left(\|\text{FFT}(r_{\text{EFR}})\|\right)/N}{1 \cdot 10^{-6}}\right) \qquad [\text{dB re. 1 } \mu\text{V}] \qquad (9)$$

Results: The corresponding model responses are shown in Fig. 5.

Discussion: The results shown in panels (a) and (b) of Fig. 5 are similar for both model versions. However, the time domain representation of the 70-dB envelope-following response of panel (c) clearly shows more inhibition (negative voltages) for **model v1.2**. This is a consequence of the corrected inhibition/excitation balance of the IC model stage.

References

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

[2] A. Osses Vecchi, F. Ernst, and S. Verhulst. "Hearing-impaired sound perception: What can we learn from a biophysical model of the human auditory periphery?" In: *International Congress on Acoustics*. Aachen, 2019, pp. 678–685.

[3] P. Nelson and L. Carney. "A phenomenological model of peripheral and central neural responses to amplitude-modulated tones". *J. Acoust. Soc. Am.* 116, (2004), 2173–2186.

[4] S. Verhulst, H. Bharadwaj, G. Mehraei, C. Shera, and B. Shinn-Cunningham. "Functional modeling of the human auditory brainstem response to broadband stimulation". *J. Acoust. Soc. Am.* 138, (2015), 1637–1659.

[5] S. Verhulst, T. Dau, and C. Shera. "Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission". *J. Acoust. Soc. Am.* 132, (2012), 3842–8.

[6] D. Pralong and S. Carlile. "The role of individualized headphone calibration for the generation of high fidelity virtual auditory space". *J. Acoust. Soc. Am.* 100, (1996), 3785–3793.

[7] S. Verhulst, F. Ernst, M. Garrett, and V. Vasilkov. "Supra-threshold psychoacoustics and envelope-following response relations: normal-hearing, synaptopathy and cochlear gain loss". *Acta Acust. united Ac.* 104, (2018), 800–803.

[8] T. Picton. "Auditory brainstem responses: peaks along the way". In: *Human auditory evoked potentials*. Plural Publishing, 2011. Chap. 8, pp. 213–245.

[9] A. Oppenheim, R. Schafer, and J. Buck. *Discrete-time signal processing*. 2nd ed. Prentice-Hall, Inc., 1999.

[10] T. Dau, O. Wegner, V. Mellert, and B. Kollmeier. "Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion". *J. Acoust. Soc. Am.* 107, (2000), 1530–1540.

[11] S. Verhulst, A. Jagadeesh, M. Mauermann, and F. Ernst. "Individual differences in auditory brainstem response wave characteristics: Relations to different aspects of peripheral hearing loss". *Trends in Hearing* 20, (2016), 1–20.

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