


A model framework for simulating spatial hearing of bilateral cochlear implant users

Hongmei Hu^{1,2,4,*} , Sebastián A. Ausili^{3,a} , Ben Williges^{4,a} , Jonas Klug^{1,2}, Rebecca C. Felsheim^{1,2} , Deborah Vickers⁴ , and Mathias Dietz^{1,2} 

¹Department of Medical Physics and Acoustics, University of Oldenburg, 26129 Oldenburg, Germany

²Cluster of Excellence “Hearing4all”, University of Oldenburg, 26129 Oldenburg, Germany

³Department of Otolaryngology, University of Miami, Miami, FL 33136, USA

⁴SOUND Lab, Cambridge Hearing Group, Clinical Neurosciences, University of Cambridge, CB20SZ Cambridge, UK

Received 28 February 2023, Accepted 18 July 2023

Abstract – Bilateral cochlear implants (CIs) greatly improve spatial hearing acuity for CI users, but substantial gaps still exist compared to normal-hearing listeners. For example, CI users have poorer localization skills, little or no binaural unmasking, and reduced spatial release from masking. Multiple factors have been identified that limit binaural hearing with CIs. These include degradation of cues due to the various sound processing stages, the viability of the electrode-neuron interface, impaired brainstem neurons, and deterioration in connectivity between different cortical layers. To help quantify the relative importance and inter-relationship between these factors, computer models can and arguably should be employed. While models exploring single stages are often in good agreement with selected experimental data, their combination often does not yield a comprehensive and accurate simulation of perception. Here, we combine information from CI sound processing with computational auditory model stages in a modular and open-source framework, resembling an artificial bilateral CI user. The main stages are (a) binaural signal generation with optional head-related impulse response filtering, (b) generic CI sound processing not restricted to a specific manufacturer, (c) electrode-to-neuron transmission, (d) binaural interaction, and (e) a decision model. The function and the outputs of different model stages are demonstrated with examples of localization experiments. However, the model framework is not tailored to a specific dataset. It offers a selection of sound coding strategies and allows for third-party model extensions or substitutions; thus, it is possible to employ the model for a wide range of binaural applications and even for educational purposes.

Keywords: Interaural time differences, Interaural level differences, Lateralization, Localization, Bilateral cochlear implant, Excitation-inhibition neuron

1 Introduction

Cochlear implants (CIs) have successfully restored the sensation of hearing to those with severe to profound deafness since the 1980s [1] and there are over 1 million implant recipients worldwide as of 2022 [2]. In the past 10–15 years, there have been improvements in advanced implant features and a large increase in the number of patients - (predominantly children) who receive bilateral CIs. These changes have led to significant improvements in CI outcomes due to the restoration of some binaural processing abilities, resulting in better speech understanding in noisy environments and better sound localization. For example, the accuracy of a bilateral CI user's localization in the

azimuthal plane is much improved over that of unilateral CI users, by exploiting interaural level differences (ILD) [3–5], and potentially making better use of envelope interaural time differences (ITD) [6, 7]. However, substantial differences still exist in many binaural tasks between bilateral CI users and normal hearing (NH) listeners [8–10]. Multiple factors have been identified that limit binaural hearing abilities of people with bilateral CIs, and the influence of these factors can depend on the individual. Factors or processing stages include the acoustic environment, the head-related impulse response (HRIR), various CI speech processing stages and their individual settings, electrode-neuron transmission, binaural interaction in brainstem neurons, as well as particularly complex and plastic cortical information decoding and decision-making stages (see, e.g., [11] for a review). These stages are mostly covered in the proposed framework. Figure 1 shows a schematic of

*Corresponding author: hongmei.hu@uni-oldenburg.de

^aThese authors contributed equally.

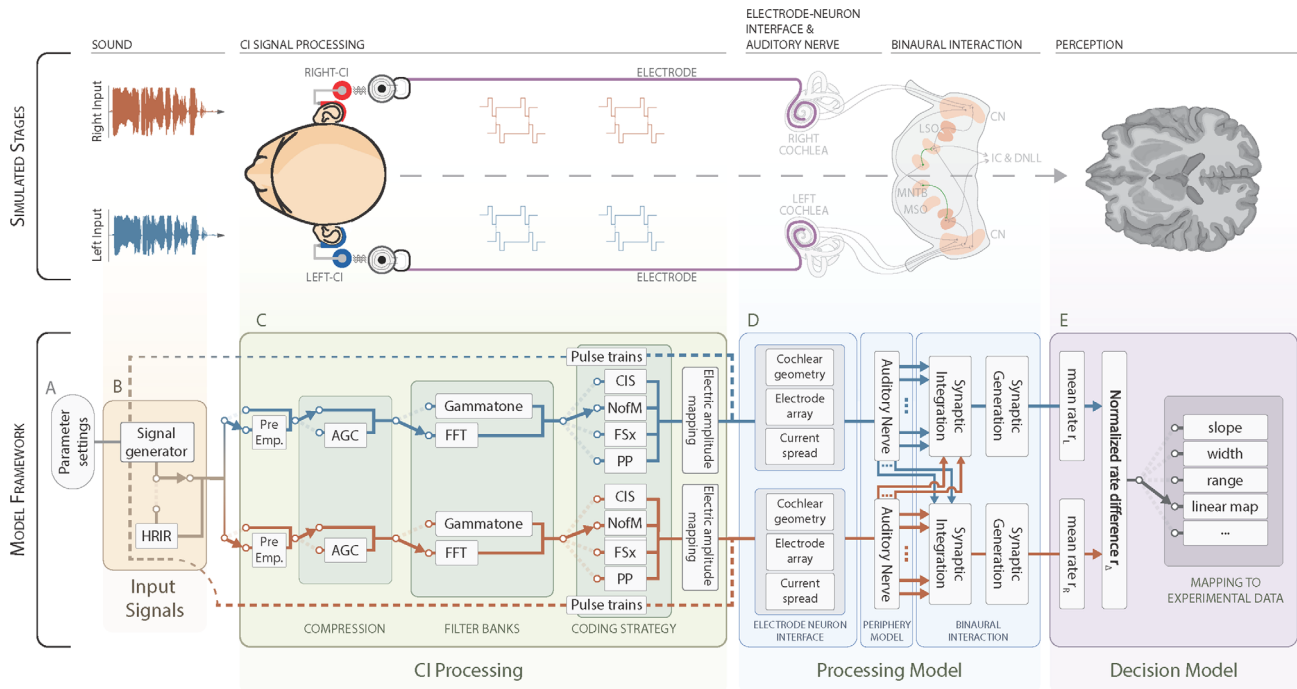


Figure 1. The upper panel of the figure (“Simulated Stages”) illustrates the “hardware components” of the different stages, and the lower panel (“Model Framework”) presents the corresponding modules along the processing chain in the current model framework. (A) The parameter settings encompass default values for the model, with provisions for user-defined settings. (B) The “Input Signals” stage incorporates functions for generating binaural signals, with or without head-related impulse responses (HRIRs). (C) The “CI Processing” stage comprises a general pre-emphasis high-pass filter, automatic gain controls (AGCs), filter banks, and coding strategies. Note that this block can be bypassed by providing pulse trains directly from the signal generator block. (D) The “Processing Model” includes an electrode-neuron interface (ENI), an auditory nerve (AN) periphery model, and an excitation-inhibition (EI) binaural interaction model. (E) The “Decision Model” maps the spiking rate to psychoacoustic behavioral data. Abbreviations: continuous interleaved sampling (CIS); N-of-M channel selection (NofM); with x -channel fine structure processing (FSx); peak picking (PP).

the model framework, which includes input signals, various CI processing stages, the electrode-neuron interface stage, the auditory nerve (AN) stage, the binaural integration stage, and simplified decision-making stages. Within such a long sound encoding/decoding pipeline, it is often not possible to identify the reason for a particular impairment or limitation. Computer models of auditory processing can be employed to help diagnose and understand the cause of a specific perceptual result or measurement by using one or a combination of several aforementioned stages.

Any simulation from the sound source to perception is comprised of many stage-specific existing models. For example, for the peripheral auditory system, a variety of different model types have been developed, such as biophysical [12–16] and phenomenological AN models [17–24], and 3D implanted cochlear models [25–31]. By combining these models with simulations of CI signal processing, it has enabled the investigation of broader research topics, such as simulating loudness perception [32, 33], sound source localization [34, 35], and speech intelligibility [36, 37].

Theoretically, a combination of verified stage-specific models in a single framework should allow for the simulation of individual CI user’s perception (Fig. 1). Yet, as is evident from the literature referenced above, models of each stage

still have some unresolved details, or their parametrization is already extremely complex even when used in isolation. It is unclear to what extent it is necessary to have very detailed parameter definitions to simulate auditory processing on a macroscopic level. These difficulties and limitations restrict the generation of multi-stage sound-to-perception type models. In fact, accurate simulations with most models are often limited to relatively narrow datasets. In the binaural context, simulating human perception with models built upon physiological data still remains an ongoing challenge even for understanding normal hearing (see, e.g., [38] for a review). However, some recently developed, fairly comprehensive models (e.g., [39–41]) have overcome some of the challenges and demonstrated the possibility for full perceptual models.

Previously, Kelvasa and Dietz [34] published a sound-to-perception model for bilateral CIs by combining a CI speech coding strategy mimicking the advanced combination encoder (ACE) strategy from Cochlear Corporation [42, 43]. They implemented several auditory processing models to estimate the direction of a sound source: First, an electrode-neuron interface, second an AN model, third a binaural interaction stage, and finally a decision stage. Like most of the above-mentioned models, this model was

able to account for selected localization data but was not generalizable to other experiments. In addition, there were some programming limitations, most notably, their Hodgkin and Huxley model stage for the lateral superior olive (LSO) is the computational bottleneck and is not provided in MATLAB. Therefore, it was deemed necessary to revise some of the model stages from Kelvasa and Dietz [34] to simulate a broader range of experiments. The first is the simulation of different CI stimulation techniques, including direct stimulation of single or multiple electrodes, utilizing audio line input, as well as free-field stimulation based on head-related impulse responses (HRIR) or even binaural room impulse responses (BRIR). A second extension allows for a much wider range of sound-coding strategies to be explored, incorporating options such as filter bank type, fixed pulse rate or local-peak-driven pulse rates, or N-of-M channel selection. The third extension replaced the Hodgkin and Huxley LSO model in the binaural integration stage with a functional and computationally more efficient model [44] as used in Klug et al. [40] and Hu et al. [45]. The fourth extension was to add a variety of decision stages that can be selected in a task-dependent manner.

In this work, the framework is demonstrated using simulations of free-field localization results of bilateral CI users, based on typical and well-established experimental findings previously conducted in both research (e.g., [46]) and clinical (e.g., [4, 5, 7, 9]) settings. The model outputs are shown at different stages along the processing chain, including HRIR-based virtual acoustic rendering, interaurally independent automatic gain controls (AGCs), and various coding strategies with different filter banks. These model simulations aim to correspond to the average CI user's performance, providing a starting point for future individualized applications.

2 Model framework and stages

The lower panel of Figure 1 illustrates the five main stages of the proposed modular framework in direct comparison to the physiological stages represented in the upper panel: (A) a parameter initialization stage, (B) a binaural signal generation stage, (C) a bilateral CI processing stage, (D) an auditory processing model stage including binaural interaction, and (E) a decision model stage. Most stages can be switched on and off, to allow for different test setups. Due to its flexible nature, the framework allows for easy third-party model extensions or substitutions.

2.1 Parameter settings

The parameter initialization stage (Fig. 1A) loads the default parameters for most stages except for the "Decision Model" stage, and, if applicable, changes them to the user-defined settings. More specifically, users can define the switches and set the parameters of stages B, C, and D (Fig. 1) based on their specific research topic, such as switching AGCs on and off, switching coding strategies, or varying cochlear insertion depth.

2.2 Input signals

The signal generation stage (Fig. 1B) contains different functions for generating binaural signals either using HRIRs or by adding ILDs and ITDs directly. For simulations with CI processing, six binaural signal generation functions are provided as example input stimuli that have been commonly used in lateralization and localization studies. These are tone bursts [6, 47], broad-band, low-pass, and high-pass noise bursts [5, 7, 48–51], click trains [6], and speech fragments [6, 52]. Furthermore, custom signals can be loaded as input. By switching the HRIR in the "Input Signals" module (Fig. 1B), the binaural cues can be introduced either by adding ITD or ILD cues independently or by applying HRIRs. In the current framework, the HRIRs were taken from Kayser et al. [53] and were those from the most frontal of the three behind-the-ear (BTE) microphones mounted on the Bruel and Kjaer artificial head and torso simulator (HATS, type 4128C) with a source distance of 3 m. This database covers the range of $\pm 90^\circ$, at 5° intervals in azimuth, and from -10° to 20° in steps of 10° in elevation. It should be noted that the HRIRs of positive azimuthal angles (right side) were simply mirrored from the negative ones (left side).

2.3 CI processing

The CI processing (Fig. 1C) consists of a pre-emphasis filter, AGC, two possible filter banks, four research coding strategies, and electric amplitude mapping. For simulating direct stimulation for single- or multiple- electrodes, this block can be bypassed by providing pulse trains with independently added ITD and ILD cues from the signal generator (Fig. 1B).

2.3.1 Pre-emphasis filter

The first step of the CI processing is a pre-emphasis filter, as a substitute mostly for the ear canal and middle ear transfer characteristics. This is implemented using a first-order Butterworth high-pass filter with a corner frequency of 1200 Hz. The present implementation permits changes to the pre-emphasis frequencies and the deactivation of the pre-emphasis filter.

2.3.2 Automatic gain control

The AGC stage compresses the input signal above a defined threshold (knee point). Here, a general broadband AGC based on Giannoulis et al. [54] is implemented in the framework as an optional stage. In brief, the digital signal is amplified with a gain that is dependent on the compression ratio and on the knee point. Temporal smoothing is achieved with a digital one pole filter, which includes options for adjusting attack and release times. The default threshold, ratio, attack time, and release time were 50 dB SPL, 2:1, 50 ms, and 100 ms, respectively (see Appendix for full details). Comparable results to the literature for both static and moving sources (e.g., [55, 56]) or commercial CI devices have been achieved with these parameters.

2.3.3 Filter banks

After the AGC, the input signal is split into M channels. For this, either a Fast Fourier Transform (FFT)-based filter bank or a gammatone-filter bank [57, 58] can be selected. The channel number and the corner frequencies for each channel can be specified by the user. For the FFT-based filter bank, the sample rate is fixed at 16 kHz and a 128-point moving Hann windowed audio signal covering the range of 250 to 7437 Hz [59]. After transforming the input signal into a spectrogram, the M -channel signal envelope is obtained by summing several (weighted) power frequency bins (please refer to [59–61] for more details). The FFT-based approach and the corresponding parameters coarsely resemble the signal processing of Cochlear Corporation devices [62].

The gammatone filter bank implementation outputs a complex-valued analytical signal in each band, allowing for the extraction of both Hilbert envelope and the temporal fine structure [58]. The default filter bank has center frequencies between 158 and 7480 Hz [63] and consists of 12 fourth-order gammatone filters. The bandwidth of each auditory filter scales with its center frequency according to Glasberg and Moore [64]. The center frequencies and bandwidths are chosen so that the -3 dB corner frequencies of adjacent filters are approximately aligned. The Hilbert envelope in each band is smoothed using a 200 Hz low-pass filter.¹ This processing and parametrization coarsely resemble elements of MED-EL signal processing, even though they do not use a gammatone filter bank but rather bell-shaped finite impulse response filters [65].

2.3.4 Coding strategies

Four CI speech coding strategies were implemented. Three of them are based on well-known commercial coding strategies, or on what is publicly known about them: (1) continuous interleaved sampling (CIS) [66], (2) N-of-M selection as part of the advanced combination encode (ACE) [42, 43], and (3) x -channel fine structure processing (FSx), typically with $x = 4$ [67]. The fourth strategy “peak picking” (PP) [52] contains elements of fundamental asynchronous stimulus timing (FAST) [68], the peak-derived timing (PDT) [69] coding strategy, and fine structure processing. Briefly, instead of a fixed rate, the stimulation time and amplitude of each channel are specified by the local temporal maxima. For the three most apical channels, the pulse timing and amplitude of the local maxima are extracted from the real part of the complex-valued gammatone-filter output, which contains the temporal fine structure of the signal. The remaining mid- to high-frequency channels of PP are similar to the FAST coding strategy, i.e., the local maxima are extracted from the envelope. It should be noted that all these research coding strategies were self-implemented for

¹ It is a 1st-order finite impulse response Butterworth filter with forward and backward filtering. The non-causal approach was used to ease temporal alignment of fine structure and envelope for sound coding strategies. A causal filter implementation would be needed for real-time applications.

different purposes and are not intended to correspond exactly to their commercial counterparts. They were implemented in a way that is not restricted to commercial hardware or computing limitations and to allow for flexible parametrization. This includes the total electrode number along the CI electrode array, the stimulation pulse rate, and allowing for arbitrarily short pulses. Consequently, the default parameters might not be optimized for all tasks. However, they have been validated in different studies with CI users and showed similar speech intelligibility in noise to the commercial strategies (e.g., [8, 52, 61]).

The implemented FSx coding strategy combines elements of the gammatone filter bank based CIS with fine structure processing as in the PP strategy in the x most apical channels. The remaining mid- to high-frequency channels of FSx are based on the low-pass filtered temporal Hilbert envelope and sampled at a fixed pulse rate. For example, with $x = 4$, the four most apical channels contain the temporal fine structure of the signal, where the pulse timing and amplitude of the local temporal maxima are used for the stimulation, related to but not identical to MED-EL FS4 and FS4p strategies.

Since speech is one of the most commonly used stimuli for comparing coding strategies, Figure 2 shows the output at selected stages until the end of the CI processing for the Chinese word “shui” (water) from the Chinese mandarin matrix sentence test (CMNmatrix) [70]. The original word was a single channel signal with a sound level of 60 dB SPL (Fig. 2A). Figures 2B–2D show the two-channel waveform of the processed signal after applying HRIR of -60° (B), pre-emphasis (C), and two independent AGCs (D). The corresponding spectrograms of the left channel signal are shown in the second row. Figures 2E–2H are the electrograms of the compressed signals (D) processed with one of the four research coding strategies: 12-channel CIS (E), FFT-based CIS with 8-of-22 selection (F), 12-channel FSx = 4 (G), 12-channel PP (H).

As expected, the three gammatone filterbank-based coding strategies (E, G, and H) show more similar stimulation patterns compared to the FFT-based CIS with 8-of-22 channel selection (F), without considering the stimulation rate. The interaural pulse time differences were clearly shown in the zoomed-in plot of FSx = 4 (G, channel 1–4) and PP coding strategies (H). In addition, the frequency-dependent delay in each individual channel can be observed in the gammatone filter bank-based electrograms (e.g., at the beginning) when compared with the FFT filter bank based electrogram. Table 1 listed some default settings of these four exemplary coding strategies used in the following demonstrations as well as in the Supplementary materials.

2.3.5 Electric amplitude mapping

To map the acoustic amplitude into the CI user’s electrical dynamic range, a frequency-independent compression is performed for all coding strategies on each channel by applying a logarithmic-shaped compression function with steepness controlled by the parameter $\alpha_c = 415.96$ [60].

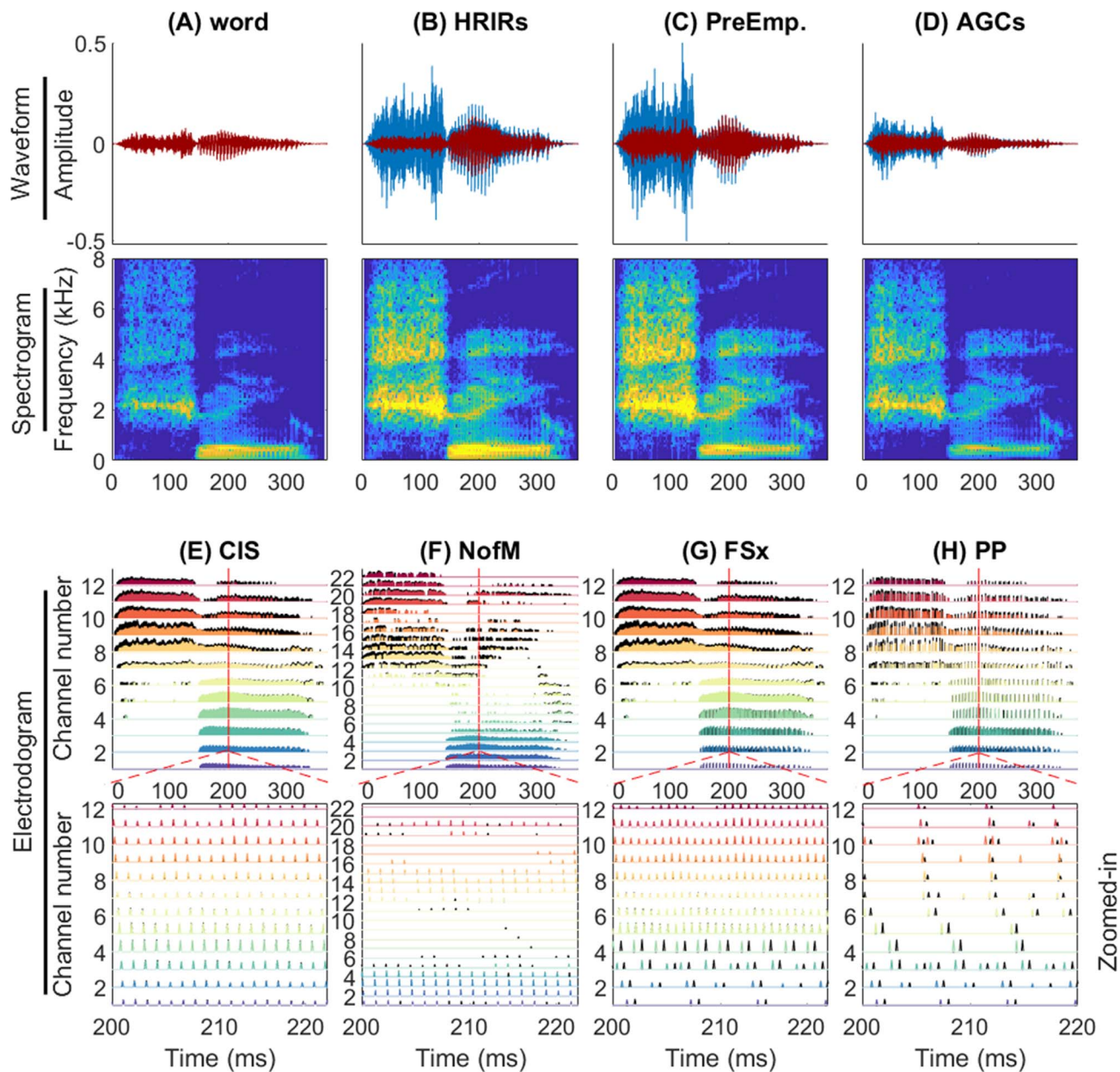


Figure 2. Illustration of input signal and cochlear implant (CI) processing. (A) Single word stimulus at 60 dB SPL. (B) Stimulus output after applying head-related impulse response (HRIR) of azimuth -60° . (C) Stimulus output after pre-emphasis. (D) Stimulus output after applying two independent automatic gain controls (AGCs). Blue and red traces represent the left and right channels, respectively. The corresponding subplots in the second row display spectrograms of the left channel signal. The electrodiagrams resulting from applying the four exemplary coding strategies to the stimulus outputs depicted in panel D are shown in column E to H. (E) 12-channel gammatone filter banks based continuous interleaved sampling (CIS). (F) 8-of-22 channel Fast Fourier Transform (FFT) filter banks based CIS. (G) FSx = 4 strategy. (H) Peak picking (PP) strategy. The bottom subplots in Figures 2E–2H present zoomed-in plots between 200 and 220 ms within the narrow red rectangular at the red line. The electrodiagram for the left CI is shown in black, while color-coded traces represent the electrodiagram for the right CI.

Unless stated otherwise, a threshold level of 25 dB SPL and a maximum level of 65 dB SPL were set as the acoustic input range of each band-pass filtered channel, and the resulting levels were then mapped to the clinical unit (CU) between the CI electrode threshold and saturation levels (e.g., 100 and 200 CU in the examples), respectively.

These values were then mapped onto current, by using the mapping function:

$$I = 17.5 \times 100^{(CU/255)} \mu\text{A}. \quad (1)$$

Finally, biphasic pulses were sent to the auditory nerve (AN) model although the currently implemented CI AN

Table 1. The default settings of the four exemplary research coding strategies: $\textcircled{4}$ 4th-order gammatone filter; $\textcircled{8}$ low-pass filtered Hilbert envelope; $\textcircled{88}$ local maxima based on the real part of the complex-valued gammatone-filter output; $\textcircled{888}$ local maxima based on the low-pass filtered Hilbert envelope.

	12-channel CIS	8-of-22-channel CIS	FSx = 4	PP
Filter bank	gammatone $\textcircled{4}$	FFT	gammatone $\textcircled{4}$	gammatone $\textcircled{4}$
Channel number	12	22	12	12
Pulse rate	900 pps ($\textcircled{88}$)	900 pps	the four most apical channels ($\textcircled{888}$); 900 pps in the remaining channels ($\textcircled{88}$)	three apical channels ($\textcircled{888}$); the remaining channels ($\textcircled{8888}$)

model is agnostic to polarity. Unless stated otherwise, the default biphasic pulse is cathodic/anodic with a phase duration of 25 μ s and an inter-phase gap of 8 μ s. The stimulation was done in an interleaved manner from apical to basal electrodes. Other electric amplitude mapping equations can be implemented to replace equation (1) to convert any clinical unit to a current value in μ A for other CI manufacturers.

2.4 Processing model

The encoded sound after CI processing is then passed into the auditory processing model stage, which starts with the electrode-neuron interface (ENI), then the AN model, and finally the binaural excitation-inhibition (EI) interaction model (see Fig. 1D). Figure 3 illustrates the exemplary settings of these three parts and their interactions. The same electrical AN model and EI model as in Hu et al. [45] were used but were extended from single EI neuron to multiple EI neurons, considering the current spread: each electrical pulse can activate many AN fibers (ANFs), and the outputs of these ANFs then serve as inputs to several EI neurons, linking the left and right hemispheres.

2.4.1 Electrode neuron interface

The ENI is a bridge between the CI electrode output and the AN model. It allows the simulation of different experimental settings. For example, users can specify the cochlear geometry (e.g., dead regions or the number of AN fibers), the CI electrode array (e.g., the length, the electrode number, or the insertion depth), and the spread of excitation in the left and right ear independently. For demonstration purposes, one predefined interface is shown in Figure 3. The exemplary interface assumes: (1) left and right sides are identical; (2) the electrode positions along the unwound cochlea can theoretically reach from the most apical end (35 mm) to the most basal end (0 mm); (3) there are 28 AN fibers per mm, i.e., in total 980 fibers, evenly distributed along the tonotopical axis from 0 mm to 35 mm. Figure 3A shows the location of a predefined example CI electrode array on the straightened and simplified cochlea (brown), where 12 electrodes are equally distributed along the tonotopic axis between 5 to 26 mm. The frequencies shown at the bottom of Figure 3A correspond to the full frequency range (20–20,000 Hz) of a normal-hearing ear

[71]. The input frequency allocation of each electrode is defined as part of the CI coding strategies.

Next, the electric field at the positions of the auditory nerve fibers was simulated, not considering the generally important spatial structure of the fibers. Here, the decline of the electric field, which is commonly referred to as the spread of excitation, is modeled by a double-sided, one-dimensional, exponentially decaying spatial-spread function controlled with the parameter λ [36]:

$$I(x_n, x_{el}) = I_0 \times e\left(-\frac{|x_n - x_{el}|}{\lambda}\right) \quad (2)$$

where I_0 is the active input current at the position x_{el} and I is the effective current due to the spread of excitation to the position of x_n . The presented framework adopted $\lambda = 9$ mm as its default value. It corresponds to a medium amount of spread of excitation, according to electrically evoked compound action potential data [72]. The parameter λ can be interpreted as a surrogate parameter inversely related to the distance between the electrodes and the AN fibers (radial distance is assumed to be constant along the cochlea). Figure 3B shows one set of exemplary spread of excitation functions of the 12 electrodes. When it is intended to use a 3D model, one possibility is to derive the excitation matrix from all electrodes to all neurons or even to all compartments of all neurons within the 3D model and then substitute the spread of the excitation stage with such a matrix.

2.4.2 Auditory nerve model

The electrical AN model of Hamacher [18] in the implementation of Fredelake and Hohmann [36] was used to simulate electrically stimulated AN fibers. This model incorporates the effects of cell membrane, membrane noise, refractory period, mean response latency, and response time jitter in a functional way. As in Hu et al. [45], all default AN parameters were the same as in Fredelake and Hohmann [36]. The main information transferred from the CI processing outputs (electrodegram, with optional current spread) to the AN fibers after the ENI stage includes a sequence of stimulated electrodes and their stimulation time, effective stimulation current I (see Eq. (2)), and phase duration of pulses. The output of the AN stage is a spike matrix containing the effective firing time of action potentials for each fiber. The response probability of each fiber

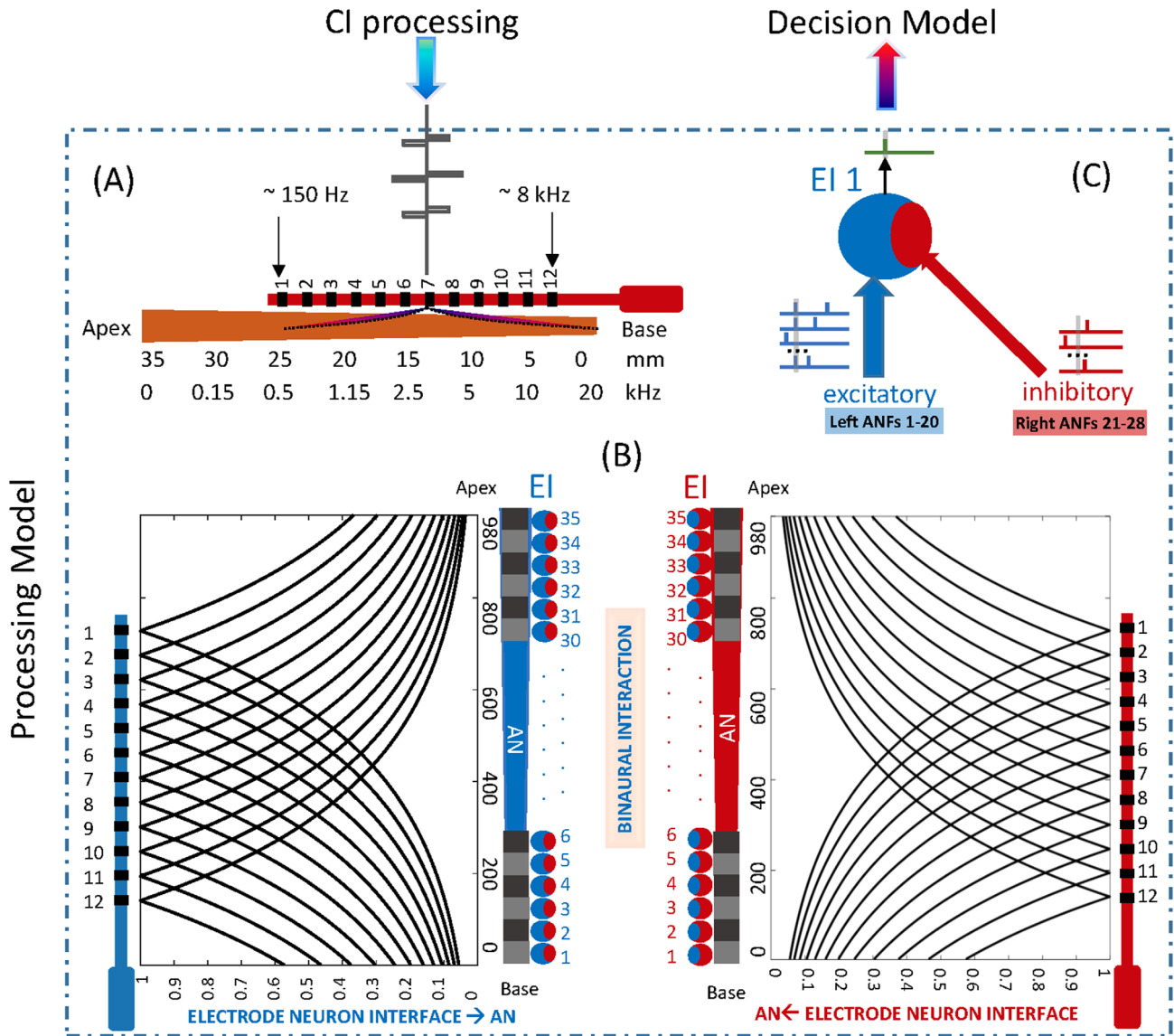


Figure 3. (A) Schematic representation of a predefined cochlear implant (CI) electrode array along a simplified unwound cochlea. The array consists of 12 electrodes equally distributed along the tonotopic axis, spanning from 5 to 26 mm from the base. The depicted frequencies correspond to the entire normal-hearing range of 20 to 20,000 Hz. (B) Spread of excitation functions corresponding to the 12 electrode locations are shown between the CI electrode array and the cochlea, with $\lambda = 9$ mm. A total of 980 Auditory Nerve Fibers (ANFs) and 35 excitation-inhibition (EI) neurons are evenly distributed along the tonotopic axis on each side. The outputs of these ANFs are classified into 35 groups, depicted as interleaved black and gray rectangles, based on their center frequencies. These groups are then transmitted to the corresponding EI neurons, illustrated as solid circles in blue and red. (C) Illustration of the left-hemispheric EI neuron 1 receiving excitatory input from the ipsilateral or left side (represented in blue, 20 ANFs, e.g. ANFs 1–20) and inhibitory input from the contralateral or right side (8 ANFs, e.g., ANFs 20–28, represented in red).

depends on two parameters: Firstly, the membrane voltage after the first pulse phase, which is influenced by the threshold potential, stimulation current and the first phase duration. Secondly, the time difference between the onset of the current pulse and the last action potential. When following the above-mentioned input–output format, the AN model can be replaced by other AN models (e.g., [19, 20, 22]).

In the example, 980 simulated AN fibers on each side were created. They were 35 groups of 28 fiber bundles sharing the same functional properties. To be more specific, a bundle of 28 fibers was first randomly generated, and then replicated 35 times along the tonotopic axis and across ears, simulating identical left and right sides. Each bundle was assigned to the excitatory and inhibitory inputs of a single EI neuron described in the next section (binaural neuron

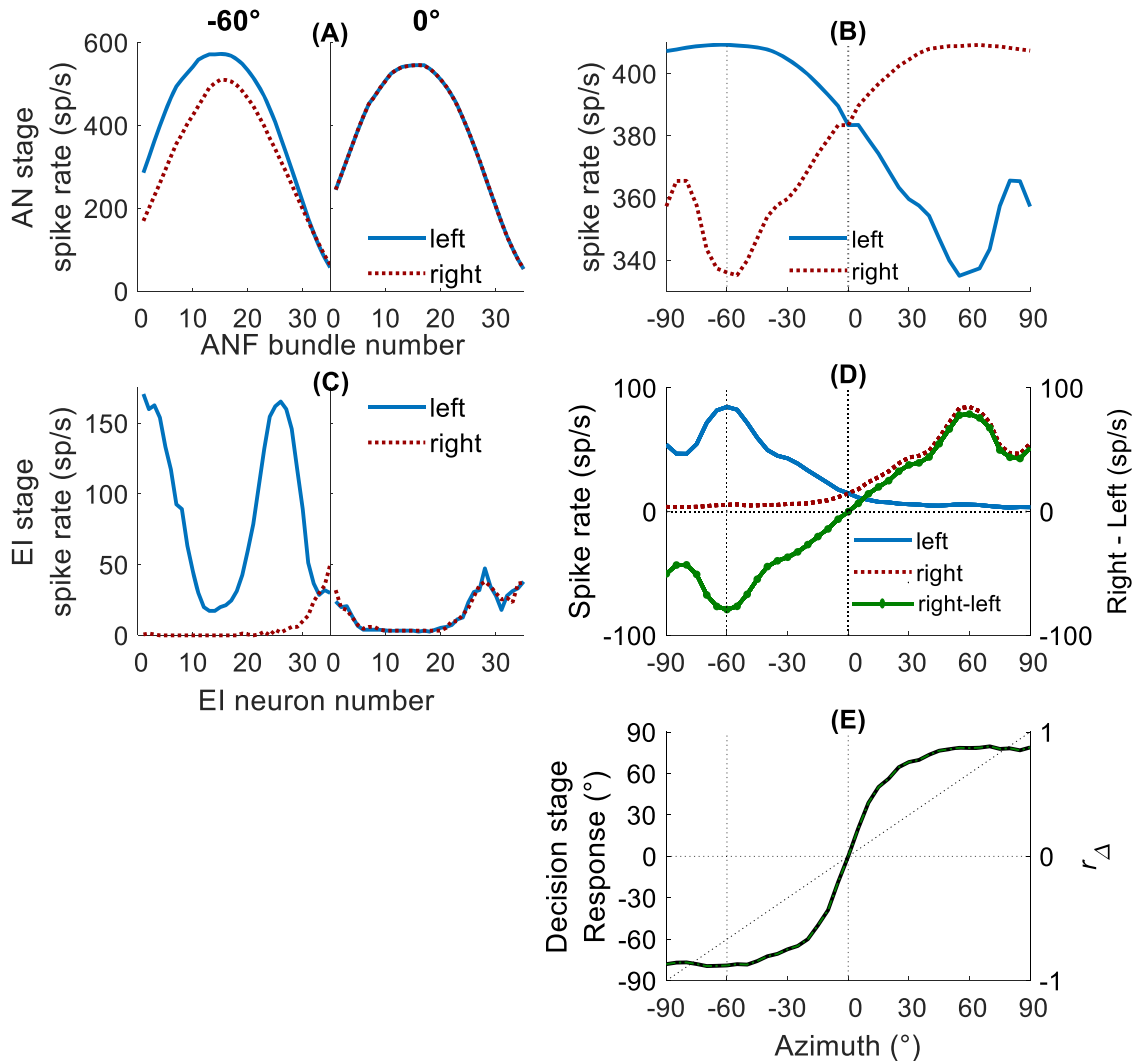


Figure 4. Simulation outputs at AN, EI, and decision stages for the same Chinese word “shui” at 60 dB SPL. (A) average spike rates of each ANF bundle along the tonotopic axis for left (blue) and right (red) sides, with azimuth of -60° (left subplot) and 0° (right subplot). (B) average spike rates across all 980 ANFs for different azimuth angles (left and right AN rate-azimuth tuning curves, in blue and red). (C) spike rate of each EI along the tonotopic axis of the left (blue) and right (red) sides, at azimuth angles of -60° and 0° . (D) average spike rates across 35 EI neurons for different azimuth angles (left and right EI rate-azimuth tuning curves, blue and red) and the corresponding spike rate difference between the right and left hemisphere (green). (E) the normalized rate difference (right y -axis) and the corresponding predicted response azimuth angles (left y -axis) after applying a linear mapping.

model). Each bundle of these 28 AN fibers is represented by a black or gray color-filled rectangle in Figure 3B. Figure 4A shows the average spike rates of each fiber bundle along the tonotopic axis from 0 to 35 mm. These spikes were evoked by the same Chinese word “shui” shown in Figure 2 (60 dB SPL) convolved with the HRIR of azimuth -60° or 0° (with AGCs). Figure 4B shows the average spike rates of all AN fibers evoked by the same stimuli presented at different azimuthal angles. The AN spike patterns are useful for understanding the input–output relations of each processing stage shown in Figure 1. For example, it enables the user to check if spatial cues are

still encoded at this level, to understand at which level the information may get corrupted or lost.

2.4.3 Binaural neuron model

As in Klug et al. [40], the single EI neuron model used in Hu et al. [45] was extended to a population EI model neurons along the tonotopic array. In the current model configuration, the outputs of the left and right ANFs (Fig. 4A, 35 bundles and in total 980 ANFs on each side) serve as the inputs to the EI stage. There are 35 EI neurons on each hemisphere in response to these 35 ANF bundles

(Fig. 3B). Each EI neuron receives 20 excitatory inputs from the neighboring ipsilateral 20 AN fibers and eight inhibitory inputs from the contralateral eight AN fibers, all placed around the same center frequencies. Figure 3C illustrates one left-hemispheric EI neuron receiving excitatory input from the ipsilateral side (20 ANFs) and inhibitory input from the contralateral side (eight ANFs). The EI spike outputs were generated according to the coincidence counting model described by Ashida et al. [44] but with the same parameters as in Klug et al. [40] and Hu et al. [45]. Briefly, ipsilateral ANF spikes are convolved with a 1.1 ms long rectangular window (of amplitude 1) and contralateral spikes with a 3.1 ms long window of amplitude 2, simulating inhibition. If the sum of both convolved signals reaches the response threshold of 3, an action potential is generated. Within the refractory period of 1.6 ms, only the first action potential leads to a spike. Figure 4C shows the spike pattern of the 35 EI neurons evoked by the 60 dB SPL word “shui” convolved with the HRIR of azimuth -60° (left subplot) and 0° (right subplot), where a reduction of on-frequency rate could be observed. Figure 4D shows the average spike rates across the 35 EI neurons evoked by the same word at different azimuth angles (blue and red curves are corresponding to left and right hemispheres). Under a fixed combination of the above-mentioned EI neuron parameters (e.g., coincidence and inhibitory window size, response threshold, refractory period), different properties of the excitatory and inhibitory AN fibers from the bundle of 28 AN fibers can also lead to different results. In the demonstrated setups, inhibition will eventually win and suppress excitation when a pulse is presented bilaterally to a single electrode pair, stimulating all fibers in the relevant bundles. However, the different durations of inhibitory and excitatory inputs result in an additional rate-dependence of this mechanism. It’s noteworthy that the rate and level-dependent outputs of the EI model, shown in Supplementary Figure 12, align with the results from [73] in a qualitative manner.

2.5 Decision model

As many questions about the human “decision stage” remain open, it is challenging to define a single comprehensive model of the decision-making process for different tasks and applications. In previous studies, various central read-out mechanisms have been used (for a review, see [74]). Kelvasa and Dietz [34] showed that the averaged hemispheric LSO model response difference is approximately proportional to azimuthal sound source localization in CI users. Klug et al. [40] adopted this simple linear mapping and demonstrated that the mean rate difference between the left-hemispheric and the right-hemispheric binaural EI model neurons together with a subject specific factor are largely sufficient to explain both envelope-ITD based and ILD based lateralization for narrow-band stimuli in normal hearing listeners. As previously emphasized, it is out of the scope of the current model framework to provide a validated decision model for all possible applications. For demonstration purposes a normalized spiking rate

difference r_Δ modified from Klug et al. [40] was employed. It was obtained as

$$r_\Delta = \frac{\sum_{n=1}^N (r_{R,n} - r_{L,n})}{\sum_{n=1}^N (r_{R,n} + r_{L,n})}, \quad (3)$$

where N is the number of EI model neurons on each side (35 in this study). The spike rates of the two EI model neurons on the right and left hemispheres (with the same n th center frequency) are $r_{R,n}$ and $r_{L,n}$, respectively. Comparing the rate difference ($\sum_{n=1}^N (r_{R,n} - r_{L,n})$; Fig. 4D, green, Right–Left)

to the normalized rate difference (Fig. 4E, r_Δ), the normalization leads to a value between -1 and 1 . The normalization was achieved by dividing the rate difference with the total spike rate across two hemispheres, $\sum_{n=1}^N (r_{R,n} + r_{L,n})$,

which requires a non-zero total spike count. If one hemispheric rate drops to zero, the decision variable is either -1 or 1 . This assumes that as soon as only one hemisphere is active exclusively, localization does not depend on the spike rate. This assumption might not be valid in other applications. For example, in a lateralization experiment, the amount of spike rate might be related to the extent of lateralization. In such a case, normalization might reduce the possible effects of certain factors on the extent of lateralization. For instance, the lateralization differences between single and multiple electric stimulations might be underestimated because of normalization. It should also be noted that such normalization can, to some extent, diminish some features that are presented in the rate difference tuning curves, such as the local peaks shown near the azimuth of $\pm 45^\circ$ in Figure 4D. In addition, as shown in Figure 4C, the 35 EI spike rates along the tonotopic axis evoked by the Chinese word “shui” were not identical, however in the calculation of the normalized r_Δ (Eq. (3)), all the EI neurons are equally weighted. Further optimization could consider tonotopically weighted rate differences (e.g., taking into account spectral dominance region [75]).

Afterwards, alternative mapping functions can be used to project the normalized spiking rate difference r_Δ on various scales to account for a range of psychoacoustic data (see Supplementary Materials). Similar to Kelvasa and Dietz [34] and Klug et al. [40], a simple linear function

$$y = c \times r_\Delta + d \quad (4)$$

was introduced to map the r_Δ function to azimuth angles. For example, $c = 90^\circ$, and $d = 0^\circ$ were used in equation (4) to map the $r_\Delta \in [-1, 1]$ to azimuth responses between $[-90^\circ, 90^\circ]$ shown in classic localization experimental data (e.g., Fig. 4E). It should be noted that no other dedicated individual calibration of the linear mapping, such as using a subject-specific factor [40], was applied in this demonstration.

3 Discussion

The goal of this research was to provide a general CI model framework for simulating the spatial hearing abilities

of bilateral CI users in a broad range of binaural experiments. The framework combines various CI processing strategies with characteristics of different ascending auditory pathway stages in a modular and open-source software. To demonstrate the outputs of different model stages, a localization experiment setup was simulated with a simple decision model. Signal representations were shown after different CI processing stages. Simulated neural response rates were shown at the level of the AN and after binaural interaction, and finally the decision variable was visualized as a function of sound source azimuth.

Beyond the examples demonstrated in this paper, the model has been applied to several psychoacoustic datasets including lateralization and localization experiments in bilateral CI users (e.g., [6, 76–80]). In general, with the same AN, binaural, and multi-task decision stages, the model was able to capture the overall results of the bilateral CI users in these experimental studies, such as the HRIR-, ITD-, and ILD-based localization/lateralization judgments, the effect of coding strategies, and the effect of level-dependent compression. These results are provided in the [Supplementary Materials](#).

The following subsections discuss the limitations of the current model framework and possible applications with further extension.

3.1 Limitations

In the CI processing stage of the current model implementation, four different coding strategies and two different filter banks were implemented. Their implementations, particularly for CIS and FSx were partially based on secondary literature due to limited published details from manufacturers. It should be noted that they were not validated to the same extent as commercial coding strategies. However, some studies using the self-implemented coding strategies demonstrated similar speech reception thresholds using direct CI stimulation when compared to participants using their own speech processors [8, 52]. Additionally, as discussed in the following [Section 3.2](#), users should be able to easily integrate their own coding strategies into the model framework. It is crucial to recognize that the current binaural stage solely incorporates the LSO circuitry. However, for a more comprehensive understanding and representation of binaural processing, it may be necessary to include the medial superior olive (MSO) circuitry in future work.

The decision stage model was introduced to mimic how central pathways extract a lateralization or localization percept from the ensemble of binaural model neuron spikes. Due to the complexity, a variety of different decoding stages have been suggested in previous studies (for a review, see [74]). In this paper, the normalized rate difference r_{Δ} and linear mapping functions were used to map the model outputs to different psychoacoustic data. With this multi-task decision model, ITD and ILD threshold, lateralization, and localization experiments were simulated, using single or multi-electrode direct stimulation or the simulated CI processor. Overall, this simplified decision

model showed plausible results in predicting the average CI user’s performance with different stimulation techniques. However, not all features of the selected experiments could be captured with this approach. For example, the model may underestimate the effect of multiple electrode stimulation compared to single electrode stimulation, the differences between the coding strategies with and without fine structure encoding, or exaggerate the detrimental effect of two independent AGCs by treating all the EI neurons equally. To be more specific, the normalized rate difference assigns the same weight to the outputs of all 35 EI neurons and this might be one of the reasons for some of the unusual simulation results included in the [Supplementary Materials](#). For example, one unusual simulation result that occurred with the use of two independent AGCs ([Supplementary Figure 5](#)) might be due to this limitation. There the model suggested a reverse in the perceived localization direction for broad-band stimuli at high input levels. However, to our knowledge, this effect has not been explicitly reported in the localization literature with bilateral CI users, although inverted ILDs have been reported in the low-frequency channels with bilaterally independent AGCs (e.g., [46, 48, 81, 82]) or N-of-M selection (e.g., [34]). Sparsely distributed reverse localization data points were shown in some Advanced Bionics bilateral CI users [46], but it was not specifically discussed by the authors. This implies that either the current simplified decision model exaggerated the effect of the inverted ILDs in the low-frequency channels or the reversed direction may happen in a minority of CI users under some extreme circumstances. To solve the issues with the decision stage, more advanced decision models such as weighted rate differences and machine learning based methods (e.g., [83, 84]) might help. Furthermore, a decision stage including a listener-specific factor (e.g., [40]) for fitting the model outputs to each participant might be required for accounting for individualized performance in the psychoacoustic data.

All the examples demonstrated here were based on the current “general purpose” parameter settings. However, making individual predictions is possible by customizing the model parameters of other stages (CI processing, ENI, AN, binaural stage) to their clinical profile, additionally to the above-mentioned decision model. For example, all the simulation results presented in this paper and the [Supplementary materials](#) assumed ideal bilateral synchronization and bilaterally symmetrical hearing (same CI fittings, ENIs, same AN, and EI neurons) in both ears. To simulate configurations such as a dominant CI, different ENIs, and ANFs (e.g., different AN properties, neuron number) for each EI neuron across the tonotopic array or even across ears need to be set. For this approach, the parameters can be easily adapted by including asymmetric neuron survival, different electrode insertion depths, the spread of excitation, CI dynamic range, two unsynchronized independent speech processors, and bilaterally mismatched fittings. Consequently, further studies involving, e.g., mismatched insertion depth [85–87], bilaterally synchronized N-of-M channel selection [34, 88–91], and bilaterally synchronized AGCs [46, 81, 82] are possible within the current framework.

3.2 Possible applications with further extension

The model framework is designed to be flexible enough to cover a wide range of binaural research topics. Although it was only tested with lateralization and localization experiments (see also [Supplementary materials](#)) in this study, by selecting different stage combinations and different decision models, its applications can go beyond the provided experimental examples. By referring to the MATLAB examples provided, users are able to replace the current functions with customized ones, such as incorporating other HRIRs or BRIRs for complex acoustic environments (e.g., [53, 92–98]), self-developed AGCs, different coding strategies, different AN models, or different decision models. Each model stage alone can be used or extended for both research and education purposes. The framework can thus be used as a tool to help validate binaural algorithms [99, 100], for designing binaural studies with real CI users or to study the implications of binaural fitting [87]. Although the current model framework is binaural by default, it can be used for monaural purposes or only include certain stages. For example, one of the main stages of the current framework, the implemented CI processing, includes most of the technical details of a typical CI speech processor (e.g., AGC, filter banks, coding strategies). Thus, this model can also be used for developing or validating new sound processing strategies (e.g., [52, 59, 61, 90, 101–106]).

Since modeling studies are not restricted by a clinic appointment timeframe, this model can also be a powerful tool for testing new concepts or experiments before starting clinical tests, where recruiting many participants with the same coding strategies or pre-processing is very time-consuming, expensive, and cumbersome. For example, in pre-clinical studies, certain specific settings might deviate from parameters typically used by clinical populations (e.g., switching off AGC or assuming synchronized AGCs, using single or multi-electrode direct stimulation instead of the participant's own speech processor), which makes it difficult to predict the outputs of CI users in real-life with their everyday processor settings. The framework can help in “generalizing” such laboratory research findings by exploring the impact of variations in the parameter settings across a range of ecologically relevant acoustic conditions.

In addition to the above-mentioned applications which focus on the CI device-related limitations (e.g., binaural cue distortions) and AN-related issues (e.g., ENI, ANFs, and EI interaction), the framework also allows for third-party model extensions or substitutions for other applications, although less straightforward. For example: (1) the demonstrated ENI assumed an unwound cochlea and all AN fibers were equally distributed between the most basal (0 mm) and the most apical (35 mm) ends. It is possible and might be necessary to incorporate 3D implanted cochlear models (e.g., [25–28, 30, 107–109]) for the ENI individualization to simulate interactions between implanted cochlear anatomy and stimulation parameters (e.g., [29, 110, 111]). (2) The current AN model does not take into account the fiber morphology and orientation towards the electric field,

as well as its first and second derivatives that are essential for simulating the activation of AN fibers. The present model is insensitive to the number and polarity of pulse phases, and to the size of the inter-phase gap. Several studies have investigated the polarity sensitivity of auditory nerve fibers using eCAP recordings in human CI users [112, 113], and to model such experiments, other AN models (e.g., [19, 20, 114]) can be incorporated. (3) Extensions to predict bilateral speech intelligibility are possible by connecting the AN spike outputs to the backend processing stages of other purposes such as an automatic speech recognition system (e.g., [37, 110, 111]) or by replacing it with more abstract internal model representations (e.g., [115, 116]). (4) Applications for predicting CI users' bilateral loudness perception (e.g., [117]) by combining, e.g., the AN fiber outputs with a physiologically-motivated binaural loudness model (e.g., [118]). (5) The electrical hearing of one side could be replaced with an acoustic hearing periphery (e.g., [40, 45, 119]) or by combining the electrical hearing with an acoustic hearing on the same side [110, 120]. In this way, the model framework can also be useful for the investigation of more complex input systems (e.g., CI users with single-sided deafness or electro-acoustic stimulation) in complex acoustic scenarios.

4 Summary and conclusion

An open-source computer model framework for simulating the spatial hearing abilities of bilateral CI listeners was developed based on the model of Kelvasa and Dietz [34]. It includes a binaural signal generation stage, a CI processing stage with a wider range of speech-coding strategies, an electrode-neuron interface stage, an auditory nerve model stage, an excitation-inhibition binaural interaction model stage, and a decision stage based on the normalized hemispheric rate differences. The framework can be used to simulate left/right discrimination, lateralization, and localization performance of CI users for single- or multiple-electrode direct stimulation, and free-field listening experiments. It can be extended or modified to model further experiments. The MATLAB code of the model framework, as well as the code to reproduce the model data and figures, are published as open source on Zenodo [121].

Data Availability Statement

The research data associated with this article are included in the [Supplementary materials](#) of this article and are also available on Zenodo [121].

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions

Conceptualization and framework design: HH, MD. Framework implementation: HH. Framework optimization and validation: HH, BW, SA, JK, RF. Paper writing: HH, BW, SA, JK, RF, DV, MD.

Funding

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program grant agreement no. 716800 (ERC Starting Grant to Mathias Dietz). DV and BW were supported by a UK Medical Research Council Senior Fellowship (MR/S002537/1).

Acknowledgments

The AGC MATLAB code used in the framework was implemented and validated as part of Jakob Diefel's bachelor thesis. The authors thank Stephan D. Ewert for comments on earlier versions of the manuscript.

Supplementary materials

The supplementary material of this article is available at <https://acta-acustica.edpsciences.org/10.1051/aacus/2023036/olm>.

Supplementary Figure 1: An overview of the four experiments selected for the purpose of demonstrating and validating the proposed framework. Experiment 1 and 2 simulate the localization (with HRIR) and lateralization (without HRIR) experiments in bilateral CI users, respectively. In addition, the effect of AGC was illustrated in experiment 1a and 1b (with and without AGC). Experiment 3 and 4 simulate the ITD just noticeable difference (JND) or ITD- and ILD-based lateralization, where single-electrode or multiple-electrode pairs were stimulated without a CI processing stage.

Supplementary Figure 2: Waveforms (column 1), spectrogram (column 2) and electrograms (column 3-6) of these six acoustic stimuli (250 Hz sinusoidal tones, BB-, LP-, HP- noise bursts, clicks, and Chinese word 'shui') at 60 dB SPL. The electrograms of four coding strategies are shown in column 3-6: 12-channel gammatone filter-bank based CIS (column3), FSx = 4 (column4), PP (column5), and the 8-of-22 channel FFT filter-bank based ACE* (column 6). The Matlab function for creating these plots is 'plot_supfig2.m'

Supplementary Figure 3: The electrograms (A, C, within 60 ms, from 60°) and the ILD-azimuth functions (B, D) of the BB stimuli without (AGC = 0) and with AGC (AGC = 1) obtained from the 12-channel FSx=4 coding strategies. The electrogram of the right CI is shown in black behind the colored left electrogram. The broadband ILDs before the CI processing were plotted as black dotted lines, while the ILDs of each individual

electrode pair were plotted as colored solid lines. The simulated source level was 70 dB SPL. The corresponding Matlab function for plots is 'plot_supfig3'.

Supplementary Figure 4: Examples of the effects of different coding strategies (different colors represent 12-channel gammatone filter-bank based CIS, 8of22-channel CIS, FSx = 4, and PP without AGC) on the azimuthal localization performance with the six stimulus types (column 1-6) at 40 dB SPL: the spike pattern along the tonotopic axis for 0° azimuth (upper panels), response azimuth over target azimuth (lower panels) functions. The corresponding Matlab functions for generating the simulation and the plots are 'gen_supfig4.m, and 'plot_subfig4.m'.

Supplementary Figure 5: Demonstration of the effect of AGC on the EI-outputs (localization) for the six stimuli at different levels with 12-channel FSx = 4 coding strategies. The first row shows the estimated direction vs the original direction without AGC, the second one with AGC. The corresponding Matlab functions for generating the simulation and the plots are 'gen_supfig5.m, and 'plot_subfig5.m'.

Supplementary Figure 6: Example of an estimated psychometric ILD curve on the r_A scale (left y-axis) and the corresponding lateralization scale (right y-axis) after the mapping. The slope, width, and range are indicated by a , ω and range.

Supplementary Figure 7: The spike pattern along the tonotopic axis for ITD = 0 and ILD = 0 (upper panels), r_A -ITD (middle panels) and r_A -ILD (bottom panels) functions of the EI-outputs for the six stimuli types with a general 12-channel CIS coding strategy without AGC at 6 different levels (represented by different colors). The corresponding Matlab functions for generating the simulation and the plots are 'gen_supfig7.m' and 'plot_subfig7.m'.

Supplementary Figure 8: Effects of different coding strategies (different colors represent 12-channel gammatone filter-bank based CIS, FSx=4, PP, and 8of22-channel FFT-based CIS; without AGC) on the ITD and ILD discrimination for the six stimuli types (column 1-6) at 60 dB SPL: the spike pattern along the tonotopic axis for ITD = 0 and ILD = 0 (upper panels), % right response-ITD (middle panels) and % right response-ILD (bottom panels) discrimination functions. The corresponding Matlab functions for generating the simulation and the plots are 'gen_supfig8.m' and 'plot_subfig8.m'.

Supplementary Figure 9: stimulation locations used in experiment 3 and 4. The marked red points are located at 5, 10, 15, 20, and 25 mm along the tonotopic axis.

Supplementary Figure 10: the normalized spike rate difference (r_A)-IPD functions (A, 380 μ A; B, 600 μ A) and the ITD JNDs-pulse rate functions (C, blue, 380 μ A; purple, 600 μ A) of single electrode stimulation at different pulse rates. The corresponding Matlab functions for the simulation and the generation of the plots are 'gen_supfig10.m' and 'plot_subfig10.m'.

Supplementary Figure 11: The average spike rate of the 35 EI-neurons along the tonotopic axis at ITD = 0 and ILD = 0 for 100 pps (A) and 1000 pps (B). On the y-axis, the spike rate is shown in spikes per second (sp/s), and the x-axis is the tonotopic location as well as the

corresponding greenwood frequency. The normalized r_{Δ} -ITD (C, left y-axis) and r_{Δ} -ITD functions (D, left y-axis) and the corresponding lateralization results (right y-axis). The dotted lines in D are the corresponding results for the 1000 pps. The stimulation level is 480 μ A. The corresponding Matlab functions for the simulation and generating the plots are 'gen_supfig11.m' and 'plot_subfig11.m'.

Supplementary Figure 12: The average spike rate of the 35 EI-neurons along the tonotopic axis for single-electrode pair stimulation located at 15 mm. Three pulse rates of 100 pps (A), 500 pps (B) and 1000 pps (C) were simulated for unmodulated pulses at eight different input levels with ITD = 0 and ILD = 0. On the y-axis the spike rate is shown in spikes per second (sp/s), and the x-axis is the location along the basilar membrane. Panel (D) shows the on-site (at 15 mm) EI neuron's spike rate-level for these three pulse rates. The corresponding Matlab functions for the simulation and generating the plots are 'gen_supfig12.m' and 'plot_subfig12.m'.

References

1. M. Hainarosie, V. Zainea, R. Hainarosie: The evolution of cochlear implant technology and its clinical relevance, *Journal of Medicine and Life* 7, 2 (2014) 1–4.
2. F.-G. Zeng: Celebrating the one millionth cochlear implant, *JASA Express Letters* 2 (2022) 077201.
3. R.Y. Litovsky, A. Parkinson, J. Arcaroli: Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear Hear* 30 (2009) 419–431.
4. P. Nopp, P. Schleich, P. D'Haese: Sound localization in bilateral users of Med-El Combi 40/40+ cochlear implants. *Ear and Hearing* 25 (2004) 205–214.
5. S.A. Ausili, M.J.H. Agterberg, A. Engel, C. Voelter, J.P. Thomas, S. Brill, A.F.M. Snik, S. Dazert, A.J. Van Opstal, E.A.M. Mylanus: Spatial hearing by bilateral cochlear implant users with temporal fine-structure processing. *Frontiers in Neurology* 11 (2020) 915.
6. B. Laback, S.-M. Pok, W.-D. Baumgartner, W.A. Deutsch, K. Schmid: Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors. *Ear and Hearing* 25 (2004) 488–500.
7. B.U. Seeber, H. Fastl: Localization cues with bilateral cochlear implants. *Journal of the Acoustical Society of America* 123 (2008) 1030–1042.
8. H. Hu, M. Dietz, B. Williges, S.D. Ewert: Better-ear glimpsing with symmetrically-placed interferers in bilateral cochlear implant users. *Journal of the Acoustical Society of America* 143 (2018) 2128–2141.
9. D.W. Grantham, D.H. Ashmead, T.A. Ricketts, R.F. Labadie, D.S. Haynes: Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants*. *Ear and Hearing* 28 (2007) 524–541.
10. R.Y. Litovsky, M.J. Goupell, S. Godar, T. Grieco-Calub, G. L. Jones, S.N. Garadat, S. Agrawal, A. Kan, A. Todd, C. Hess, S. Misurelli: Studies on bilateral cochlear implants at the University of Wisconsin's binaural hearing and speech laboratory. *Journal of the American Academy of Audiology* 23 (2012) 476–494.
11. M. Dietz: Models of the electrically stimulated binaural system: a review. *Network: Computation in Neural Systems* 27 (2016) 186–211.
12. G.E. O'Brien, J.T. Rubinstein: The development of biophysical models of the electrically stimulated auditory nerve: single-node and cable models. *Network: Computation in Neural Systems* 27 (2016) 135–156.
13. F. Rattay: The basic mechanism for the electrical stimulation of the nervous system. *Neuroscience* 89 (1999) 335–346.
14. J. Boulet: Biophysical and phenomenological models of cochlear implant stimulation. Ph.D. Thesis. McMaster University, 2016.
15. N.S. Inennov, J.T. Rubinstein: Stochastic population model for electrical stimulation of the auditory nerve. *IEEE Transactions on Biomedical Engineering* 56 (2009) 2493–2501.
16. N. Mangado, J. Pons-Prats, M. Coma, P. Mistrik, G. Piella, M. Ceresa, M.Á.G. Ballester: Computational evaluation of cochlear implant surgery outcomes accounting for uncertainty and parameter variability. *Section of the Journal Frontiers in Physiology. Computational Physiology and Medicine* 9 (2018) 498.
17. I.C. Bruce, L.S. Irlicht, M.W. White, S.J.O. Leary, S. Dynes, E. Javel, G.M. Clark: A stochastic model of the electrically stimulated auditory nerve: pulse-train response. *IEEE Transactions on Biomedical Engineering* 46 (1999) 630–637.
18. V. Hamacher: Signalverarbeitungsmodelle Des Elektrisch Stimulierten Gehörs. Ph.D. Thesis. RWTH Aachen, Aachen, 2004.
19. C.D.F. Horne, C.J. Sumner, B.U. Seeber: A phenomenological model of the electrically stimulated auditory nerve fiber: temporal and biphasic response properties. *Frontiers in Computational Neuroscience* 10 (2016) 8.
20. S.N. Joshi, T. Dau, B. Epp: A model of electrically stimulated auditory nerve fiber responses with peripheral and central sites of spike generation. *Journal of the Association for Research in Otolaryngology: JARO* 18 (2017) 323–342.
21. J.H. Goldwyn, J.T. Rubinstein, E. Shea-Brown: A point process framework for modeling electrical stimulation of the auditory nerve. *Journal of Neurophysiology* 108 (2012) 1430–1452.
22. S. Tabibi, J. Boulet, N. Dillier, I.C. Bruce: Phenomenological model of auditory nerve population responses to cochlear implant stimulation. *Journal of Neuroscience Methods* 358 (2021) 109212.
23. M.J. van Gendt, M. Siebrecht, J.J. Briaire, S.M. Bohte, J.H. M. Frijns: Short and long-term adaptation in the auditory nerve stimulated with high-rate electrical pulse trains are better described by a power law. *Hearing Research* 398 (2020) 108090.
24. M. Takanen, I.C. Bruce, B.U. Seeber: Phenomenological modelling of electrically stimulated auditory nerve fibers: a review. *Network: Computation in Neural Systems* 27 (2016) 157–185.
25. T. Hanekom, J.J. Hanekom: Three-dimensional models of cochlear implants: a review of their development and how they could support management and maintenance of cochlear implant performance. *Network: Computation in Neural Systems* 27 (2016) 67–106.
26. I.M. Lei, C. Jiang, C.L. Lei, S.R. de Rijk, Y.C. Tam, C. Swords, M.P.F. Sutcliffe, G.G. Malliaras, M. Bance, Y.Y.S. Huang: 3D printed biomimetic cochleae and machine learning co-modelling provides clinical informatics for cochlear implant patients. *Nature Communications* 12 (2021) 6260.
27. S. Bai, J. Encke, M. Obando-Leitón, R. Weiß, F. Schäfer, J. Eberharter, F. Böhnke, W. Hemmert: Electrical stimulation in the human cochlea: a computational study based on high-resolution micro-CT scans, *Frontiers in Neuroscience* 13 (2019) 1312.

28. J.H.M. Frijns, J.J. Briaire, J.J. Grote: The importance of human cochlear anatomy for the results of Modiolus-Hugging multichannel cochlear implants. *Otology and Neurotology* 22 (2001) 340–349.
29. W. Nogueira, D. Schurzig, A. Büchner, R.T. Penninger, W. Würfel: Validation of a cochlear implant patient-specific model of the voltage distribution in a clinical setting. *Frontiers in Bioengineering and Biotechnology Section Computational Physiology and Medicine* 4 (2016) 84.
30. R.K. Kalkman, J.J. Briaire, J.H.M. Frijns: Current focussing in cochlear implants: an analysis of neural recruitment in a computational model. *Hearing Research* 322 (2015) 89–98.
31. C. Jiang, S. Singhal, T. Landry, I.V. Roberts, S.R.D. Rijk, T. Brochier, T. Goehring, Y.C. Tam, R.P. Carlyon, G.G. Malliaras, M.L. Bance: An instrumented cochlea model for the evaluation of cochlear implant electrical stimulus spread. *IEEE Transactions on Biomedical Engineering* 68 (2021) 2281–2288.
32. C.M. McKay, K.R. Henshall, R.J. Farrell, H.J. McDermott: A practical method of predicting the loudness of complex electrical stimuli, *Journal of the Acoustical Society of America* 113 (2003) 2054–2063.
33. C.M. McKay: Applications of phenomenological loudness models to cochlear implants. *Frontiers in Psychology* 11 (2021) 611517.
34. D. Kelvasa, M. Dietz: Auditory model-based sound direction estimation with bilateral cochlear implants. *Trends in Hearing* 19 (2015) 2331216515616378.
35. M. Nicoletti, C. Wirtz, W. Hemmert: Modeling sound localization with cochlear implants. In: J. Blauert, Ed. *The Technology of Binaural Listening*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013: 309–331.
36. S. Fredelake, V. Hohmann: Factors affecting predicted speech intelligibility with cochlear implants in an auditory model for electrical stimulation. *Hearing Research* 287 (2012) 76–90.
37. T. Brochier, J. Schlittenlacher, I. Roberts, T. Goehring, C. Jiang, D. Vickers, M. Bance: From microphone to phoneme: an end-to-end computational neural model for predicting speech perception with cochlear implants. *IEEE Transactions on Biomedical Engineering* 69 (2022) 3300–3312.
38. M. Dietz, G. Ashida: Correction To: Computational Models of Binaural Processing: With 93 Illustrations. In: R.Y. Litovsky, M.J. Goupell, R.R. Fay, A.N. Popper, Eds. *Binaural Hearing*, Springer International Publishing: Cham, 2021: C1–C1.
39. M. Takanen, O. Santala, V. Pulkki: Visualization of functional count-comparison-based binaural auditory model output. *Hearing Research* 309 (2014) 147–163.
40. J. Klug, L. Schmors, G. Ashida, M. Dietz: Neural rate difference model can account for lateralization of high-frequency stimuli. *Journal of the Acoustical Society of America* 148 (2020) 678–691.
41. J. Encke, M. Dietz: A hemispheric two-channel code accounts for binaural unmasking in humans. *Communications Biology* 5 (2022) 1122.
42. J.F. Patrick, P.A. Busby, P.J. Gibson: The development of the nucleus freedom cochlear implant system. *Trends in Amplification* 10 (2006) 175–200.
43. G. Clark: *Cochlear implants: fundamentals and applications*. Springer NY, New York, 2003.
44. G. Ashida, J. Kretzberg, D.J. Tollin: Roles for coincidence detection in coding amplitude-modulated sounds. *PLOS Computational Biology* 12 (2016) e1004997.
45. H. Hu, J. Klug, M. Dietz: Simulation of ITD-dependent single-neuron responses under electrical stimulation and with amplitude-modulated acoustic stimuli. *Journal of the Association for Research in Otolaryngology* 23 (2022) 535–550.
46. M.T. Pastore, R. Pulling Kathryn, C. Chen, A. Yost William, F. Dorman Michael: Effects of bilateral automatic gain control synchronization in cochlear implants with and without head movements: sound source localization in the frontal hemifield. *Journal of Speech, Language, and Hearing Research* 64 (2021) 2811–2824.
47. M. Eklöf, B. Tideholm: The choice of stimulation strategy affects the ability to detect pure tone inter-aural time differences in children with early bilateral cochlear implantation. *Acta Oto-Laryngologica* 138 (2018) 554–561.
48. M.F. Dorman, L. Loisel, J. Stohl, W.A. Yost, A. Spahr, C. Brown, S. Cook: Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear and Hearing* 35 (2014) 633–640.
49. M. Sparreboom, S.A. Ausili, E.A.M. Mylanus: Lateralization of interaural level differences in children with bilateral cochlear implants. *Cochlear Implants International* 23 (2022) 125–133.
50. M. Sparreboom, S. Ausili, J.H. Agterberg Martijn, A.M. Mylanus Emmanuel: Bimodal fitting and bilateral cochlear implants in children with significant residual hearing: the impact of asymmetry in spatial release of masking on localization. *Journal of Speech, Language, and Hearing Research* 64 (2021) 4030–4043.
51. S.A. Ausili, B. Backus, M.J.H. Agterberg, A.J. van Opstal, M.M. van Wanrooij: Sound localization in real-time vocoded cochlear-implant simulations with normal-hearing listeners. *Trends in Hearing* 23 (2019) 2331216519847332.
52. B. Williges, T. Jürgens, H. Hu, M. Dietz: Coherent coding of enhanced interaural cues improves sound localization in noise with bilateral cochlear implants. *Trends in Hearing* 22 (2018) 2331216518781746.
53. H. Kayser, S.D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, B. Kollmeier: Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *EURASIP Journal on Advances in Signal Processing* 2009 (2009) 298605.
54. D. Giannoulis, M. Massberg, J.D. Reiss: Digital dynamic range compressor design – a tutorial and analysis. *Journal of the Audio Engineering Society* 60 (2012) 399–408.
55. B. Vaerenberg, P.J. Govaerts, T. Stainsby, P. Nopp, A. Gault, D. Gnansia: A uniform graphical representation of intensity coding in current-generation cochlear implant systems. *Ear and Hearing* 35 (2014) 533–543.
56. A.W. Archer-Boyd, R.P. Carlyon: Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. *Journal of the Acoustical Society of America* 145 (2019) 1389–1400.
57. R.D. Patterson, I. Nimmo-Smith, J. Holdsworth, P. Rice: An efficient auditory filterbank based on the gammatone function, in: A meeting of the IOC Speech Group on Auditory Modelling at RSRE, 14–15 December 1987.
58. V. Hohmann: Frequency analysis and synthesis using a gammatone filterbank. *Acta Acustica united with Acustica* 88 (2002) 433–442.
59. W. Nogueira, A. Büchner, T. Lenarz, B. Edler: A psychoacoustic, “Nofm”-type speech coding strategy for cochlear implants. *EURASIP Journal on Applied Signal Processing* 18 (2005) 3044–3059.
60. B.A. Swanson: Pitch perception with cochlear implants. Ph.D. Thesis, Faculty of Medicine, Dentistry & Health Sciences, Otolaryngology Eye and Ear Hospital, The University of Melbourne, 2008.
61. H. Hu, M.E. Lutman, S.D. Ewert, G. Li, S. Bleeck: Sparse nonnegative matrix factorization strategy for cochlear implants. *Trends in Hearing* 19 (2015) 2331216515616941.

62. B.A. Swanson: Pitch perception with cochlear implants, Department of Otolaryngology, University of Melbourne, 2008.
63. A. Nobbe, P. Schleich, P. Nopp, C. Zierhofer: Vergleich Einer Feinstruktur-Strategie Mit Der CIS+Strategie Bzgl. Tonhöhenunterscheidung Bei Med-El Benutzern, in Presented at the 9th DGA Jahrestagung, 2006.
64. B.R. Glasberg, B.C.J. Moore: Derivation of auditory filter shapes from notched-noise data. *Hearing Research* 47 (1990) 103–138.
65. A. Nobbe, P. Schleich, C. Zierhofer, P. Nopp: Frequency discrimination with sequential or simultaneous stimulation in Med-El Cochlear implants. *Acta Oto-Laryngologica* 127 (2007) 1266–1272.
66. B.S. Wilson, C.C. Finley, D.T. Lawson, R.D. Wolford, D.K. Eddington, W.M. Rabinowitz: Better speech recognition with cochlear implants. *Nature* 352 (1991) 236–238.
67. I. Hochmair, P. Nopp, C. Jolly, M. Schmidt, H. Schöfer, C. Garnham, I. Anderson: Med-El cochlear implants: state of the art and a glimpse into the future. *Trends in Amplification* 10 (2006) 201–219.
68. Z.M. Smith, A. Kan, H.G. Jones, M. Buhr-Lawler, S.P. Godar, R.Y. Litovsky: Hearing better with interaural time differences and bilateral cochlear implants. *Acoustical Society of America Journal* 135 (2014) 2190.
69. R.J.M. van Hoesel, R.S. Tyler: Speech perception, localization, and lateralization with bilateral cochlear implants. *Journal of the Acoustical Society of America* 113 (2003) 1617–1630.
70. H. Hu, X. Xi, L.L.N. Wong, S. Hochmuth, A. Warzybok, B. Kollmeier: Construction and evaluation of the Mandarin Chinese Matrix (Cmmatrix) sentence test for the assessment of speech recognition in noise. *International Journal of Audiology* 57 (2018) 838–850.
71. D.D. Greenwood: A cochlear frequency-position function for several species – 29 years later. *Journal of the Acoustical Society of America* 87 (1990) 2592–2605.
72. M.L. Hughes, P.J. Abbas: Electrophysiologic channel interaction, electrode pitch ranking, and behavioral threshold in straight versus perimodiolar cochlear implant electrode arrays. *Journal of the Acoustical Society of America* 119 (2006) 1538–1547.
73. J.J. Tsai, K. Koka, D.J. Tollin: Varying overall sound intensity to the two ears impacts interaural level difference discrimination thresholds by single neurons in the lateral superior olive. *Journal of Neurophysiology* 103 (2010) 875–886.
74. M. Dietz, J.-H. Lestang, P. Majdak, R.M. Stern, T. Marquardt, S.D. Ewert, W.M. Hartmann, D.F.M. Goodman: A framework for testing and comparing binaural models. *Hearing Research* 360 (2018) 92–106.
75. J. Raatgever, On the binaural processing of stimuli with different interaural phase relations. Dissertation, Technische Hogeschool Delft, Netherlands, 1980.
76. B. Laback, K. Egger, P. Majdak: Perception and coding of interaural time differences with bilateral cochlear implants. *Hearing Research* 322 (2015) 138–150.
77. A. Ihlefeld, R.P. Carlyon, A. Kan, T.H. Churchill, R.Y. Litovsky: Limitations on monaural and binaural temporal processing in bilateral cochlear implant listeners. *Journal of the Association for Research in Otolaryngology* 16 (2015) 641–652.
78. A. Kan, H.G. Jones, R.Y. Litovsky: Lateralization of interaural timing differences with multi-electrode stimulation in bilateral cochlear-implant users. *Journal of the Acoustical Society of America* 140 (2016) EL392–EL398.
79. K. Egger, P. Majdak, B. Laback: Channel interaction and current level affect across-electrode integration of interaural time differences in bilateral cochlear-implant listeners. *Journal of the Association for Research in Otolaryngology* 17 (2016) 55–67.
80. O.A. Stakhovskaya, M.J. Goupell: Lateralization of interaural level differences with multiple electrode stimulation in bilateral cochlear-implant listeners. *Ear and Hearing* 38 (2017) e22–e38.
81. P.G. Mayo, M.J. Goupell: Acoustic factors affecting interaural level differences for cochlear-implant users. *Journal of the Acoustical Society of America* 147 (2020) EL357–EL362.
82. R.T. Dwyer, C. Chen, P. Hehrmann, N.C. Dwyer, R.H. Gifford: Synchronized automatic gain control in bilateral cochlear implant recipients yields significant benefit in static and dynamic listening conditions. *Trends in Hearing* 25 (2021) 23312165211014139.
83. A. Francl, J.H. McDermott: Deep neural network models of sound localization reveal how perception is adapted to real-world environments. *Nature Human Behaviour* 6 (2022) 111–133.
84. P.-A. Grumiaux, S. Kitić, L. Girin, A. Guérin: A survey of sound source localization with deep learning methods. *Journal of the Acoustical Society of America* 152 (2022) 107–151.
85. H. Hu, M. Dietz: Comparison of interaural electrode pairing methods for bilateral cochlear implants. *Trends in Hearing* 19 (2015) 2331216515617143.
86. M.J. Goupell, J.H. Noble, S.A. Phatak, E. Kolberg, M. Cleary, O.A. Stakhovskaya, K.K. Jensen, M. Hoa, H.J. Kim, J.G.W. Bernstein: Computed-tomography estimates of interaural mismatch in insertion depth and scalar location in bilateral cochlear-implant users. 2021. medRxiv. 2021.02.26.21252533.
87. J.G.W. Bernstein, K.K. Jensen, O.A. Stakhovskaya, J.H. Noble, M. Hoa, H.J. Kim, R. Shih, E. Kolberg, M. Cleary, M.J. Goupell: Interaural place-of-stimulation mismatch estimates using CT scans and binaural perception, but not pitch, are consistent in cochlear-implant users. *Journal of Neuroscience* (2021) JN-RM-0359-21.
88. W.O. Gray, P.G. Mayo, M.J. Goupell, A.D. Brown: Transmission of binaural cues by bilateral cochlear implants: examining the impacts of bilaterally independent spectral peak-picking, pulse timing, and compression. *Trends in Hearing* 25 (2021) 23312165211030411.
89. A. Kan, Z.E. Peng, K. Moua, R.Y. Litovsky: A systematic assessment of a cochlear implant processor's ability to encode interaural time differences, in: 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2018.
90. M. Dietz, B. Backus: Sound Processing for a Bilateral Cochlear Implant System, No. 15173203.9. EU Patent Application Patent. 2015.
91. T. Gajecki, W. Nogueira: The effect of synchronized linked band selection on speech intelligibility of bilateral cochlear implant users. *Hearing Research* 396 (2020) 108051.
92. T. Wendt, S. Van De Par, S.D. Ewert: A computationally-efficient and perceptually-plausible algorithm for binaural room impulse simulation. *Journal of the Audio Engineering Society* 62 (2014) 748–766.
93. C. Kirsch, T. Wendt, S. Par, H. Hu, S. Ewert: Computationally-efficient simulation of late reverberation for inhomogeneous boundary conditions and coupled rooms. *Journal of the Audio Engineering Society* 71 (2022) 186–201.
94. S. van de Par, S.D. Ewert, L. Hladek, C. Kirsch, J. Schütze, J. Llorca-Bofí, G. Grimm, M.M.E. Hendrikse, B. Kollmeier, B.U. Seeber: Auditory-visual scenes for hearing research. *Acta Acustica* 6, sup2 (2022) 55.

95. P. Majdak, F. Zotter, F. Brinkmann, J. De Muynke, M. Mihocic, M. Noisternig: Spatially oriented format for acoustics 2.1: introduction and recent advances. *Journal of the Audio Engineering Society* 70 (2022) 565–584.
96. S.D. Ewert, O. Buttler, H. Hu: Computationally efficient parametric filter approximations for sound-source directivity and head-related impulse responses, in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, 2021:1–6.
97. H. Hu, L. Zhou, H. Ma, Z. Wu: HRTF personalization based on artificial neural network in individual virtual auditory space. *Applied Acoustics* 69 (2008) 163–172.
98. F. Denk, S.M.A. Ernst, S.D. Ewert, B. Kollmeier: Adapting hearing devices to the individual ear acoustics: database and target response correction functions for various device styles. *Trends in Hearing* 22 (2018) 2331216518779313.
99. R.M. Baumgärtel, H. Hu, M. Krawczyk-Becker, D. Marquardt, T. Herzke, G. Coleman, K. Adiloğlu, K. Bomke, K. Plotz, T. Gerkmann, S. Doclo, B. Kollmeier, V. Hohmann, M. Dietz: Comparing binaural pre-processing strategies II: speech intelligibility of bilateral cochlear implant users. *Trends in Hearing* 19 (2015) 2331216515617917.
100. R.M. Baumgärtel, M. Krawczyk-Becker, D. Marquardt, C. Völker, H. Hu, T. Herzke, G. Coleman, K. Adiloğlu, S.M.A. Ernst, T. Gerkmann, S. Doclo, B. Kollmeier, V. Hohmann, M. Dietz: Comparing binaural pre-processing strategies I: instrumental evaluation. *Trends in Hearing* 19 (2015) 2331216515617916.
101. A.E. Vandali, R.J.M. van Hoesel: Development of a temporal fundamental frequency coding strategy for cochlear implants. *Journal of the Acoustical Society of America* 129 (2011) 4023–4036.
102. B.U. Seeber, S. Li: Neural model based sound coding for cochlear implants (Version 0), in *Presented at the 19th International Symposium on Hearing: Psychoacoustics, Physiology of Hearing, and Auditory Modelling, from the Ear to the Brain (ISH2022)*, Lyon, France, 2022.
103. S. Tabibi, A. Kegel, W.K. Lai, N. Dillier: A Bio-Inspired Coding (BIC) strategy for cochlear implants, *Hearing Research* 388 (2020) 107885.
104. M.J. Fumero, A. Eustaquio-Martín, J.M. Gorospe, R. Polo López, M.A. Gutiérrez Revilla, L. Lassaletta, R. Schatzer, P. Nopp, J.S. Stohl, E.A. Lopez-Poveda: A state-of-the-art implementation of a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. *Hearing Research* 409 (2021) 108320.
105. E.A. Lopez-Poveda, A. Eustaquio-Martín, J.S. Stohl, R.D. Wolford, R. Schatzer, B.S. Wilson: A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Ear and Hearing* 37 (2016) e138–e148.
106. E.A. Lopez-Poveda, A. Eustaquio-Martín, F.M. San-Victoriano: Binaural pre-processing for contralateral sound field attenuation and improved speech-in-noise recognition. *Hearing Research* 418 (2022) 108469.
107. T. Potrusil, A. Heshmat, S. Sajedi, C. Wenger, L. Johnson Chacko, R. Glueckert, A. Schrott-Fischer, F. Rattay: Finite element analysis and three-dimensional reconstruction of tonotopically aligned human auditory fiber pathways: a computational environment for modeling electrical stimulation by a cochlear implant based on micro-CT. *Hearing Research* 393 (2020) 108001.
108. A.M. Croner, A. Heshmat, A. Schrott-Fischer, R. Glueckert, W. Hemmert, S. Bai: Effects of degrees of degeneration on the electrical excitation of human spiral ganglion neurons based on a high-resolution computer model. *Frontiers in Neuroscience* 16 (2022) 914876.
109. R.K. Kalkman, J.J. Briare, D.M.T. Dekker, J.H.M. Frijns: The relation between polarity sensitivity and neural degeneration in a computational model of cochlear implant stimulation. *Hearing Research* 415 (2022) 108413.
110. L. Zamaninezhad, V. Hohmann, A. Büchner, M.R. Schädler, T. Jürgens: A physiologically-inspired model reproducing the speech intelligibility benefit in cochlear implant listeners with residual acoustic hearing. *Hearing Research* 344 (2017) 50–61.
111. T. Jürgens, V. Hohmann, A. Büchner, W. Nogueira: The effects of electrical field spatial spread and some cognitive factors on speech-in-noise performance of individual cochlear implant users – a computer model study, *PLOS One* 13 (2018) e0193842.
112. O. Macherey, R.P. Carlyon, A. van Wieringen, J.M. Deeks, J. Wouters: Higher sensitivity of human auditory nerve fibers to positive electrical currents. *Journal of the Association for Research in Otolaryngology* 9 (2008) 241–251.
113. J.A. Undurraga, A. van Wieringen, R.P. Carlyon, O. Macherey, J. Wouters: Polarity effects on neural responses of the electrically stimulated auditory nerve at different cochlear sites. *Hearing Research* 269 (2010) 146–161.
114. M. Takanen, B.U. Seeber: A phenomenological model reproducing temporal response characteristics of an electrically stimulated auditory nerve fiber. *Trends in Hearing* 26 (2022) 23312165221117079.
115. T. Bibberger, S.D. Ewert: Towards a simplified and generalized monaural and binaural auditory model for psychoacoustics and speech intelligibility. *Acta Acustica* 6 (2022) 23.
116. S. Jørgensen, S.D. Ewert, T. Dau: A multi-resolution envelope-power based model for speech intelligibility. *Journal of the Acoustical Society of America* 134 (2013) 436–446.
117. H. Hu, L. Hartog, B. Kollmeier, S.D. Ewert: Spectral and binaural loudness summation of equally loud narrowband signals in single-sided-deafness and bilateral cochlear implant users. *Frontiers in Neuroscience* 16 (2022) 1–18.
118. I. Pieper, M. Mauermann, B. Kollmeier, S.D. Ewert: Toward an individual binaural loudness model for hearing aid fitting and development. *Frontiers in Psychology* 12 (2021) 1–23.
119. I. Pieper, M. Mauermann, B. Kollmeier, S.D. Ewert: Toward an individual binaural loudness model for hearing aid fitting and development. *Frontiers in Psychology* 12 (2021).
120. M.S.A. Zilany, I.C. Bruce, P.C. Nelson, L.H. Carney: A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. *Journal of the Acoustical Society of America* 126 (2009) 2390–2412.
121. M. Imsiecke, B. Krüger, A. Büchner, T. Lenarz, W. Nogueira: Interaction between electric and acoustic stimulation influences speech perception in ipsilateral ears users. *Ear and Hearing* 41 (2020) 868–882.
122. H. Hu, S.A. Ausili, B. Williges, J. Klug, R.C. Felsheim, D. Vickers, M. Dietz, Matlab Code and Results for the Manuscript Submitted to Acta Acustica “a Model Framework for Simulating Spatial Hearing of Bilateral Cochlear Implants Users” (Version 1.0), Zenodo. 2023. <https://doi.org/10.5281/zenodo.7471961>.

Appendix

A.1 AGC

The automatic gain control (AGC) stage compresses the input signal above a defined threshold C_{th} . The implemented AGC was based on Giannoulis et al. [54] and slightly modified in order to fit into the processing chain as part of Jakob Dießel's bachelor thesis. Briefly, the pre-emphasized digital signal $x_0(n)$, where n is the time index at a sampling rate of f_s , is first converted into $x(n)$ in dB SPL. Then the gain $g_c(n)$ is calculated using

$$g_c(n) = \begin{cases} 0, & x(n) < C_{\text{th}} \\ C_{\text{th}} + (x(n) - C_{\text{th}}) \times R - x(n), & x(n) \geq C_{\text{th}} \end{cases}, \quad (\text{A.1})$$

where R is the compression ratio and C_{th} the compression threshold (or knee-point), respectively. In the second step the gain $g_c(n)$ is smoothed using a digital one-pole filter with the following parameters

$$g_s(n) = \begin{cases} \alpha_A \times g_s(n-1) + (1 - \alpha_A) \times g_c(n), & g_c \leq g_s \\ \alpha_R \times g_s(n-1) + (1 - \alpha_R) \times g_c(n), & g_c > g_s \end{cases}. \quad (\text{A.2})$$

Its coefficients α_A and α_R are defined by the “rise time for the step response to go from 10% to 90% to the final value” (Giannoulis et al. [54]) according to

$$\alpha_A = e^{\left(\frac{-\log(9)}{f_s \times T_A}\right)}, \quad (\text{A.3})$$

$$\alpha_R = e^{\left(\frac{-\log(9)}{f_s \times T_R}\right)}. \quad (\text{A.4})$$

Here, T_A is the attack time and T_R the release time. In the last step, the gain is linearized and multiplied with the input signal $x_0(n)$ to generate a digital signal $x_{\text{out}}(n)$ as follows:

$$x_{\text{out}}(n) = 10^{\left(\frac{g_s(n)}{20}\right)} \times x_0(n). \quad (\text{A.5})$$

The default C_{th} , R , T_A , and T_R were 50 dB SPL, 2:1, 50 ms, and 100 ms, respectively. With these parameters, Dießel showed that the implemented compressor leads to results comparable to the literature for both static and moving sources (e.g., [55, 56]) or commercial CI devices. In the current study, two independent AGCs with these default parameters were used to simulate the effect of non-synchronized AGCs on the localization of static spatial sounds generated virtually with the Kayser et al. [53] HRIR database.

Cite this article as: Hu H. Ausili SA. Williges B. Klug J. Felsheim RC, et al. 2023. A model framework for simulating spatial hearing of bilateral cochlear implant users. Acta Acustica, 7, 42.