2014 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, SEPT. 21–24, 2014, REIMS, FRANCE

A PROBABILISTIC APPROACH TO HEARING LOSS COMPENSATION

Mojtaba Farmani[1]

Bert de Vries[1,2]

[1]Eindhoven University of Technology
Department of Electrical Engineering
Eindhoven, The Netherlands
m.farmani@tue.nl

[2]GN ReSound, Research Dept.
Het Eeuwsel 6, 5612 AS Eindhoven
The Netherlands
bdevries@ieee.org

ABSTRACT

Modern hearing aids use Dynamic Range Compression (DRC) as the primary solution to combat Hearing Loss (HL). Unfortunately, common DRC based solutions to hearing loss are not directly based on a proper mathematical or algorithmic description of the hearing loss problem. In this paper, we propose a probabilistic model for describing hearing loss, and we use Bayesian inference for deriving optimal HL compensation algorithms. We will show that, for a simple specific generative HL model, the inferred HL compensation algorithm corresponds to the classic DRC solution. An advantage to our approach is that it is readily extensible to more complex hearing loss models, which by automated Bayesian inference would yield complex yet optimal hearing loss compensation algorithms.

Index Terms— Dynamic range compression, Bayesian inference, Kalman filter, hearing loss, hearing aids

1. INTRODUCTION

The audible area of acoustic waves for human listeners is limited both in frequency range (from around 20 Hz to 20 kHz) and in power levels [1], see Fig. 1. The lower boundary for normal listeners is often called the *hearing threshold* (HT) and the upper boundary is associated with the *Uncomfortable Loudness* (UCL) level.

Most forms of Hearing Loss (HL) causes the hearing threshold level to be increased in a frequency-dependent way while the UCL levels stay at about the same levels (Fig. 1). As a result, the growth of loudness, i.e. the psychological perception of power levels, is distorted for hearing-impaired listeners. For instance, most hearing impaired patients experience limited ability to hear moderate-level sounds like speech, while the perception of very loud sounds is barely affected. More precisely, an increase of sound level creates a bigger growth of loudness for a hearing-impaired than for a normal hearing person. In the auditory sciences, this phenomenon is called the *recruitment* problem. The separation of loudness recruitment region from normal loudness growth is marked by a *Recruitment Threshold* (R) [2, p. 3]. Incoming sounds with levels greater than R are normally perceived [3].

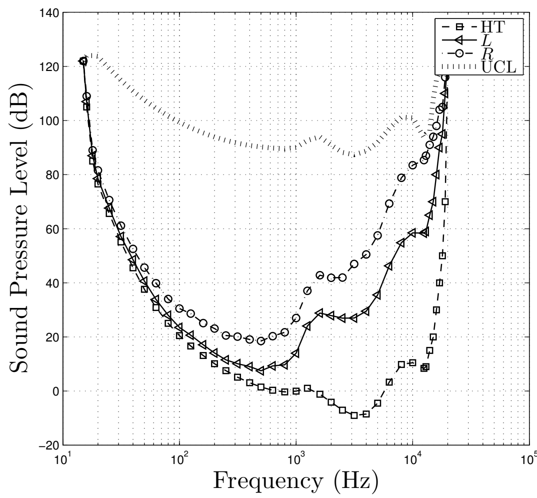

Fig. 1. Hearing thresholds (HT) and uncomfortable loudness (UCL) levels as a function of frequency for normal listeners, and impaired hearing (L) and recruitment thresholds (R) for hearing impaired patients. The area between the HT and UCL curves is the normal audible range. The area between the L and UCL curves correspond to the residual audible range for impaired listeners. In the area between the L and R curves, distorted sound is experienced as a result of recruitment.

As can been seen from Fig. 1, the audible range for a hearing-impaired person, also known as the *residual range*, is smaller than the normal audible range. In order to compensate for hearing loss, hearing aid algorithms implement a mapping function from the normal audible to the residual hearing range [2, p. 287]. In modern hearing aids, this mapping algorithm is realized by a dynamic range compressor (DRC).

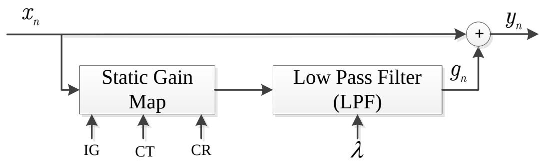

Fig. 2. General architecture of a dynamic range compressor.

1.1. Dynamic range compression

A DRC is a dynamic system that adapts a scalar gain on the basis of changes in the input signal level. Fig. 2 shows the general architecture of a classic DRC, where x_n, g_n and y_n are input, gain and output signal levels in dB, respectively.

While different DRC circuits vary in their details, in this paper we will use a generic DRC circuit as our reference [4]. A DRC consists of both a 'static gain' and a 'dynamic' module. The static gain unit is described by

$$h(x) = \begin{cases} \text{IG} & \text{if } x < \text{CT} \\ \text{IG} - (1 - \frac{1}{\text{CR}})(x - \text{CT}) & \text{otherwise} \end{cases} \quad (1)$$

The parameters of the DRC static unit go by the names 'initial gain' (IG), 'compression threshold' (CT) and 'compression ratio' (CR). The specific choice of static gain parameters relates to the physical interpretation of the compressor.

The dynamic aspects of the DRC are usually realized by a one-pole low-pass filter, as described by

$$g_n = \lambda g_{n-1} + (1 - \lambda) h(x_n), \quad (2)$$

where λ is a 'forgetting factor'. The physical time interval to reach $1 - 1/e$ of the final value in response to a step input is called the time constant (τ) [4]. The forgetting factor is directly related to the time constant by $\lambda = \exp(\frac{-1}{\tau f_s})$, where f_s is the sampling frequency.

The goal of the low-pass filter is to smooth gain changes, since gain changes may lead to audible signal distortion. Commonly, DRC systems in hearing aids react with different time constants to increases ('attack') and decreases ('release') of the input signal level, but in this paper we will assume the same time constants for both attack and release. The DRC output is simply $y_n = g_n + x_n$, where all signals are in dB, cf. Fig. 3 .

1.2. Problem statement and our approach

Unfortunately, in the case of dynamic range compression as a hearing loss compensation method, a direct mathematical link to the underlying hearing loss problem is often not available. For instance, the hearing loss model as described before is completely static while the DRC circuit contains a dynamic module. As a result, there exist no objective methods to estimate time constants for DRC circuits in hearing aids. More

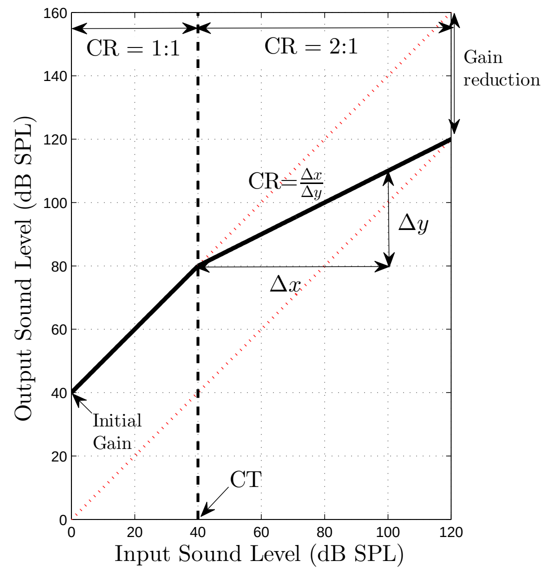

Fig. 3. Static input-output function of a hearing dynamic range compressor determines the relation between the input sound level and the amplified sound level.

generally, in order to evaluate the performance of any proposed hearing loss compensation method, we need a proper mathematical description of the hearing loss problem first. Indeed the literature on a comparative evaluation of DRC algorithms for HL compensation is almost entirely based on subjective testing.

We propose a new framework for hearing loss compensation that is based on a probabilistic model for hearing loss. In this framework, the *solution* to the problem is automatically derived by online Bayesian inference, which for simple hearing loss models can be solved by Kalman filtering. Extending inference to the parameters of the system leads to a principled parameter estimation method and extending inference one more layer up the hierarchy allows for principled DRC model comparison. In contrast to [5], both signal processing and parameter estimation are approached within the same inference framework. In this paper, we report work in progress on the model specification and state estimation stages.

2. MODEL SPECIFICATION

Hearing loss compensation is generally accomplished by independently operating modules in the frequency channels of a filter bank. Since the filter bank structure is not essential for this paper, we will ignore the specifics of the filter bank and refer the reader to a well-working filter bank for hearing aids such as described by Kates et al. [6]. In the following, signals are considered to refer to logarithmic power levels in separate frequency channels.

We consider HL compensation as a hidden state estimation problem. The unobserved state that we would like to estimate is the gain that should be applied to the input sound level in order to so as to compensate for a patient's hearing loss problem.

2.1. State transition model

The state transition model is represented by the conditional probability $p(g_n|g_{n-1})$, where g_n is the gain in dB at the nth time step. Since rapidly changing gains have unfavorable effects on the sound quality [2, p. 305], we assume that the difference between two consecutive gains should be small. This can be modeled by zero-mean Gaussian noise,

$$g_n = g_{n-1} + w_n \tag{3}$$

where n is the discrete time index, g is the provided gain in dB, and w_n is the process noise which is governed by a white Gaussian noise process $\mathcal{N}(0, \sigma_w^2)$, where σ_w^2 is the variance of the process noise.

2.2. Observation model

It is desired that the aided hearing-impaired user perceives similar loudness levels as an unaided normal-hearing person. Loudness is a subjective (and technically not measurable) quantity, but has a direct relation with sound level; therefore, we will instead consider sound pressure levels and require that the input sound level for the normal-hearing listener approximately equals the sound level for the compensated impaired listener:

$$\mathrm{f}(s_n + g_n) \approx s_n, \tag{4}$$

where s_n is the input signal level in dB HL[1], g_n is the estimated gain in dB and function $\mathrm{f}(.)$ is a model for the user's hearing loss. Again, we model "almost equal" by a Gaussian noise term, leading to

$$s_n \;=\; \mathrm{f}(s_n + g_n) + v_n, \tag{5}$$

where $v_n \sim \mathcal{N}(0, \sigma_v^2)$ is the observation noise, which models our uncertainties about the hearing loss model too. Note that the input power level s_n is actually observed (measured) and hence known.

We get to choose the hearing loss model f. In this paper, for illustrative purposes, we selected a (slightly simplified version of a) hearing loss model as proposed by Zurek et al. [3], (Fig. 4):

$$\mathrm{f_z}(x) = \begin{cases} \frac{R}{R-L}(x-L) & \text{if } x \le R \\ x & \text{otherwise} \end{cases} \tag{6}$$

[1]Sound level relative to the average hearing threshold for the population of normal hearing people.

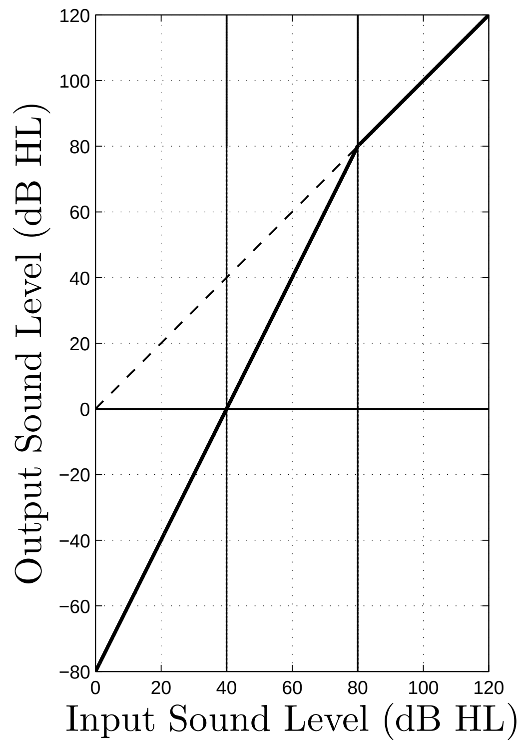

Fig. 4. Input-output function of the simplified version of Zurek's hearing loss simulator in a frequency band with a hearing threshold of $L = 40$ dB HL and a recruitment threshold of $R = 80$ dB HL [3]

.

where x is the sound level in dB HL, and L and R represent the impaired hearing threshold and the recruitment threshold of the user in dB HL, respectively. Zurek's model uses hearing threshold shift and loudness recruitment to simulate hearing loss, and studies have shown that the model is well-matched to the perception of impaired listeners [3].

3. GAIN ESTIMATION BY BAYESIAN INFERENCE

Our probabilistic hearing loss compensation problem is described by

$$g_n = g_{n-1} + w_n \tag{7a}$$
$$s_n = \mathrm{f_z}(s_n + g_n) + v_n \tag{7b}$$
$$w_n \sim \mathcal{N}(0, \sigma_w^2) \tag{7c}$$
$$v_n \sim \mathcal{N}(0, \sigma_v^2) \tag{7d}$$

where s_n is observed and the gain g_n is an unobserved state variable. Hearing loss compensation consists of inference for g_n. With a probabilistic approach, the conditional probability density function $p(g_n|s^n, \theta)$ constitutes the solution of the estimation problem, where $\theta = \{\sigma_w, \sigma_v\}$ and $s^n =$

$\{s_1, s_2, ..., s_n\}$. The posterior of our estimation problem by use of Bayes rule is given by:

$$p(g_n|s^n,\theta) = \frac{p(s_n|g_n,\theta)}{p(s_n|s^{n-1},\theta)} \times p(g_n|s^{n-1},\theta). \quad (8)$$

In Eq. 8, the terms on the right-hand side can be computed from the system Eq. 7 (see the following section). To make the calculations practical and fast, we approximate the probability density functions as Gaussian distributions, which in terms of working out Eq. 8 leads to a Kalman filter [7, 8]. For Zurek's hearing loss model $\mathrm{f_z}$, this system describes a piecewise linear-Gaussian state space model.

3.1. Kalman filter estimation for the compensation gain

In general, a Kalman filter algorithm has two stages. The first part concerns *prediction* over the state based on the previously estimated state:

$$\hat{g}_{n|n-1} = \hat{g}_{n-1}, \quad (9a)$$

$$\sigma^2_{g,n|n-1} = \sigma^2_{g,n-1} + \sigma^2_w, \quad (9b)$$

while the second part is about *correcting* the predicted state based on related observations (in our case: s_n):

$$B_n = \begin{cases} \frac{R}{R-L} & \text{if } s_n + \hat{g}_{n-1} \leq R \\ 1 & \text{else} \end{cases} \quad (10a)$$

$$K_n = \sigma^2_{g,n|n-1} B_n \left(\sigma^2_v + B^2_n \sigma^2_{g,n|n-1}\right)^{-1}, \quad (10b)$$

$$\hat{g}_n = \hat{g}_{n|n-1} + K_n\Big(s_n - \mathrm{f_z}(s_n + \hat{g}_{n|n-1})\Big), \quad (10c)$$

$$\sigma^2_{g,n} = (1 - K_n B_n)\sigma^2_{g,n|n-1}. \quad (10d)$$

Note that $\hat{g}_{n|n-1}$ refers to the mean of the predicted (or a priori) state distribution, and $\sigma^2_{g,n|n-1}$ is the variance of the prediction which models our uncertainty about our prediction based on the previous state estimation.

The factor B_n is the slope of the piecewise hearing loss model. Generally, for a nonlinear hearing loss model, in the *extended* Kalman filter, B_n refers to the derivative of the hearing loss model at the working point.

We call the algorithm described by Eq. 9 and Eq. 10 the *Kalman Hearing Loss Compensation* (KHLC) algorithm.

4. ANALYSIS OF KHLC

4.1. Static characteristics

In Fig. 5 the steady-state gain-vs-input curve of the proposed KHLC algorithm is shown. It can be seen that the initial gain compensates fully for the hearing loss when the sound level is at the absolute hearing threshold. For all higher sounds levels, this amount of gain would be excessive due to recruitment. In KHLC, the static gain linearly decreases (in log-log domain) until the recruitment threshold. When the input is larger than the recruitment threshold, the user can hear like a normal-hearing person (according to our model), and the provided gain equals zero.

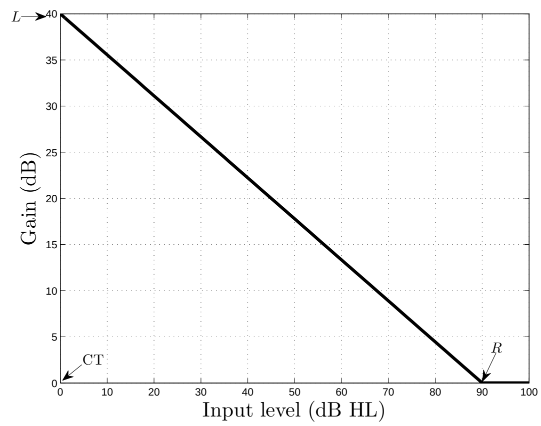

Fig. 5. Steady-state gain-vs-input curve of the KHLC algorithm with hearing threshold $L = 40$ dB HL and recruitment threshold $R = 90$ dB HL based on Zurek's hearing loss model [3].

It is clear from Fig. 5 that KHLC is a compression system. To define the compression ratio of the system, Fig. 6 depicts the steady-state behavior of output versus input signal levels. The compression ratio equals $\mathrm{CR} = \frac{R}{R-L}$ when the input sound level is between 0 dB HL and the recruitment threshold R. When the input sound level is more than R, the amount of the provided gain would be 0 dB; therefore, the compression ratio equals $\mathrm{CR} = 1$.

4.2. Dynamic characteristics

In the correction stage of the Kalman filter, the *Kalman gain* Eq. 10c, K_n balances the weighing of observations s_n versus model predictions. When K_n approaches zero , it means that we trust more our predictions, and observations have little impact on our estimates. In other words, the sensitivity to changes in the observations would be low, which is equivalent to a long time constant. On the other hand, when K_n approaches 1, our observations strongly affect the gain estimation, which relates to a short response time.

When KHLC is in steady state, i.e. $\sigma^2_{g,n} = \sigma^2_{g,n-1}$, it follows from Eq. 9b and Eq. 10d that

$$\sigma^2_{g,n-1} = \frac{1 - K_n B_n}{K_n B_n} \sigma^2_w \quad (11)$$

and Eq. 10b evaluates to

$$K_n = \frac{1}{2}\left(-B_n \frac{\sigma^2_w}{\sigma^2_v} + \sqrt{B^2_n (\frac{\sigma^2_w}{\sigma^2_v})^2 + 4}\right). \quad (12)$$

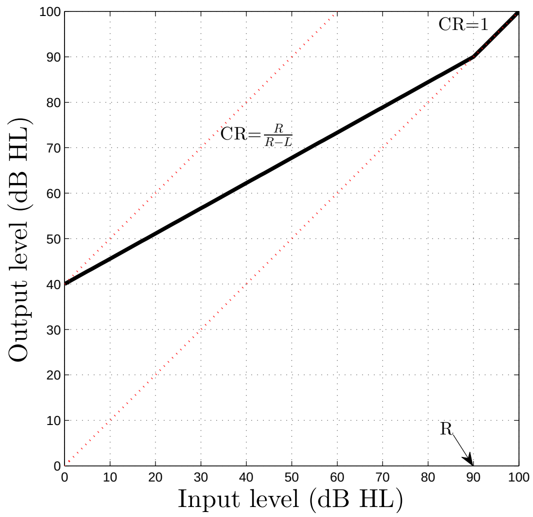

Fig. 6. Steady-state input-output curve of the KHLC algorithm with a hearing threshold of $L = 40$ dB HL and a recruitment threshold of $R = 90$ dB HL based on Zurek's hearing loss model [3].

σ_w	σ_v	$r = \frac{\sigma_w}{\sigma_v}$	Attack time (ms)
0.15	1	0.15	0.44
0.15	2	0.075	0.81
0.1	2.5	0.04	1.56
0.1	5	0.02	3.00
0.2	20	0.01	6.00

Table 1. Attack time of KHLC algorithm for different noise ratios.

Apparently, the dynamic behavior of the KHLC system is predominantly determined by the standard deviation ratio σ_w/σ_v rather than the individual standard deviations.

The dynamic characterization of hearing aid algorithms is described by International Electrotechnical Commission (IEC) standards [2, p. 173]. When the input signal level increases from 55 to 80 dB SPL, the time it takes for the output to stabilize within 2 dB of its ultimate level is called *attack time*. Table 1 illustrates the KHLC attack times to an input sound level step of 55 to 80 dB SPL for different amounts of process and observation noise variances while $L = 29$ dB HL and $R = 90$ dB HL. The table confirms that the attack time has an inverse linear relationship with the fraction σ_w/σ_v.

5. ALTERNATIVE HEARING LOSS MODELS

In Section 4 it was shown that Zurek's HL model in the KHLC algorithm leads to a compression ratio CR $= \frac{R}{R-L}$ and a compression threshold $C = 0$ dB HL.

In hearing aids, most classic DRCs have compression threshold larger than 0 dB HL, which is apparently not con-

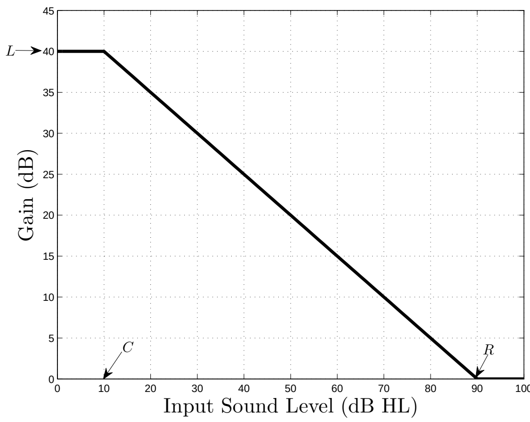

Fig. 7. Steady-state gain-vs-input curve of the KHLC gain based on the f_a hearing loss model with hearing threshold $L = 40$ dB HL and recruitment threshold $R = 90$ dB HL, and $C = 10$ dB HL.

sistent with Zurek's hearing loss model. This is remarkable and it illuminates the risks of letting go of a direct mathematical relationship between the problem (hearing loss) and the proposed solution (DRC). In contrast, in the KHLC framework the solution is inferred from a problem description. We will now amend Zurek's model to account for compression thresholds.

Our proposed alternative hearing loss model $f_a(.)$ is defined by

$$f_a(x) = \begin{cases} x - L & \text{if } x < C + L \\ \alpha x - \beta & \text{else if } C + L \leq x \leq R \\ x & \text{otherwise} \end{cases} \quad (13)$$

where $\alpha = \frac{R-C}{R-(C+L)}$ is also the compression ratio, $\beta = \frac{L.R}{R-(C+L)}$, and C determines the compression threshold in dB HL. As before, R represents the recruitment threshold and L is the impaired-hearing threshold.

Fig. 7 depicts the static gain-vs-input curve of the KHLC algorithm based on hearing loss model f_a. It can be seen that the provided gain is constant for the sound levels with levels less than the compression threshold (i.e. C = 10 dB HL). The gain decreases when sound levels are higher than C and less than the recruitment threshold. When the input sound level is more than recruitment threshold, the user can hear like a normal-hearing person (according to this model), and the inferred gain equals zero.

The KHLC algorithm with the alternative hearing loss model f_a leads to the same gains that are prescribed by the classic DRC algorithm. To show this, we simulated step responses both with a KHLC algorithm (with $L = 50$ db HL, $R = 110$ db HL, $C = 10$ dB, $\sigma_w = 0.1$, and $\sigma_v = 20$) and the classic DRC algorithm that we defined in Section 1.1 (with parameters $CR = 2$, $CT = 10$ db HL, $\tau = 6$ ms. Fig. 8 depicts the responses of both systems to a step input signal.

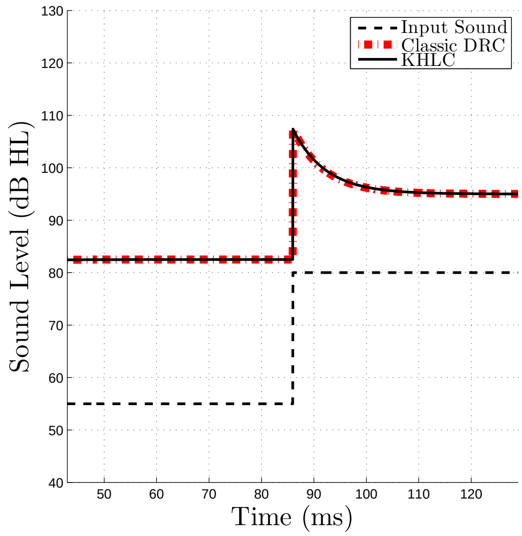

Fig. 8. Comparison of the amplified signal by the inferred KHLC gain versus the classic DRC algorithm.

It can be seen that the inferred gain in the KHLC algorithm corresponds to the gain in the DRC algorithm. The benefit of the proposed method lies in the fact that our compensation gain is *inferred* by optimal estimation on the basis of a hearing loss model that can be exchanged en bloc by the designer. As a result, in the design of our gain compensation system we can make use of the large literature of hearing loss models, e.g. [9, 10, 11]. Through gain inference, we are guaranteed that the solution (the gain tracks) optimally correspond to the problem statement. As discussed, this feature is unfortunately lost in today's dynamic range compression based solutions to hearing loss.

6. CONCLUSIONS

In this paper, we formulated the hearing loss compensation as a probabilistic inference problem. This is the first step towards a new class of dynamic range compressors which are directly inferred by Kalman filtering from explicit problem definitions in the form of hearing loss models.

In contrast to classic dynamic range compression solutions to hearing loss, in our approach there is a direct mathematical link between the problem statement and the solution. Moreover, our approach opens the door to designing hearing loss compensation algorithms with the benefit of access to the large literature on both Kalman filtering and hearing loss models. Our approach is a principled method that allows extension of inference to parameter estimation (i.e. the difficult 'fitting' problem in hearing aids) as well as inference over competing hearing loss models. The latter topics will be considered in future work.

7. REFERENCES

[1] William B Snow, "Audible frequency ranges of music, speech and noise," *The Journal of the Acoustical Society of America*, pp. 155–166, 1931.

[2] Harvey Dillon, *Hearing Aids*, Thieme, 2012.

[3] Patrick M Zurek and Joseph G Desloge, "Hearing loss and prosthesis simulation in audiology," *The Hearing Journal*, vol. 60, no. 7, 2007.

[4] Dimitrios Giannoulis, Michael Massberg, and Joshua D Reiss, "Digital dynamic range compressor design - A tutorial and analysis," *Journal of the Audio Engineering Society*, vol. 60, pp. 399–408, 2012.

[5] Tom Heskes and Bert de Vries, "Incremental utility elicitation for adaptive personalization," in *Belgium-Netherlands Conference on Artificial Intelligence*, 2005, pp. 127–134.

[6] James M. Kates and Kathryn Hoberg Arehart, "Multichannel Dynamic-Range Compression Using Digital Frequency Warping," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, no. 18, pp. 3003–3014, 2005.

[7] J.F.G. de Freitas, M. Niranjan, and A.H. Gee, "Hierarchical bayesian models for regularisation in sequential learning," *Neural computation*, vol. 12, no. 4, pp. 933–53, Apr. 2000.

[8] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall PTR, 1993.

[9] Jean-Pierre Gagne and Norman P Erber, "Simulation of sensorineural hearing impairment," *Ear and hearing*, vol. 8, no. 4, pp. 232–243, 1987.

[10] Brian CJ Moore and Brian R Glasberg, "Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech," *The Journal of the Acoustical Society of America*, vol. 94, no. 4, pp. 2050–2062, 1993.

[11] Dong Wook Kim, Young Cheol Park, Won Ky Kim, Seon Joon Park, Won Doh, Seung Woo Shin, and Dae Hee Youn, "Simulation of hearing impairment with sensorineural hearing loss," in *Engineering in Medicine and Biology Society*. IEEE, 1997, vol. 5, pp. 1986–1989.