# Modeling Loudness Perception in Electrical Hearing with a Phenomenological Auditory Nerve Model

Niklas Löcherer, Bernhard U. Seeber

*Audio Information Processing, 80333 Munich, Germany, Email: niklas.loecherer@tum.de*

## Introduction

Horne et al. [4] extended the stochastic leaky integrate and fire (SLIF) neuron, a phenomenological model for electrically stimulated auditory nerve fibers (ANF), to reproduce the input-output functionality as well as temporal properties like latency and jitter of the neurons when being stimulated with both single monophasic and biphasic electrical pulses. To determine the single fiber's response to ongoing pulse trains, the ANF model was further developed to also take into account the recovery and refractory behavior of the neurons and spike-rate adaptation [10, 11]. The current work combines a MATLAB implementation of the ANF model used at our institute with a widely accepted loudness model for electrical stimulation developed by McKay et al. [7, 8]. As their model considers both temporal and spatial loudness summation effects but models the neuronal response only coarsely, we suppose that a combination with our more accurate prediction of the temporal neuronal response pattern can deliver explanations for loudness related phenomena in temporally and spatially more complex stimulation setups involving arbitrary time-variant multi-electrode stimuli. To initially validate the basic functionality of the model fusion, a procedure with a reduced set of variables inspired by Cohen [3] and Werner et al. [14] was used to fit individual loudness growth functions (LGF) including their definition of maximum comfortable level (MCL) based on the firing probability width along the cochlea.

## ANF model structure

The neuronal population was for simplification modeled as a linear array of 300 equally spaced nerve fibers covering a length of 33 mm along the cochlea (≈1 fiber/0.1 mm) Comparable to the work of Werner et al. [14], all fibers were simulated with a set of fixed parameters including the membrane time constant $\tau$ (248.41 µs), latency and jitter function [4] as well as their relative spread, which is defined as the ratio between the standard deviation and mean of the temporally fluctuating threshold membrane potential. A simplified model was used for the spatial spread of excitation (SOE), in which an electrode is assumed to behave as a point source imprinting a current to the neural distribution. As we treat the auditory nerve tissue to act as homogeneous resistive medium [1], the current spread away from the electrode can typically be described by applying a linear attenuation in dB to the electrical potential or the stimulation current depending on the distance of the considered ANF from the site of stimulation. Corresponding to the values fitted by Werner et al. [14], a symmetrical current decay rate of 1 dB/mm was utilized throughout all simulations. Measurements of the radial distances between the electrodes and the inner wall of scala tympani obtained in [3] were used to calculate the peak value of the current weight function.

## The loudness model

A widely accepted approach for calculating loudness in electrical stimulation, which considers temporal as well as spatial interaction effects of pulse train stimuli, is the model by McKay et al. [6, 7, 8]. As outlined by McKay et al. [6], several temporal properties of normal hearing (e.g. masking, temporal resolution) can be described with a phenomenological temporal integration (TI) model, which uses a sliding time window to sum up the weighted peripheral nerve activity. A central loudness decision device uses the integration results as basis for the loudness judgement. The initial model [7] was later extended to also account for spatial loudness summation, i.e. how stimulation at different sites of the cochlea within a certain temporal interval will influence the overall loudness perception [8]. The first step of this model extension consists of building up an excitation density array $E(x,t)$, representing the neural activity at cochlea places $x$ at time instant $t$, which is the output of the previously described sliding temporal integration of the spike responses of the locally stimulated nerve fibers along the cochlea. In our approach, the neural spikes predicted by the ANF model get summed up by a window with an equivalent rectangular duration (ERD) of 7 ms, mathematically described as follows:

$$W(t) = (1-w)\cdot exp\left(\frac{t}{T_{b1}}\right) + w\cdot exp\left(\frac{t}{T_{b2}}\right), t < 0$$

$$W(t) = exp\left(-\frac{t}{T_a}\right), t \geq 0 \qquad (1)$$

The time constants $T_a, T_{b1}$ and $T_{b2}$ as well as the weighting factor $w$ were fitted by [9] to describe the effect of forward and backward masking in normal hearing listeners ($\mathrm{T_a} = 3.5\ \mathrm{ms}, \mathrm{T_{b1}} = 4.6\ \mathrm{ms}, \mathrm{T_{b2}} = 16.6\ \mathrm{ms}$ and $w = 0.17$). The sliding temporal integration effectively corresponds to a convolution of the fibers' responses (spike trains) with the reversed time window $W(-t)$. In the next step the resultant neural excitation density pattern $E(x,t)$ is transformed into channel-specific instantaneous loudness values by grouping and summing up the temporal integration results of the single fibers inside 22 neural subpopulations (channels). This approach is inspired by recent cortical entrainment measurements by Thwaites et al. [12], who reported evidence for a cortical representation of channel-specific instantaneous loudness happening prior to the formation of overall instantaneous loudness. While [8] hypothesized an additional nonlinear specific-loudness transform being applied to the excitation density at each cochlea place, we decided to skip this transformation, as its form and existence remains unclear and has to be investigated more closely in future work. Therefore, comparable to classical acoustic loudness models, like [2, 15], the overall instantaneous loudness at time instant $t$ is simply obtained by integrating the channel-specific instantaneous loudness contributions across all cochlea places $x$. At that point, it should also be noted that [7] and [8] did not describe the excitation density $E$ to be time-dependent, as they argued that any kind of long-term integration of neural excitation, which typically provokes loudness to increase with stimulus duration up to about 100 ms [16], only slightly affected their data recorded at very low pulse rates. In our approach, short-term loudness is calculated as described in [2] by applying a first-order low-pass filter with a cut-off frequency of 8 Hz to the overall instantaneous loudness time signal. Overall loudness is calculated as the 99th percentile short-term loudness.

## Fitting individual loudness growth

A fitting procedure with a reduced set of variables but resembling the approach by Cohen [3] was applied to the modeled neural distribution with the aim of reproducing individual loudness growth data measured in the same study. Differences in shape of the individual LGFs were for simplicity assumed to exclusively arise from discrepancies in the distribution of the fiber threshold currents across the cochlea. As measured by Van den Honert and Stypulkowski [13] and also considered in the fitting procedure of Cohen [3], the fiber threshold (electric current in $\mu A$ to provoke a firing probability of 50%) is assumed to be normally distributed across the nerve fiber array (mean $m$, standard deviation $SD_m$). We found out during our model evaluations, that LGF curvature can be steered by the ratio $SD_m/m$. [13] measured maximum variations in the mean fiber threshold of around 500 μA across preparations when using monopolar intracochlear stimulation in cats and the SD varied between 0.24 and 0.47 of the mean. This means that the fiber threshold is not only stochastically varying with respect to time (μ=104.5 μV, σ=4.595 μV fitted in [4]) but also in the spatial domain. For simplification of the fitting procedure used in the current study, the SOE function was held constant for the different electrodes (E6 = basal, E12 = middle, E18 = apical), which differs from the fitting approach by Werner et al. [14], who also optimized the shape parameters (center point, peak amplitude and decay rate) of the stimulation weight functions. Furthermore, differences in the neural distribution across subjects were not in the scope of this initial validation of the connected models. As the ANF model output as well as the distribution of the fiber thresholds in the spatial dimension and therefore the excitability of all nerve fibers was varied between repeated trials, multiple iterations i.e. combinations of stimulation current, $m$ and $SD_m$ of the spatial fiber threshold distribution were needed to achieve a meaningful estimate of the provoked loudness growth. A definition of maximum comfortable level (MCL) similar to [5, 14] was included in the fitting procedure. Loudness is assumed to be related to the number of fibers being activated by an electrical stimulus, i.e. related to the amount of neural activity. Both [5] and [14] describe MCL as the current needed to provoke a so-called excitation width (EW) of 4 mm along the cochlea. Whereas Kalman et al. [5] defined EW as spatially integrated fiber activity, Werner et al. [14] used the half peak bandwidth (HPBW) of the spatial firing probability distribution. The latter approach was also used in the current study.

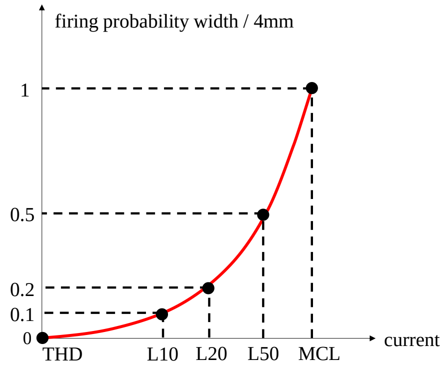


**Figure 1:** Schematic growth of firing probability width (normalized), reference curve for fitting procedure

At every iteration of the fitting process, the simulation result was compared in terms of the produced firing probability width, evaluated at five current levels corresponding to threshold level and 10/20/50/100% MCL. The mean $m$ and standard deviation $SD_m$ of the simulated spatial fiber threshold distribution, which across 30 iterations produced the lowest root-mean-square error (RMSE) from the reference curve in Figure 1, were then used to finally simulate the subject's LGF using the introduced loudness model.

## Results

Figure 2 shows the results of the loudness growth fitting procedure for two exemplary subjects taken from [3], in which psychophysical LGFs were measured using 300-ms biphasic pulse bursts at a pulse rate of 250 pulses/s (25 μs/phase, 25 μs inter-phase gap). For visualization purposes the measured values (dots) were expressed in terms of percent dynamic range (0% = Thd, 100% = MCL) and the output of the TI loudness model was normalized to the maximum of the simulated EW growth (dashed). Furthermore, the plots include the percentage of activated fibers (triangles). A fiber is called "active", if its total probability of emitting at least one spike response exceeds 50%. The model output (diamonds) was fitted with a simple power function (solid), which in both examples shows good agreement with the measured data in terms of the coefficient of determination $R^2$. Both subjects exhibit a different electric dynamic range and therefore steepness of the LGF. These differences arise from the fitted discrepancies in the ratio $SD_m/m$ (S1: 0.25, S3: 0.29). The values fitted by Cohen [3] (S1 E6: m = 637 μA, $SD_m/m$ = 0.2; S3 E12: m = 517 μA, $SD_m/m$ = 0.26) show good accordance with our results. As considered by our fitting procedure, a higher ratio $SD_m/m$ tends to reduce the curvature and therefore increase the integrated area under the LGF. In addition to the trends observed in the growth of EW and amount of activated fibers, the combined spatial and temporal integration mechanism in the loudness model also predicts the power law characteristic of LGF measurements.

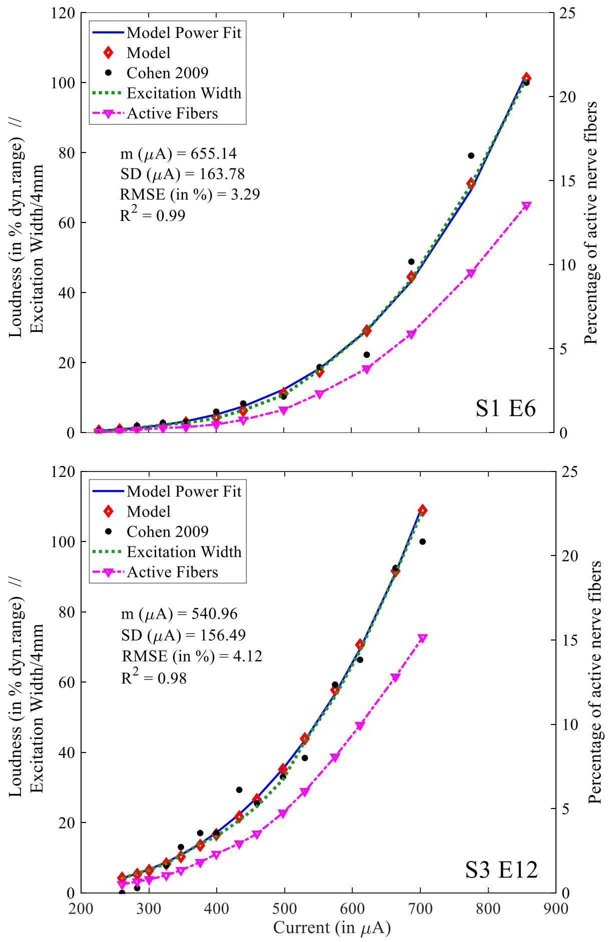


**Figure 2:** Loudness growth for two subjects from [3] (E6 = basal electrode, E12 = middle electrode)

## Conclusion

The investigated combination of a phenomenological ANF model [4, 10, 11] and a loudness model for electrical stimulation [7, 8] succeeded in modelling large individual differences observed in psychophysical LGFs [3]. A fitting procedure for the mean and SD of a normal distribution of fiber thresholds across the cochlea was sufficient to achieve accordance with the measurements and to implicitly produce the power law characteristic of loudness growth observed in electrical hearing without introducing any additional explanatory nonlinearities within the excitation-to-loudness transformation. In future work, the combination of both models will be further examined for its capability to explain temporal and spatial interaction effects and the influence of various electrical pulse parameters on the loudness perception in electrical stimulation.

## References


[1] Bruce, I. C., White, M. W., Irlicht, L. S., O'Leary, S. J., Dynes, S., Javel, E., Clark, G. M.: A stochastic model of the electrically stimulated auditory nerve: Single-pulse response. IEEE Transactions on Biomedical Engineering, 46(6), 617-629, 1999.

[2] Chalupper, J., Fastl, H.: Dynamic Loudness Model (DLM) for Normal and Hearing-Impaired Listeners. Acta Acustica, 88(3), 378-386, 2002.

[3] Cohen, L.T.: Practical model description of peripheral neural excitation in cochlear implant recipients: A paper series. Hearing Research, 247-248, 87-121, 2009.

[4] Horne, C. D. F., Sumner, C. J. and Seeber B. U.: A Phenomenological Model of the Electrically Stimulated Auditory Nerve Fiber: Temporal and Biphasic Response Properties. Frontiers in Computational Neuroscience, 10, 8, 2016.

[5] Kalkman, Randy K., Briaire, Jeroen J. and Frijns, Johan H. M.: Current focussing in cochlear implants: An analysis of neural recruitment in a computational model. Hearing Research, 322, 89-98, 2015.

[6] McKay, C. M., Lim, H. H. and Lenarz, T.: Temporal Processing in the Auditory System. J Assoc Res Otolaryngol, 14(1), 103-124, 2013.

[7] McKay, C. M. and McDermott, H. J.: Loudness perception with pulsatile electrical stimulation: The effect of interpulse intervals. J Acoust Soc Am, 104(2), 1061-1074, 1998.

[8] McKay, C. M., Remine, M. D. and McDermott, H. J.: Loudness summation for pulsatile electrical stimulation of the cochlea: Effect of rate, electrode separation, level, and mode of stimulation. J Acoust Soc Am, 110(3), 1514-1524, 2001.

[9] Oxenham, A. J.: Forward masking: adaptation or integration? J Acoust Soc Am, 109(2), 732-741, 2001.

[10] Takanen, M. and Seeber, B. U.: A Phenomenological Model for Predicting Responses of Electrically Stimulated Auditory Nerve Fiber to Ongoing Pulsatile Stimulation. Conference on Implantable Auditory Prostheses (CIAP), 91, 2017.

[11] Takanen, M., Weller, J. N. and Seeber, B. U.: Modeling Refractoriness In Phenomenological Models of Electrically Stimulated Auditory Nerve Fibers. Fortschritte der Akustik – DAGA ’17, 468-470, 2017.

[12] Thwaites, A., Schlittenlacher, J., Nimmo-Smith, I., Marslen-Wilson, W. D. and Moore, B. C. J.: Tonotopic representation of loudness in the human cortex. Hearing Research, 344, 244-254, 2017.

[13] Van den Honert, C., Stypulkowski, P. H.: Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve. Hearing research, 29(2-3), 195-206, 1987.

[14] Werner, K., Leibold, C., Seeber, B. U.: Individual Fitting of an Auditory Nerve Model for CI Users: Spatial and Temporal Responses, DAGA '18, 2018.

[15] Zwicker, E. and Scharf, B.: A model of loudness summation. Psychological Review, 72(1), 3-26, 1965.

[16] Zwislocki, J. J.: Temporal Summation of Loudness: An Analysis. J Acoust Soc Am, 46(2B), 431-441, 1969.