DERIVATIVE-FREE OPTIMIZATION OF HEARING AID PARAMETERS

Shu-Hsien Chu, Mingyi Hong, Zhi-Quan Luo

University of Minnesota
Dept. of Electrical and Computer Engineering
Minneapolis, MN 55455 USA

Kelly Fitz, Martin McKinney, Tao Zhang

Starkey Hearing Technologies
Signal Processing Research Department
Eden Prairie, MN 55344, USA

ABSTRACT

Loudness restoration approaches to hearing aid fitting prescribe gain and compression so as to restore the loudness perceived by a hearing-impaired listener to that perceived by a listener with normal-hearing. Restoring the loudness perception to normal is complicated by the spread of excitation at high stimulus levels that causes intense stimuli at low frequencies to be "heard" and to contribute to the perceived loudness at high frequencies, producing excess loudness growth and poor sound quality. We apply derivative-free optimization algorithms to find a configuration of hearing aid gain and compression parameters that restores specific loudness perception of a hearing impaired listener to that of a normal hearing listener, while simultaneously minimizing the across-frequency spreading of excitation, and ensuring the feasibility of the resulting hearing aid parameters.

Index Terms— Hearing aids, loudness, dynamic range compression, spread of excitation, derivative-free optimization

1. INTRODUCTION

Modern digital hearing aids apply *dynamic range compression*, a form of automatic gain control, to treat the abnormal perception of loudness that is typical of sensorineural hearing loss. By amplifying quiet sounds more than loud sounds, compressive amplification can provide audibility and comfort for patients with hearing loss over a wider dynamic range than linear amplification [1]. In *multichannel compression*, the signal is filtered into several frequency channels,and compression is applied independently to the signal in each channel, allowing compression to be prescribed differently in each channel according to the patient's hearing loss. Various strategies have been proposed for configuring multichannel hearing aid compressors to compensate for hearing loss [2, 3, 4, 5]. *Loudness restoration* approaches prescribe compression intended to restore the loudness perceived by a hearing-impaired listener to the loudness that would be perceived by a normal-hearing listener [6]. Not only the overall loudness, but also the *specific loudness*, the loudness density as a function of frequency, is restored for a range of stimulus spectra. The loudness restoration principle is a component of several hearing aid fitting rationales, and is the foundation of the Cambridge Loudness Restoration (CAMREST) algorithm [6].

At high stimulus levels, significant upward spread of excitation causes intense stimuli at low frequencies to be "heard" and to contribute to the perceived loudness at high frequencies. This may produce excess loudness and loudness growth at high frequencies, and may also produce distortion and poor sound quality, because of the abnormal response of the cochlea to off-frequency stimulation [7]. Our objective in this work is to compute hearing aid compressor parameters that restore loudness perception to normal while simultaneously minimizing the spread of excitation caused by excessive amplification.

Moore [6] used a model of loudness perception to iteratively adjust hearing aid gains to restore loudness for speech spectra at two different levels, and the compressor parameters were inferred from the resulting gain profiles. More recently, Jepsen and Nordahn [8] used an excitation pattern (closely related to specific loudness) model to devise hearing aid gain and compression parameters predicted maximize speech intelligibility. Genetic algorithms have also been employed to search for perceptually optimal hearing aid parameters settings [9, 10], though not, to our knowledge, to achieve a balance between competing objectives, as in the present work.

We incorporate a loudness model into the cost function of a non-convex optimization problem. The model is used to assess both loudness restoration and spread of excitation. Because the model is non-linear and non-invertible, we employ derivative-free techniques to find optimal hearing aid parameters. We present a two-stage approach that encourages the global optimality of the final solution. Our work builds on the work of Moore, generalizing the loudness restoration approach to any number of training spectra, and admitting any number, variety, and weighting of additional constraints and optimality criteria. The present work complements empirical results recently presented by the authors [11] by providing details of the formulation of the optimization problem, and of the two-stage algorithm of deriving optimal hearing aid gain and compression.

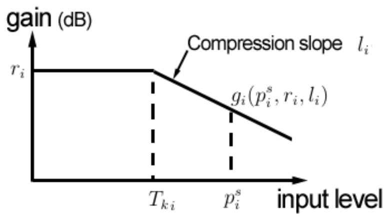

Fig. 1. Compression gain of channel i is determined by the input level p and parameters r and l.

2. PROBLEM FORMULATION

2.1. Hearing aid compression

The hearing aid is modeled as a compressor with $\mathcal{C}$ non-overlapping channels. Within each channel i, the gain, g_i, applied to a stimulus spectrum, s, is computed from the spectrum level in the channel, p_i^s, according to

$$g_i(p_i^s, r_i, l_i) = \begin{cases} r_i, & \text{if } p_i^s \leq T_i \\ r_i + l_i(p_i^s - T_i), & \text{if } p_i^s > T_i \end{cases} \quad (1)$$

where r_i is the maximum channel gain, l_i is the compression slope, and T_i is the compression threshold, a constant level, below which the applied gain in the channel is fixed at r_i. The parameters r_i and l_i are bounded within the feasible ranges $[LB^r \; UB^r]$ and $[LB^l \; UB^l]$ respectively. The input-gain function is depicted in Figure 1. The $\mathcal{C}$ by 1 vectors $\boldsymbol{g}$, $\boldsymbol{p}^s$, $\boldsymbol{r}$, $\boldsymbol{l}$ denote the collection of g_i, p_i^s, r_i, r_i from each channel i, respectively, and the problem solution consists of the optimal sets $\boldsymbol{r}$ and $\boldsymbol{l}$.

2.2. Restoration of loudness

We use the loudness model developed by Moore et al. [12, 13] to estimate the specific loudness for listeners with hearing impairment. This model predicts the loudness deficit due to hearing impairment, as well as the spread of excitation at high stimulus levels. Hearing loss is described by the loss of sensitivity to pure tones of various frequencies relative to the sensitivity of a listener having "normal" hearing (no impairment). All predictions of loudness are implicitly dependent on the hearing loss specification, but for brevity, we omit this dependency in our notation.

Since our objective is to find the single set of parameters $(\boldsymbol{r}, \boldsymbol{l})$ that restore loudness for a variety of sounds, we select a variety of stimulus spectra, describing, for example, the long-term average spectra of speech at different levels of vocal effort, and optimize over this stimulus set.

For a single stimulus spectrum $\boldsymbol{p}^s$, the amplified output of hearing aid is described by $\boldsymbol{p}^s + \boldsymbol{g}(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l})$, where the gain $\boldsymbol{g}(\cdot)$ is determined from the spectrum level $\boldsymbol{p}^s$ and the compression parameters $\boldsymbol{r}$ and $\boldsymbol{l}$, as described by Equation 1. Let $\bar{\boldsymbol{\nu}}(\boldsymbol{p}^s)$ be the loudness predicted for a listener with normal hearing in response to $\boldsymbol{p}^s$, and $\boldsymbol{\nu}(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l})$ be the loudness predicted for a listener with hearing loss for the amplified spectrum $\boldsymbol{p}^s + \boldsymbol{g}(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l})$. Optimal loudness restoration is achieved by minimizing the average difference between aided impaired loudness and normal loudness, defined

$$D(\boldsymbol{r}, \boldsymbol{l}) = \frac{1}{\mathcal{S}} \sum_{s=1}^{\mathcal{S}} \|(\boldsymbol{\nu}(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l}) - \bar{\boldsymbol{\nu}}(\boldsymbol{p}^s))\|_1 \quad (2)$$

2.3. Spread of excitation

The loudness induced at excitation frequency f by amplification in channel i can be characterized by the accumulated variation in loudness with respect to the parameters $(\boldsymbol{r}, \boldsymbol{l})$ as

$$\begin{aligned} & \Lambda_{i,f}(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l}) \quad (3) \\ \equiv & \int_{l_i}^{0} \int_{0}^{r_i} |\frac{\partial \nu_f}{\partial r_i}(\boldsymbol{p}^s, \tilde{\boldsymbol{r}}_i, \boldsymbol{l})| + |\frac{\partial \nu_f}{\partial l_i}(\boldsymbol{p}^s, \boldsymbol{r}, \tilde{\boldsymbol{l}}_i)| d\tilde{r}_i d\tilde{l}_i \end{aligned}$$

where $\tilde{\boldsymbol{r}}_i = [r_1, \ldots, r_{i-1}, \tilde{r}_i, r_{i+1} \ldots, r_{\mathcal{C}}]$, all channels but i invariant, and similarly $\tilde{\boldsymbol{l}}_i = [l_1, \ldots, l_{i-1}, \tilde{l}_i, l_{i+1} \ldots, l_{\mathcal{C}}]$.

Equation 3 is not separable with respect to channel, since $\boldsymbol{\nu}$ is not separable. However, $\boldsymbol{\nu}$ is monotonically increasing with respect to spectrum level, and $\boldsymbol{g}(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l})$ is also monotonically increasing w.r.t the parameters $\boldsymbol{r}$ and $-\boldsymbol{l}$. If we further require that $g_i(p_i^s, r_i, l_i) = 0$, for $r_i = 0$, then the integration can be simplified to

$$\begin{aligned} & \Lambda_{i,f}(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l}) \quad (4) \\ = & [\nu_f(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l}) - \nu_f(\boldsymbol{p}^s, \tilde{\boldsymbol{r}}_i, \boldsymbol{l} | \tilde{r}_i = 0)] \\ & + [\nu_f(\boldsymbol{p}^s, \tilde{\boldsymbol{r}}_i, \tilde{\boldsymbol{l}}_i | \tilde{r}_i = 0, \tilde{l}_i = 0) - \nu_f(\boldsymbol{p}^s, \tilde{\boldsymbol{r}}_i, \boldsymbol{l} | \tilde{r}_i = 0)] \\ = & \nu_f(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l}) - \nu_f(\boldsymbol{p}^s, \tilde{\boldsymbol{r}}_i, \tilde{\boldsymbol{l}}_i | \tilde{r}_i = 0, \tilde{l}_i = 0) \end{aligned}$$

that is, the loudness difference between parameter configuration $(\boldsymbol{r}, \boldsymbol{l})$ and the configuration without gain in channel i, i.e., $\tilde{r}_i = 0, \tilde{l}_i = 0$.

We define the non-locality penalty, $L(\boldsymbol{r}, \boldsymbol{l})$, as the frequency-weighted induced loudness combined across channels and averaged over training spectra,

$$L(\boldsymbol{r}, \boldsymbol{l}) \equiv \frac{1}{\mathcal{S}} \sum_{\substack{1 \leq s \leq \mathcal{S} \\ 1 \leq i \leq \mathcal{C} \\ f \in \mathcal{F}_{loud}}} \omega_{i,f} \Lambda_{i,f}(\boldsymbol{p}^s, \boldsymbol{r}, \boldsymbol{l}) \quad (5)$$

where $\omega_{i,f}$ is a frequency distance weighting that places increasing weight on loudness contribution at frequencies more distant from the center frequency of compressor channel i, and $\boldsymbol{\mathcal{F}}_{loud}$ is the set of excitation frequencies at which specific loudness is estimated by the model.

2.4. Problem formulation

We propose a two-stage algorithm that finds compressor parameters $(\boldsymbol{r}, \boldsymbol{l})$ that simultaneously minimize the aided impaired loudness deviation from normal, $D(\boldsymbol{r}, \boldsymbol{l})$ (2), and the spread of excitation, $L(\boldsymbol{r},\boldsymbol{l})$ (5). The problem is formulated

$$\min_{\boldsymbol{r},\boldsymbol{l}} D(\boldsymbol{r},\boldsymbol{l}) + \alpha L(\boldsymbol{r},\boldsymbol{l}). \tag{6}$$

One-dimensional enumeration of the objective function verifies that this problem is indeed non-convex.

3. TWO-STAGE ALGORITHM

The quality of the solution to a non-convex optimization problem often depends critically on the initial point. We therefore propose a two-stage scheme that first solves, in parallel, the particular gain-only subproblems with respect to each training spectrum. From the particular solutions, we can derive a suitable initial point for the universal (compressive) optimization problem (6).

Algorithm 1 Proposed Two-Stage Optimization Scheme

Stage 1: Find particular (gain-only) solutions $\boldsymbol{x}^s$

1: **for all** $p^s \in \mathcal{S}$ **do**

2: Initial point $\boldsymbol{x}_0 \leftarrow \frac{1}{3}$ of hearing loss

3: Find particular solutions to gain-only subproblems

$$\hat{\boldsymbol{x}}^s \leftarrow \arg\min_{\boldsymbol{x}^s} D_G(\boldsymbol{x}^s) + \alpha L_G(\boldsymbol{x}^s)$$

4: **end for**

Stage 2: Find the universal (compressive) solution $(\hat{\boldsymbol{r}},\hat{\boldsymbol{l}})$

5: Initial point $(\boldsymbol{r}_0,\boldsymbol{l}_0) \leftarrow LeastSquare(\hat{\boldsymbol{x}}^s)$

6: Find universal solution to compressive gain problem

$$(\hat{\boldsymbol{r}},\hat{\boldsymbol{l}}) \leftarrow \arg\min_{\boldsymbol{r},\,\boldsymbol{l}} D(\boldsymbol{r},\boldsymbol{l}) + \alpha L(\boldsymbol{r},\boldsymbol{l})$$

7: **return** $(\hat{\boldsymbol{r}},\hat{\boldsymbol{l}})$

3.1. Gain-only particular solutions

We first consider the optimal gain for each training spectrum individually. This is a simpler problem that does not involve compression, but only finds the optimal amplification with respect to a single spectrum. The compression function $\boldsymbol{g}(\boldsymbol{p}^s,\boldsymbol{r},\boldsymbol{l})$ is replaced in the problem formulation of Equation 6 by a single gain parameter $\boldsymbol{x}^s$ that is optimized directly, yielding

$$\begin{aligned}
&\min_{0\le \boldsymbol{x}^s \le UB^r} D_G(\boldsymbol{x}^s) + \alpha L_G(\boldsymbol{x}^s) \\
\text{s.t.}\quad & D_G(\boldsymbol{x}^s) = \|\boldsymbol{\nu}(\boldsymbol{p}^s+\boldsymbol{x}^s) - \bar{\boldsymbol{\nu}}(\boldsymbol{p}^s)\|_1 \\
& L_G(\boldsymbol{x}^s) = \frac{1}{\mathcal{C}} \sum_{\substack{1\le i\le \mathcal{C}\\ f\in \mathcal{F}_{loud}}} \omega_{i,f}\left[\boldsymbol{\nu}(\boldsymbol{p}^s+\boldsymbol{x}^s) - \boldsymbol{\nu}(\boldsymbol{p}^s+\tilde{\boldsymbol{x}}_i^s)\right]
\end{aligned} \tag{7}$$

where $\tilde{\boldsymbol{x}}_i^s = [x_1^s,\ldots,x_{i-1}^s,0,x_{i+1}^s,\ldots,x_{\mathcal{C}}^s]$.

3.2. Compressive universal solution

The particular solutions $\hat{\boldsymbol{x}}^s$ are used in the second stage of the algorithm to determine an initial point for the universal compressive gain problem (Equation 6). The method of least squares fitting is applied to find parameters $(\boldsymbol{r}_0,\boldsymbol{l}_0)$ that best reproduce $\hat{\boldsymbol{x}}^s$ for all $s = 1\ldots\mathcal{S}$.

$$\begin{aligned}
(\boldsymbol{r}_0,\boldsymbol{l}_0) &= \arg\min_{\boldsymbol{r},\boldsymbol{l}} \sum_{s=1}^{\mathcal{S}} \|\boldsymbol{g}(\boldsymbol{p}^s,\boldsymbol{r},\boldsymbol{l}) - \hat{\boldsymbol{x}}^s\|^2 \\
&= \left[\arg\min_{r_i,l_i} \sum_{s=1}^{\mathcal{S}} (g_i(p_i^s,r_i,l_i) - \hat{x}_i^s)^2\right]_{i=1,2,\ldots,\mathcal{C}}
\end{aligned} \tag{8}$$

which is a vector least square problem with components corresponding to channels. If the spectrum levels are always greater than the compression threshold, $p_i^s \ge T_i, \forall i, s$, Equation 8 is a linear least square problem solved by

$$\begin{bmatrix} r_i \\ l_i \end{bmatrix} = (X^T X)^{-1} X^T \begin{bmatrix} x_i^1 \\ x_i^2 \\ \vdots \\ x_i^{\mathcal{S}} \end{bmatrix},\; X = \begin{bmatrix} 1 & (p_i^1 - T_i) \\ 1 & (p_i^2 - T_i) \\ \vdots & \vdots \\ 1 & (p_i^{\mathcal{S}} - T_i) \end{bmatrix}$$

The solution will produce gains $\boldsymbol{g}(\boldsymbol{p}^s,\boldsymbol{r}_0,\boldsymbol{l}_0)$ that exactly match $\hat{\boldsymbol{x}}^s$ only for $\mathcal{S} \le 2$. From this initial point, we apply derivative-free optimization (DFO) [14] to find the optimal compressive solution that balances loudness restoration and spread of excitation according to the problem formulation in Equation 6.

4. RESULTS

We illustrate the performance of our proposed optimization scheme by applying different weighting criteria to the competing objectives for a typical sloping audiogram, describing a hearing loss that is mild (10 dB HL) in the low frequencies, and increases to moderate (65 dB HL) at high frequencies. We applied the proposed two-stage optimization approach to find optimal loudness-restoring compression parameters for three speech-shaped training spectra, plotted in Figure 3.

The specific loudness predictions plotted in Figure 2 illustrate the quality of loudness restoration achieved through this approach. Loudness is plotted for each of the three training spectra, and for several different cost function weights for the spread of excitation penalty, α (see Equation 6). Figure 2 shows that while the match to normal loudness is degraded by increasing the α, the spread of excitation is reduced in Figure 4, as intended. In Figure 2, The average restoration error for $\alpha = 0,\ 0.01,\ 0.05$ are $0.88,\ 0.96,\ 2.29$ and the average weighted spread of excitation are $181.98,\ 94.2,\ 7.4$ respectively. The tradeoff between restoration of normal specific loudness and spread of excitation is, thus, governed by the weighting parameter α.

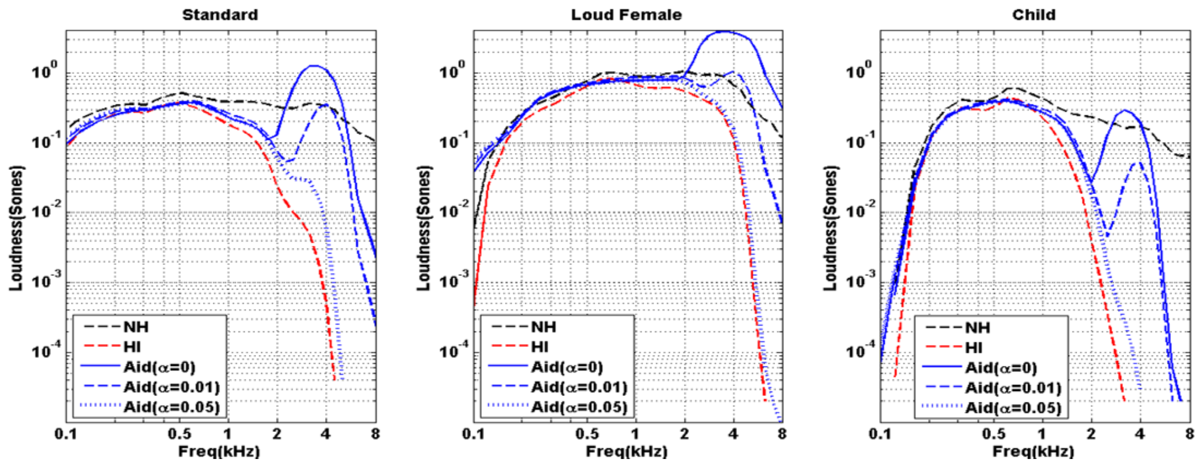

Fig. 2. Loudness Matching

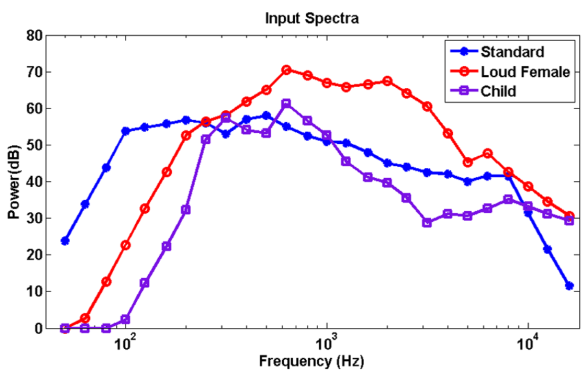

Fig. 3. Long-term average speech spectra used in simulations: 'Standard' male speech at 65 dB SPL, 'Loud Female' speech at 75 dB, and 'Child' speech at 65 dB (spectra derived from [15, 16])

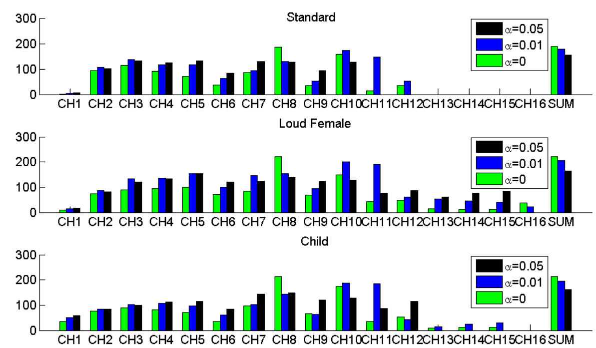

Fig. 4. The amount of weighted nonlocality contributed by each channel for each spectrum. The corresponding weight setups from left to right are $\alpha = 0, 0.01$, and 0.05.

The quality of the final solution is ensured using the two stage approach, which solves a set of simpler optimization problems (Equation 7) to find a suitable initial point for the solving universal problem (Equation 6).

5. CONCLUSION

We have shown that it is feasible to apply non-invertible, non-differentiable perceptual models in a hearing aid fitting rationale. We devised a two-stage optimization approach that guides a derivative-free optimization solver to the neighborhood of a globally optimal solution. By first identifying an initial point that is a nearly optimal for simpler subproblems, we are assured that the solver does not get stuck at a local optimum far from the global optimal solution. Although currently too computationally demanding for clinical application, this approach opens up new avenues for research in hearing fitting strategies.

In this work, we have demonstrated that derivative-free optimization utilizing such a model in the cost function, can be used to balance the competing objectives of restoring normal specific loudness and minimizing spread of excitation. Previously, there has been no way to systematically trade off spread of excitation against loudness restoration, so little is known about how best to balance those objectives, or how to set α to achieve a good hearing aid fitting. We have explored elsewhere [11] the introduction of other cost function, such as a gain non-smoothness penalty, and have also shown that the loudness model can be further tuned to the particular pathology of the patient. Actually, any number and variety of competing constraints all can be incorporated into the optimization by extending the definition of the cost function (Equation 6), without altering the basic structure of the algorithm. This flexibility and expandability are key strengths of our approach.

6. REFERENCES

[1] B. Edwards, "Springer handbook of auditory research," in *Springer Handbook of Auditory Research*, pp. 339–421. Springer-Verlag, New York, 2004.

[2] D. Byrne, H. Dillon, T. Ching, R. Katsch, and G. Keidser, "NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures," *Journal of the American Academy of Audiology*, vol. 12, no. 1, pp. 37–51, Jan. 2001.

[3] B. C. J. Moore and C. Füllgrabe, "Evaluation of the CAMEQ2-HF method for fitting hearing aids with multichannel amplitude compression," *Ear and Hearing*, vol. 31, no. 5, pp. 657–666, Oct. 2010.

[4] B. C. J. Moore, B. R. Glasberg, and M. A. Stone, "Use of a loudness model for hearing aid fitting: III. A general method for deriving initial fittings for hearing aids with multi-channel compression," *British Journal of Audiology*, vol. 33, no. 4, pp. 241–258, Aug. 1999.

[5] 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.

[6] B. C. J. Moore, "Use of a loudness model for hearing aid fitting: IV. Fitting hearing aids with multi-channel compression so as to restore 'normal' loudness for speech at different levels," *British Journal of Audiology*, vol. 34, no. 3, pp. 165–177, 2000.

[7] B. C. J. Moore, "Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids," *Trends in Amplification*, vol. 5, no. 1, pp. 1–34, 2001.

[8] M. L. Jepsen and M. A. Nordahn, "Prediction of optimal gain for speech intelligibility, considering differences in cochlear damage," in *International Hearing Aid Research Conference*, Lake Tahoe, CA, Aug. 2012, Widex A/S, pp. 67–68.

[9] E. A. Durant, G. H. Wakefield, D. J. Van Tasell, and M. E. Rickert, "Efficient perceptual tuning of hearing aids with genetic algorithms," *IEEE Transactions on Speech and Audio Processing*, vol. 12, no. 2, pp. 144–155, 2004.

[10] H. Takagi and M. Ohsaki, "Interactive evolutionary computation-based hearing aid fitting," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 3, pp. 414–427, 2007.

[11] K. Fitz, S. H. Chu, M. Y. Hong, M. McKinney, T. Zhang, and Z. Q. Luo, "Optimization of loudness-restoring hearing aid fittings," in *International Hearing Aid Research Conference*, Lake Tahoe, California, Aug. 2012, Starkey Hearing Technologies, University of Minnesota, p. 53.

[12] B. C. J. Moore, B. R. Glasberg, and T. Baer, "A model for the prediction of thresholds, loudness, and partial loudness," *Journal of the Audio Engineering Society*, vol. 45, no. 4, pp. 224–240, 1997.

[13] American National Standards Institute, American National Standards Institute, New York, *Procedure for the computation of loudness of steady sounds*, 2007, ANSI S3.4-2007.

[14] M. J. D. Powell, "UOBYQA: Unconstrained optimization by quadratic approximation," *Math. Program.*, vol. 92, no. 3, pp. 555–582, 2002.

[15] R. M. Cox and J. N. Moore, "Composite speech spectrum for hearing aid gain prescriptions," *Journal of Speech, Language, and Hearing Research*, vol. 31, no. 1, pp. 102, 1988.

[16] *Verifit user's guide version 3.10*, chapter 19 Speech Map Technical Details, p. 108, Audioscan, June 2012.