Neurophysiological insights into hearing aid amplification schemes

Faheem Dinath[1] and Ian C. Bruce[1,2]

[1]School of Biomedical Engineering and [2]Department of Electrical & Computer Engineering
McMaster University, Hamilton, ON, Canada
ibruce@ieee.org

McMaster University

AUDITORY ENGINEERING LABORATORY

Abstract

Phonemic compression schemes for hearing aids have thus far been developed and evaluated based on perceptual criteria such as speech intelligibility, sound comfort, and loudness equalization. Finding compression parameters that optimize all of these perceptual metrics has proved difficult. The goal of this study was to find optimal single-band gain adjustments based on the response of auditory-nerve fibers to speech. Sentences from the TIMIT database were processed by either the NAL-R or the DSL amplification scheme, and deviations from these linear prescriptions were obtained by adjusting the overall gain from 40 dB below to 40 dB above the prescribed gains in 5 dB steps. Neural responses were obtained using the cat auditory-periphery model of Zilany and Bruce [1,2]. Sentences were analyzed on a phoneme by phoneme basis to find the gain adjustment that minimized the difference in neural response to the amplified phoneme in the impaired model and the unamplified phoneme in the normal model. The optimal gain adjustments were found to depend on whether the error metric included the spike timing information of the neural responses (i.e., a time-resolution of several microseconds) or just the mean firing rates (i.e., a time-resolution of several milliseconds). To optimize the mean firing rates, gain adjustments on the order of $+10$ dB were required above the prescribed linear gains in general. In contrast, gain adjustments on the order of -10 dB or more below the prescribed linear gains tend to optimize the responses including spike timing information. Wide dynamic range compression appears to be more beneficial in optimizing the spike timing information than the mean rate information. These results motivate the development of novel nonlinear amplification schemes that simultaneously optimize both spike-timing and mean-rate neural representations.

I. INTRODUCTION

- Amplification prescriptions have a foundation in early empirical studies showing that the most comfortable gain at a particular frequency equals approximately half the hearing threshold at the same frequency. This is referred to as the "half-gain" rule. That is, for every 1 dB increase in hearing threshold, the most comfortable gain is increased by 0.5 dB [3].
- Popular linear hearing aid prescriptions, including the NAL-R (National Acoustic Laboratories) and DSL (Desired Sensation Level) are based on modifications of the half-gain rule and on judgments of speech intelligibility, sound comfort, and loudness equalization [4].
- The goal of this study is to find optimal single-band gain adjustments around the NAL-R and DSL prescribed gains by using the neural representation of speech rather than using perceptual feedback. It is sensible to use single-band gain adjustment, as opposed to multi-gain adjustment, because this reduces the number of working variables and therefore reduces the complexity of the adjustment. We will start by describing the auditory-periphery model in this study and show how it computes the speech neurogram. This is followed by a detailed description of the gain optimization strategy and discussion of the results.

II. METHODS

A. Models

- The auditory-periphery model used in this study was the cat auditory nerve model developed by Zilany and Bruce [1,2]. The model describes the auditory pathway from the middle ear through to the auditory nerve.
- The outer ear is modelled after a head-related transfer function described by Wiener and Ross [5].

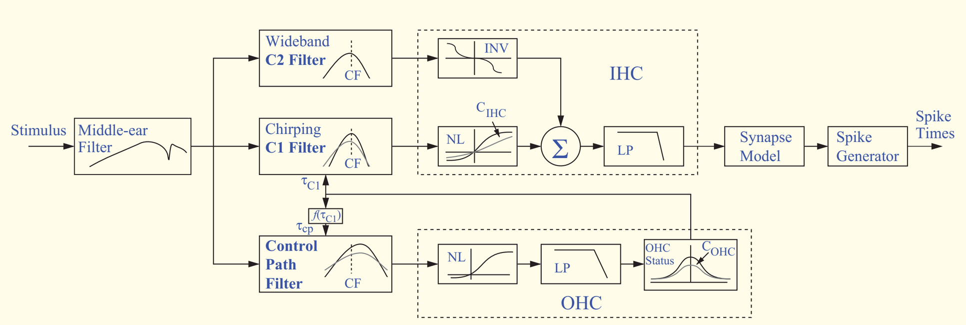

Figure 1: *Zilany and Bruce cat auditory nerve model*

- Input to the middle ear consist of speech waveforms with instantaneous pressures in units of Pascal, sampled at a rate of 500 kHz.
- A high sampling rate is necessary to ensure stability in the middle ear filter and proper statistics in the neural spike generator. In response, the model derives the spike timing information for an auditory nerve fiber with a specific characteristic frequency (CF).
- Model parameters $C_{\rm IHC}$ and $C_{\rm OHC}$, which control the level of inner and outer hair cell impairment, respectively, can be adjusted to provide a desired hearing threshold shift at a specific CF. A $C_{\rm IHC}$ or $C_{\rm OHC}$ of 0 produces full impairment whereas 1 produces normal function.

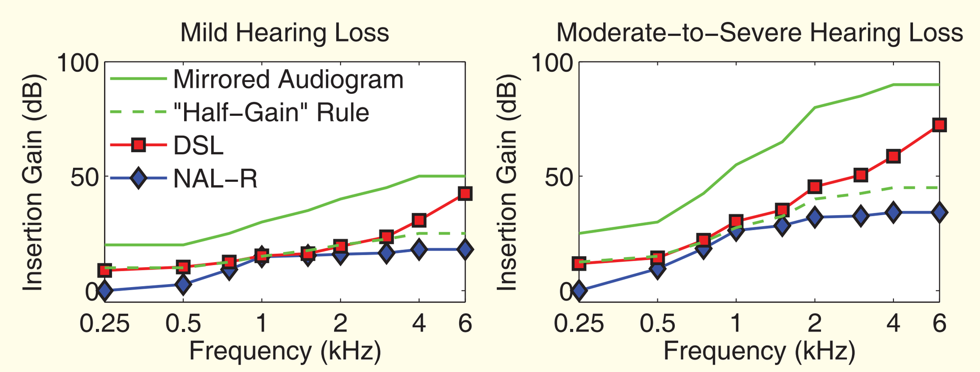

Figure 2: *The two example hearing loss profiles and corresponding insertion gains used in this poster. Left panel: a mild high-frequency hearing loss; Right panel: a moderate-to-severe high-frequency hearing loss. The mirrored audiogram is shown in solid green, and its half gain as the dotted green line. Insertion gains prescribed by DSL are shown in red and those for NAL-R are shown in blue.*

- The goal of the NAL-R linear amplification prescription is to maximize speech intelligibility for moderate SPLs by equalizing perceived loudness over the frequency range important for speech (250–8000 Hz) [4].
- Gain prescribed by NAL-R is in terms of insertion gain (IG), that is, the gain provided by the hearing aid above the gain normally supplied by the outer ear's natural amplification. [6]
- The Desired Sensation Level prescription differs from the NAL-R procedure in that it does not try to make speech equally loud, but rather comfortably loud.
- Although first developed for use in pediatric audiology, ongoing research and modifications has expanded the role of DSL for use with adults [7].
- Gain provided by DSL is expressed in terms of the real ear aided gain (REAG), that is, the total gain supplied by the hearing aid.

B. Stimuli

- Speech recordings were taken from the TIMIT corpus of prompted utterances. The corpus consists of 450 phonetically-compact and 1890 phonetically-diverse read speech sentences in a wide variety of American dialects.
- For consistency and good sound pressure level coverage, speech sentences were normalized to 45, 65, or 85 dB SPL before being presented to the model.
- The neural representation of speech in the auditory nerve is visualized by a "neurogram". A neurogram is similar to the spectrogram, except that it displays the neural response as a function of CF and time.
- The neurogram can include the spike timing information of the neural responses by maintaining a small time bin size (Fig 3D), or the spike timing information can be excluded by computing the moving average with a window of several milliseconds to give only the average discharge rate as a function of time (Fig 3C).
- In this study, 30 CFs spaced logarithmically between 250 and 8000 Hz were chosen. The neural response at each CF is composed of 50 AN fiber responses. In accordance to Liberman and Kiang [8], 60% of fibers were chosen to be high spontaneous rate (>18 spikes/s), 20% medium (0.5 to 18 spikes/s), and 20% low (<0.5 spikes/s).

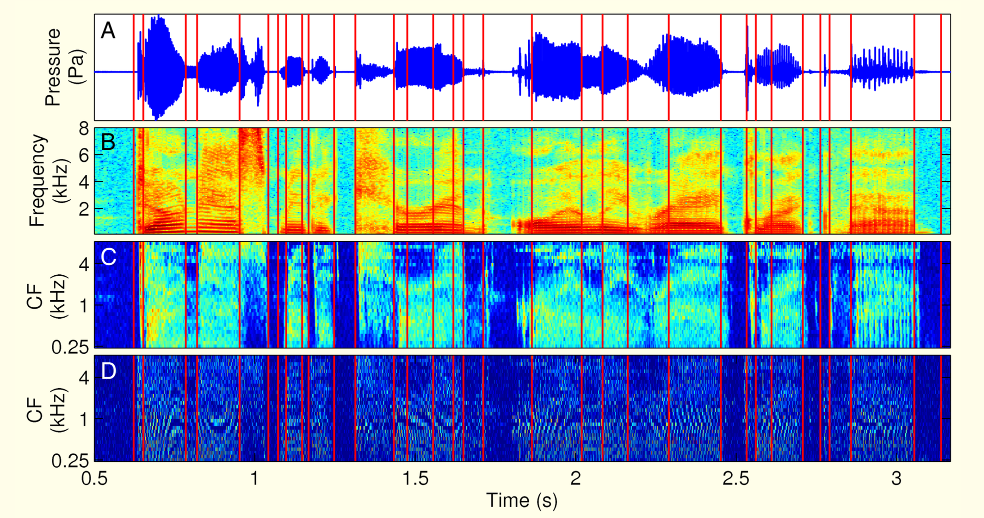

Figure 3: *An example sentence from the TIMIT database and the corresponding spectrogram and neurograms. (A) Time-domain pressure waveform; (B) Spectrogram; (C) Neurogram based on the average discharge rate; (D) Neurogram based on the spiking timing information. Phoneme boundaries are indicated by the vertical red lines.*

C. Gain optimization strategy

- Optimal single-band gain adjustments around the hearing aid prescription gains were obtained though the gain adjustment strategy shown in Fig. 4 below.
- The gain adjustment strategy compares neural responses to speech sentences on a phoneme-by-phoneme basis for the impaired and normal models. In order to avoid the confounding and complicating effects of compression attack and release times, a constant gain adjustment was applied for the duration of each phone, using the known phone boundaries from the TIMIT transcriptions.

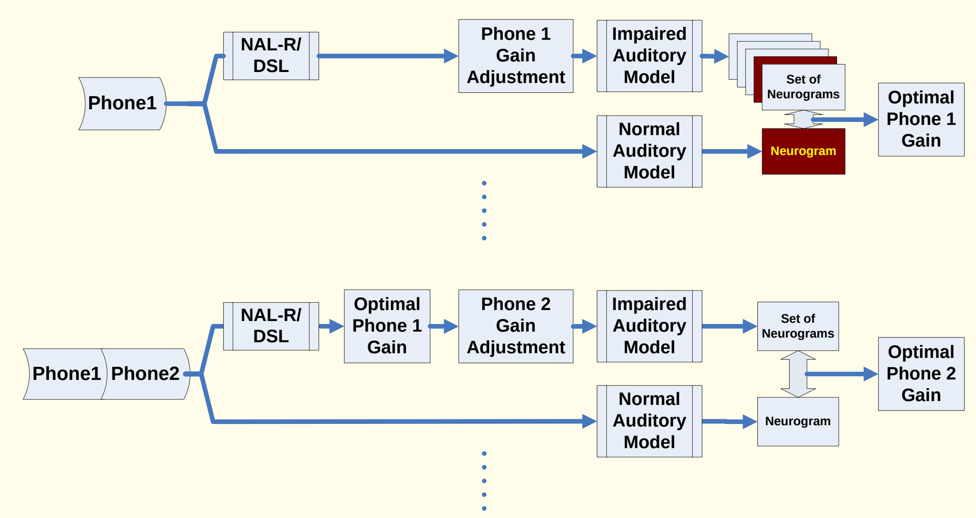

Figure 4: *Flow diagram of gain adjustment strategy.*

- The strategy begins by passing the first phone through the normal model to derive the normal neurogram. In the impaired pathway, the phone is passed though either the NAL-R or DSL amplification prescription before a single-band gain adjustment is applied. Gain adjustments range from -40 to $+40$ dB in 5-dB increments resulting in 17 uniquely amplified phones. The phones are passed through the impaired model, producing a set of 17 neurograms. The gain adjustment that minimizes the mean absolute error between the normal and impaired neurograms is deemed the optimal gain adjustment for that phone.
- For each amplification prescription, optimal gain adjustments were found by comparing either the neurograms with spike timing information or the average discharge rate neurograms.
- The second and all subsequent phones are analyzed in the same manner as the first, however, due to adaptation in the auditory-periphery model, all prior phones are prepended. The range of gain adjustments is applied only to the current phone and all previous phones are amplified with their optimal gain adjustments.

III. RESULTS

- The figures below were generated using 4 test sentences from the TIMIT database. Two sentences were presented to the auditory model having mild hearing impairment and the remaining 2 sentences were presented to a model with mild-to-severe hearing impairment. Each sentence was delivered to the gain optimization strategy at 3 different sound pressure levels, thereby providing a number of diverse phone types and sound pressure levels for examination.

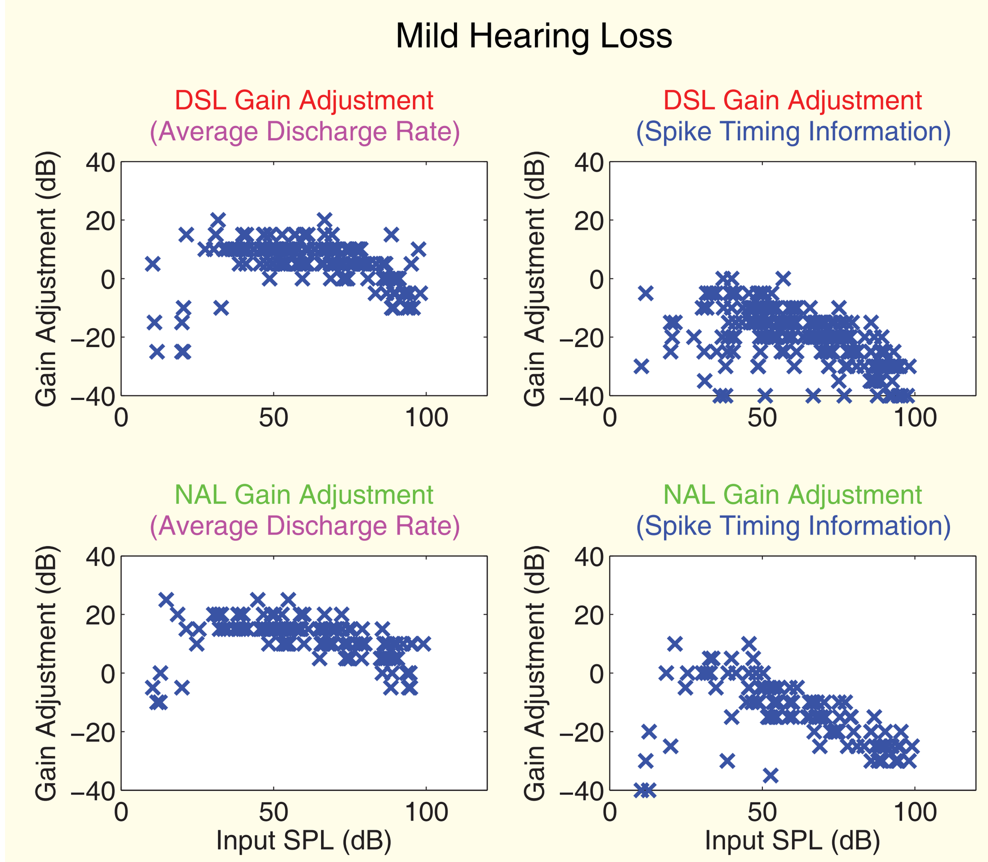

Figure 5: *Optimal gain adjustments versus phoneme input sound pressure for the case of mild hearing loss.*

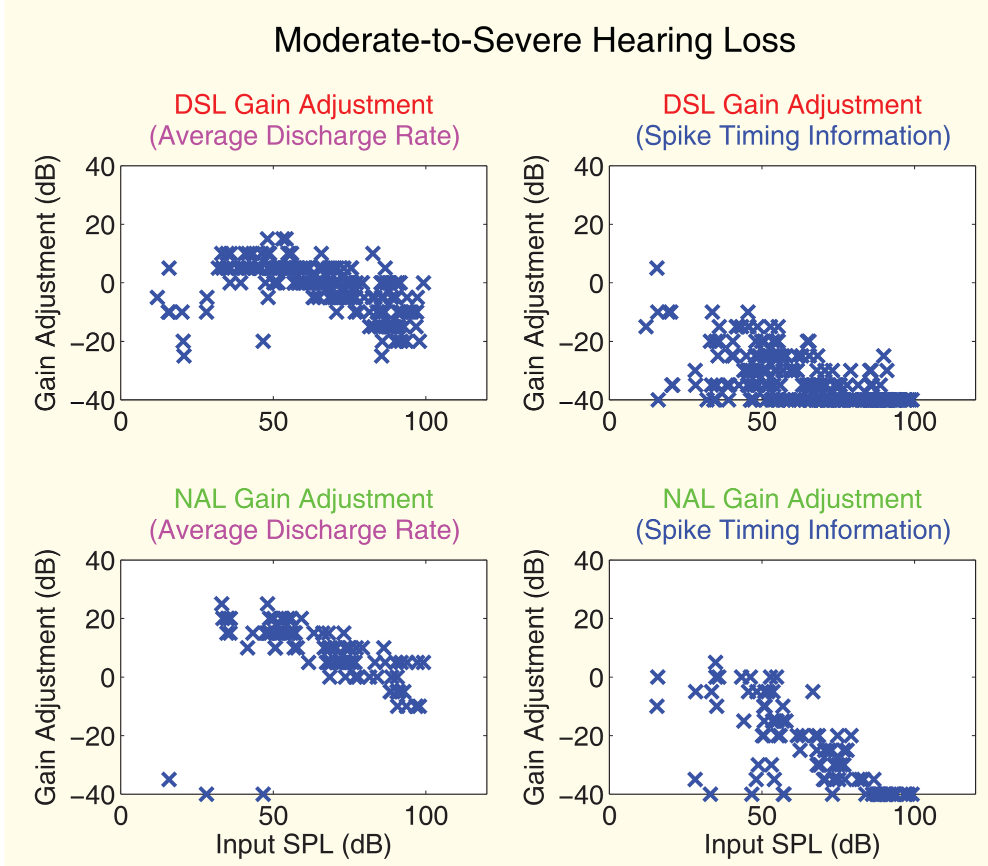

Figure 6: *Optimal gain adjustments versus phoneme input sound pressure for the case of moderate-to-severe hearing loss.*

IV. CONCLUSIONS AND FUTURE WORKS

- The results indicate that the NAL-R and DSL amplifications schemes tend find a balance between optimizing the spike timing information and average discharge rate information.
- It was found that positive gain adjustments above the prescribed gains better restored the mean discharge rate representation of speech. This is consistent with the physiological data of Heinz and Young [9], where they found that on average there is no steepening of auditory nerve fiber rate-level curves with hearing impairment. Consequently, their data would argue for hearing aid gains closer to mirroring the audiogram to restore mean discharge rates.
- The optimal gain adjustments for NAL-R were generally higher than those for DSL, consistent with the lower insertion gains of the NAL-R prescription relative to DSL (see Fig. 2).
- Wide dynamic range compression appears to be required more to optimize the spike timing information than the average discharge rate information.
- It appears that linear amplification schemes or standard single-band compression schemes cannot simultaneously optimize both the spike timing information and the average discharge rate information in the neural response to speech. This motivates: a) studies to further understand why the spike timing and average rate information are optimized at different levels of gain, and b) development of alternative nonlinear amplification strategies to produce simultaneous optimization of both forms of neural coding.

V. ACKNOWLEDGMENTS

The authors thank M. S. A. Zilany for computer programming and technical assistance. Funded by NSERC Discovery Grant 261736 & the Barber-Gennum Chair Endowment (I. C. Bruce) and an NSERC Undergraduate Summer Research Award (T. J. Zeyl).

REFERENCES

[1] M. S. A. Zilany and I. C. Bruce, "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," *J. Acoust. Soc. Am.*, vol. 120, no. 3, pp. 1446–1466, Sept. 2006.

[2] ——, "Representation of the vowel /ε/ in normal and impaired auditory nerve fibers: Model predictions of responses in cats," *J. Acoust. Soc. Am.*, vol. 122, no. 1, pp. 402–417, July 2007.

[3] S. F. Lybarger, "Selective amplification–a review and evaluation," *J. Am. Audiol. Soc.*, vol. 3, no. 6, pp. 258–266, May–Jun 1978.

[4] H. Dillon, *Hearing Aids.* New York: Thieme Medical Publishers, 2001.

[5] F. M. Wiener and D. A. Ross, "The pressure distribution in the auditory canal in a progressive sound field," *J. Acoust. Soc. Am.*, vol. 18, no. 2, pp. 401–408, Oct. 1946.

[6] D. Byrne and H. Dillon, "The national acoustic laboratories' (nal) new procedure for selecting the gain and frequency response of a hearing aid," *Ear Hearing*, vol. 7, pp. 257–265, 1986.

[7] S. Scollie, R. Seewald, L. Cornelisse, S. Moodie, M. Bagatto, D. Laurnagaray, S. Beaulac, and J. Pumford, "The desired sensation level multistage input/output algorithm," *Trends in Amplification*, vol. 9, no. 4, pp. 159–197, 2005.

[8] M. C. Liberman, "Auditory nerve response from cats raised in a low noise chamber," *J. Acoust. Soc. Am.*, vol. 63, pp. 442–455, 1978.

[9] M. G. Heinz and E. D. Young, "Response growth with sound level in auditory-nerve fibers after noise-induced hearing loss," *J. Neurophysiol.*, vol. 91, pp. 784–795, Feb. 2004.