UNIVERSITY OF CALIFORNIA SAN DIEGO
SAN DIEGO STATE UNIVERSITY

Multirate Signal Processing for Wide Dynamic Range Compression and Feedback Control in Hearing Aids

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Engineering Science (Electrical & Computer Engineering)

by

Alice Sokolova

Committee in charge:

University of California San Diego
- fred harris, Co-Chair
- Harinath Garudadri
- Rajesh Gupta
- William Hodgkiss

San Diego State University
- Baris Aksanli, Co-Chair
- Yusuf Ozturk

2023

The dissertation of Alice Sokolova is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

Co-Chair

Co-Chair

University of California San Diego

San Diego State University

2023

DEDICATION

To family, friends, and future endeavors,

And thus concludes a lengthy journey
The time has come to turn the page,
And close this chapter of my story,
Marked with sacrifice and change.

The work presented in this thesis
I vow to solemnly defend.
Let this committee be my witness,
The search for knowledge never ends.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to first thank my advisors, Professor fred harris and Professor Baris Aksanli, and my Principal Investigator Dr. Harinath Garudadri for guiding my journey during my Ph.D studies. I would like to thank Professor harris for inspiring me in my high school years to apply to San Diego State University and pursue a degree in Electrical Engineering, and for instilling in me an interest in Digital Signal Processing which has led me to this day. I would like to thank Professor Aksanli for introducing me to academic research during my undergraduate years, and for guiding me through the intricacies of scholarly research. I would like to thank Dr. Harinath Garudadri for directing my research and inspiring me with his passion for hearing loss intervention. I would also like to thank my dissertation committee members, Professors Rajesh Gupta, William Hodgkiss, and Yusuf Ozturk for their feedback and discussions of my Ph.D. work.

My research was made possible by funding from the National Institute of Health, NIH/National Institute on Deafness and Other Communication Disorders (NIDCD), under Grant R21DC015046 and Grant R33DC015046 (Self-fitting of Amplification: Methodology and Candidacy), Grant R01DC015436 (A Real-time, Open, Portable, Extensible Speech Laboratory), and Grant R44DC020406 (An Open-source Speech Processing Platform (OSP) for Research on Hearing Loss and Related Disorders); Division of Information and Intelligent Systems, University of California San Diego, under Grant IIS-1838897 (A Framework for Optimizing Hearing Aids In Situ Based on Patient Feedback, Auditory Context, and Audiologist Input); Qualcomm Institute, University of California San Diego (UCSD); San Diego State University “fred harris Excellence Scholarship; Halıcıoğlu Data Science Institute, UCSD; Wrethinking, the Foundation.

The material in this dissertation is based on the following publications.

Chapter 2 and Chapter 3, in part, are a reprint of the material as it appears in: A. Sokolova, D. Sengupta, M. Hunt, R. Gupta, B. Aksanli, f. harris, and H. Garudadri, “Real-Time Multirate Multiband Amplification for Hearing Aids”. *IEEE Access*, 2022; A. Sokolova, D. Sengupta, K.L. Chen, R. Gupta, B. Aksanli, f. harris, H. Garudadri. ”Multirate Audiometric Filter Bank for

Hearing Aid Devices", Asilomar Conference on Systems, Signals, and Computers, 2021. The dissertation author was the primary investigator and author of this material.

Chapter 4, in part, is a reprint of the material as it appears in A. Sokolova, B. Aksanli, f. harris, and H. Garudadri, "A Curvilinear Transfer Function for Wide Dynamic Range Compression with Expansion". *Journal of the Audio Engineering Society*, 2023. (under review) The dissertation author was the primary investigator and author of this material.

Chapter 5, in part, is a reprint of the material as it appears in A. Sokolova, V. Rallapalli, A. Yellamsetty, M. Hunt, B. Aksanli, f. harris, and H. Garudadri, "Validation of Frequency Warping (Freping) as a new tool for feedback control in hearing aids". *Asilomar Conference on Systems, Signals, and Computers*, 2023. The dissertation author was the primary investigator and author of this material.

My co-authors (Prof. Baris Aksanli, Kuan-Lin Chen, Harinath Garudadri, Prof. Rajesh Gupta, Prof. fred harris, Martin Hunt, Prof. Varsha Rallapalli, Dhiman Sengupta, Prof. Anusha Yellamsetty, listed in alphabetical order) have all kindly approved the inclusion of the aforementioned publications in my dissertation.

VITA

2018	Bachelor of Science in Electrical Engineering, San Diego State University, San Diego, CA, USA
2019	Lab Instructor, San Diego State University, CA, USA
2021	Instructional Assistant, University of California, San Diego, CA, USA
2022	Teaching Assistant, San Diego State University, CA, USA
2022	Lecturer, San Diego State University, CA, USA
2023	Doctor of Philosophy in Engineering Science (Electrical & Computer Engineering), University of California, San Diego, USA and San Diego State University, USA

PUBLICATIONS

Alice Sokolova, Baris Aksanli, fred harris, Harinath Garudadri. "A Curvilinear Transfer Function for Wide Dynamic Range Compression with Expansion", *Journal of the Audio Engineering Society*, 2023. (under review)

Alice Sokolova, Varsha Rallapalli, Anusha Yellamsetty, Martin Hunt, Baris Aksanli, fred harris, Harinath Garudadri. "Validation of Frequency Warping (Freping) as a new tool for feedback control in hearing aids", *Asilomar Conference on Systems, Signals, and Computers*, 2023.

Alice Sokolova, Baris Aksanli, fred harris, Harinath Garudadri. "Consolidating Compression and Revisiting Expansion: An alternative amplification rule for Wide Dynamic Range Compression", *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2023.

Alice Sokolova, Dhiman Sengupta, Martin Hunt, Rajesh Gupta, Baris Aksanli, fred harris, Harinath Garudadri. "Real-Time Multirate Multiband Amplification for Hearing Aids", *IEEE Access*, 2022.

Alice Sokolova, Dhiman Sengupta, Kuan-Lin Chen, Rajesh Gupta, Baris Aksanli, fred harris, Harinath Garudadri. "Multirate Audiometric Filter Bank for Hearing Aid Devices", *Asilomar Conference on Systems, Signals, and Computers*, 2021. (best paper nominee)

Alice Sokolova, Mohsen Imani, Andrew Huang, Ricardo Garcia, Justin Morris, Tajana Rosing, Baris Aksanli. "MACcelerator: Approximate Arithmetic Unit for Computational Acceleration", *International Symposium on Quality Electronic Design*, 2021.

Alice Sokolova, Mohsen Imani, Ricardo Garcia, Andrew Huang, Fan Wu, Baris Aksanli, Tajana Rosing. "ApproxLP: Approximate Multiplication with Linearization and Iterative Error Control",

Design Automation Conference, 2019.

Alice Sokolova, Baris Aksanli. ”Demographical Energy Usage Analysis of Residential Buildings”, *Journal of Energy Resources Technology, Transactions of the ASME*, 2019.

Alice Sokolova, Baris Aksanli. ”Demographical Energy Usage Analysis of Residential Buildings”, *ASME International Energy Sustainability and Power Conference*, 2018. (best paper nominee)

ABSTRACT OF THE DISSERTATION

Multirate Signal Processing for Wide Dynamic Range Compression and Feedback Control in Hearing Aids

by

Alice Sokolova

Doctor of Philosophy in Engineering Science (Electrical & Computer Engineering)

University of California San Diego, 2023
San Diego State University, 2023

Professor fred harris, Co-Chair
Professor Baris Aksanli, Co-Chair

Hearing loss affects millions of people in the world, and has a deep impact on quality of life. Consequences of hearing loss can include difficulty of communication, social isolation, and even an increased risk of developing dementia. While hearing loss treatment is not restorative, hearing aids aim to compensate for the hearing loss and restore audibility of sound using a combination linear and nonlinear amplification methods.

One of the fundamental features of modern hearing aids is Wide Dynamic Range Compression (WDRC). Compression seeks to compensate for hearing loss by reducing the dynamic

range of an audio signal and mapping the signal onto the remaining dynamic range of a hearing impaired individual, which is generally narrower than that of a normal hearing individual. Compression has been shown to improve loudness comfort, audibility, and intelligibility of sound for hearing impaired individuals, however, there remains significant user dissatisfaction with compressive hearing aids. This dissertation describes an end-to-end WDRC system which offers solutions to many ongoing hearing aid research challenges, and consists of the following subsystems: frequency decomposition, magnitude estimation, automatic gain control, feedback management, and user interface. Our system is open source and is part of a greater initiative to democratize hearing aid research.

Frequency decomposition is the first stage of compression. Since hearing, whether normal or impaired, is inherently frequency dependent, different frequency components must be processed differently, necessitating frequency channelization. We designed a multirate eleven band half-octave filter bank for audiometric signal decomposition. Our filter bank surpasses other non-proprietary hearing aids, offering more channels with narrower bandwidth at lower computational cost. Our novel multirate processing structure dramatically reduces the complexity of the system by processing each channel at a sampling rate proportional to its frequency. Our multirate channelizer also enables all subsequent hearing aid processing to be performed at multiple sampling rates, reducing power consumption of the entire processing chain.

The following stage of WDRC is magnitude estimation and automatic gain control. The conventional approach uses a peak detector to track a smoothed version of the signal envelope and apply gain. However, a peak detector produces an inaccurate estimate of the signal envelope with ripple. Moreover, the effects of attack and release times in hearing aids are not well explored in academic literature. We derived a closed-form solution relating attack and release time to other hearing aid parameters. Based on this derivation, we designed a highly accurate envelope detector based on the Hilbert Transform, followed by a precise automatic-gain-control algorithm yielding highly accurate dynamic performance.

Central to the automatic-gain-control algorithm is the input-output transfer function,

which maps input signal magnitude to output magnitude. The conventional WDRC input-output transfer function is a piece-wise segmented function, defined by four or more parameters. Conventional WDRC also overamplifies low-level signals, which reduces output signal-to-noise ratio and causes the hearing aid to sound noisy. We designed a novel curvilinear transfer function for WDRC which simplifies the parameterization of compression, reduces distortion caused by hard knee points, and improves signal-to-noise ratio by reducing amplification of low-level noisy.

Lastly, a major difficulty in hearing aid research is acoustic feedback management. Adaptive filters are highly effective at cancelling feedback, however, there is a delay in their response to a change of acoustic conditions, such as a hand touching the ear, and adaptive filters may cancel useful components of the input signal as well as feedback. We designed a feedback cancellation algorithm which uses variable frequency shifting to successfully eliminate feedback in realistic hearing aid test conditions.

Chapter 1

Introduction

Approximately 15% of American adults (35 million) report some trouble hearing [2]. Among this population, about 29 million U.S. adults could benefit from hearing aids [3]. However, only approximately one in four hearing impaired individuals adopt hearing aids [4]. There is a wide range of reasons for hearing aid disuse, including the high cost of hearing aids, social stigma, and unattractive appearance [5], however, most user dissatisfaction stems from insufficient hearing benefits offered by hearing aids [4, 5, 6].

Modern hearing aids are complex devices, performing far more complicated tasks than uniform amplification of a wideband audio signal. Modern digital hearing aids offer a wide variety of features, including noise suppression, beam forming, feedback cancellation, Bluetooth audio streaming, and more. However, the core functionality of all hearing aids is amplification. This dissertation will focus on the most common amplification strategy in modern hearing aids called Wide Dynamic Range Compression (WDRC). WDRC is a nonlinear amplification strategy which reduces the dynamic range of an audio signal to accommodate the reduced dynamic range of hearing impaired individuals. An alternative description is that compression applies higher gain to quieter signals to restore audibility of quiet sounds for hearing impaired users, and limits the volume of loud signals which may cause discomfort to the user. Most modern hearing aids are digital multichannel compressive devices [7, 8], meaning amplification is performed independently for various frequency channels. WDRC has been shown to be

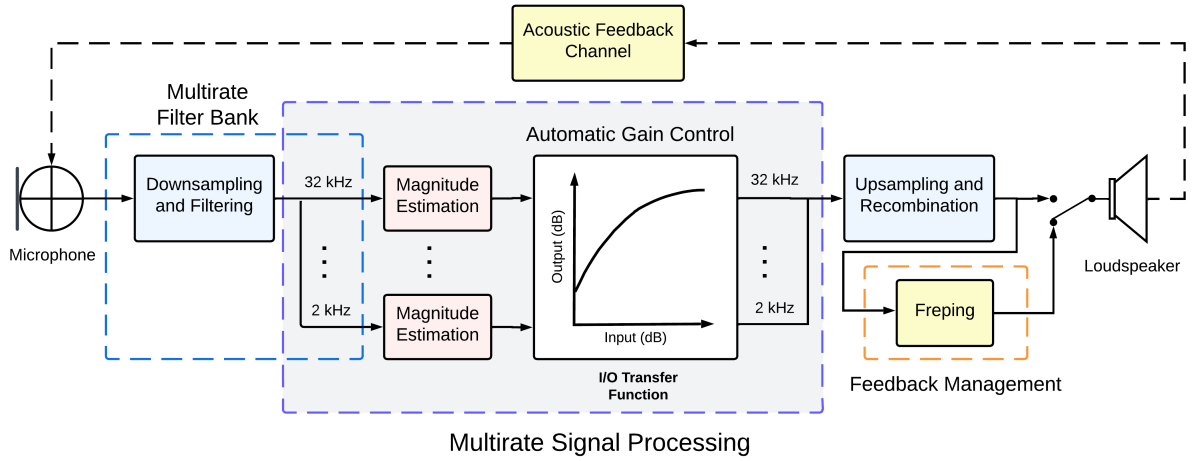

Figure 1.1. Block diagram of our multirate compressive hearing aid, performing frequency decomposition, automatic gain control, and feedback management

effective in increasing audibility of low-level signals, improving loudness-comfort. However, there remains significant dissatisfaction with the performance of compressive hearing aids, such as speech intelligibility, sound quality, and speech recognition in noise [5, 6, 7]. This dissertation seeks to improve upon all aspects of WDRC by presenting a novel multirate WDRC system with proposed solutions for multiple WDRC design challenges.

Figure 1.1 shows the algorithmic structure of our multirate compressive hearing aid. This dissertation focuses on the algorithmic structure of the hearing aid, so details on signal acquisition and analog-to-digital conversion of the microphone signal are omitted. Our multirate hearing aid consists of three main subsystems, which are frequency decomposition, automatic gain control (AGC), and feedback management. The first stage in our signal processing chain is frequency decomposition, which is accomplished with a multirate filter bank. The outputs of the filter bank are eleven frequency channels at proportional sampling rates. The envelope of each multirate frequency channel is then computed and passed to the AGC subsystem. The AGC subsystem computes the target output levels as a function of input levels and dynamically adjusts the gain in each subband in accordance with the designated attack and release time constants. The amplified channels are then restored to the original sampling rate and recombined to produce

the compressed output signal, which is delivered to the receiver (loudspeaker) of the hearing aid. Optionally, if the hearing aid is entering feedback, our feedback management system called Freping can be enabled as the last processing stage. Each subsystem of our compressive hearing aid is designed to address pertinent challenges in hearing aid research, and will be discussed in more detail throughout this chapter and this dissertation.

Our multirate compressive hearing aid amplification system is part of a greater effort to democratize hearing aid research. The proprietary "blackbox" nature of commercial hearing aids poses numerous limitations for developing and evaluating novel hearing intervention techniques. In a 2014 workshop conducted by the National Institute on Deafness and Other Communication Disorders (NIDCD), a member of the National Institutes of Health (NIH), it was identified that the hearing loss research community has a critical need for an open source, non-proprietary research platform which supports acoustic signal processing in real time [9]. In response to the needs identified by the NIH, Open Speech Platform (OSP) was founded at University of California, San Diego (UCSD) [10]. OSP is a suite of open source software and hardware tools for hearing aid research, which include a portable, wearable hearing aid processing device, wearable ear-level transducers, and a complete library of master hearing aid software [1, 11, 12]. The multirate compression system described in this dissertation serves as the WDRC software of the OSP hearing aid.

The work in this dissertation also has high applicability to the growing field of over-the-counter hearing aids. The audiology community recognizes the need for self-adjusted hearing aids [13] to meet the unmet needs of hearing impaired users, such as users without access to audiological healthcare professionals, users hindered by the high cost of prescription hearing aids, or users desiring greater customization of their hearing aids, such as musicians. On August 17, 2022, the Food and Drug Administration (FDA) issued a ruling legalizing over-the-counter hearing aids for individuals with perceived mild to moderate hearing loss [14]. This dissertation includes contributions which enable greater customization of hearing aids for user desiring greater personalization of their hearing aid.

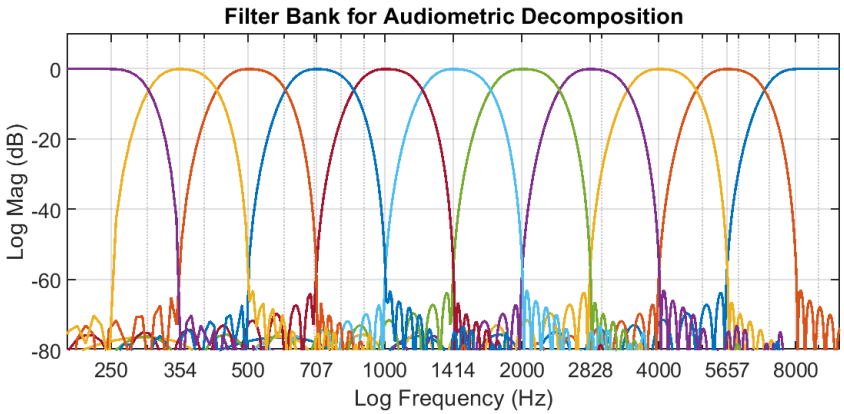

Figure 1.2. Multirate filter bank for frequency decomposition in hearing aids

1.1 Frequency Decomposition

Most modern compressive hearing aids use multichannel compression [15], meaning the audio signal is separated into frequency bands, where each frequency band is compressed independently. The use of multichannel compression is motivated by the fact that hearing loss is generally frequency dependent, with higher frequencies generally more affected by hearing loss than lower frequencies. A higher number of channels with finer frequency resolution allows audiologists to more accurately compensate for hearing loss at various frequencies and accommodate unusual hearing loss patterns, such as cookie bite [7]. However, the complexity of subband decomposition in hearing aids is limited by the processing capabilities of embedded processors, battery resources, and strict latency requirements for real-time operation. The tolerable algorithmic delay in hearing aids is commonly considered to be 10 ms [16]. These constraints make it difficult to design sufficiently narrowband filters to meet audiological needs.

Due to the logarithmic nature of the human cochlea, frequency decomposition in hearing aids is performed on the logarithmic scale. For example, the American Speech-Language-Hearing Association (ASHA) defines a set of ten audiometric frequencies used for pure-tone audiometry, which approximate half-octave center frequencies in the range of 250 Hz to 8000 Hz [17]. Due to these logarithmic properties, the filters for low frequency bands have extremely narrow bandwidths, which require very high complexity to implement. However, we circumvent

these high complexity requirements by utilizing multirate signal processing. Figure 1.2 shows our eleven band half-octave filter bank. We use downsamplers to map each octave of subbands to a sampling rate proportional to the frequency of the channel group. This increases the relative bandwidth of the filter and reduces the complexity of the filter logarithmically. This technique allows us to implement very fine resolution frequency decomposition on a power-limited battery-operated device.

1.2 Automatic Gain Control

Automatic Gain Control (AGC) in WDRC is the process of applying nonlinear gain to each frequency channel to reduce its dynamic range. Figure 1.3 illustrates the output magnitude as a function of input magnitude of conventional WDRC. In the typical implementation of compression, low-level signals receive linear amplification, signals above the compression threshold are compressed at a fixed rate referred to as the compression ratio (CR), and signals above the maximum output power (MPO) threshold are hard limited to prevent discomfort to the hearing impaired user. The gain and compression parameters of WDRC are programmed by an audiologist based on each individual's hearing loss pattern, and are termed as the hearing aid prescription.

The AGC subsystem of WDRC consists of signal magnitude estimation, the input-output magnitude transfer function, and the AGC loop itself. The performance of WDRC is evaluated based on static behavior, which is the ability of the WDRC system to meet target gain values for inputs of constant magnitude, and dynamic behavior, which is the ability of the WDRC system to respond to changes in the input with the desired response speed. The American National Standards Institute (ANSI) defines a standard for the performance of AGC in hearing aids [18]. However, many commercial and non-proprietary hearing aids fail to provide high accuracy in meeting desired static and dynamic behavior. For example, a hearing aid system may use approximations and "fudge factors" to roughly calibrate the performance of WDRC. Inaccuracy

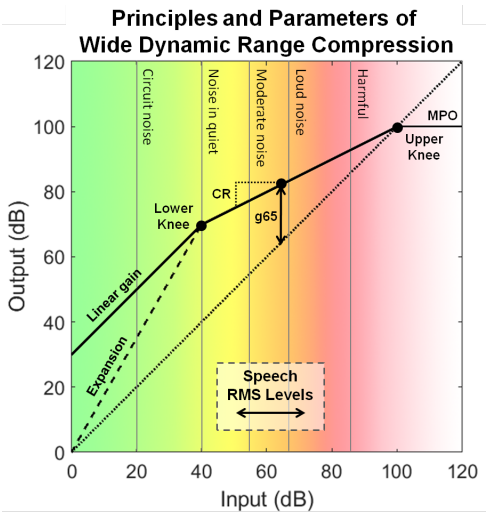

Figure 1.3. Output levels of WDRC as a function of input levels with governing parameters.

in meeting desired WDRC behavior introduces significant uncertainty in prescribing hearing aid parameters to users and exploring the perceptual effects of these parameters.

In this dissertation, we redesign the magnitude estimation subsystem and AGC loop to improve accuracy and reduce the uncertainty of WDRC. We explore the relationship between the static and dynamic parameters of WDRC and derive a closed-form solution for expressing response times as a function of static parameters. We leverage this derivation to design an AGC algorithm with highly accurate performance based on the ANSI hearing aids standards. We also challenge the conventional WDRC static behavior, governed by the input-output transfer function, and propose a novel input-output transfer function which addresses sources of user dissatisfaction with compressive hearing aids.

1.3 Feedback Management

Feedback management in hearing aids is a challenging issue, complicated by the growing use of open fittings in modern devices, smaller form factors, and shorter distances between speaker and microphone [19]. At the same time, it is one of the most prominent causes of dissatisfaction for hearing aid users [4]. The issue of feedback may limit the maximum gain a hearing aid can provide, causing it to fail in meeting the prescribed hearing aid gains.

Numerous approaches have been explored for feedback management in hearing aids, the most prominent of which is adaptive filtering. Adaptive filtering has been highly successful in preventing feedback by estimating the signal in the feedback path and subtracting it and subtracting it from the microphone input signal. However, an imperfect estimate of the feedback signal may cause parts of the desired input signal to be cancelled out along with the feedback signal, negatively affecting sound quality [19]. For this reason, it is sometimes advisable to disable adaptive feedback cancellation in situations when output sound quality is a priority, such as when listening to music through hearing aids [20]. Adaptive filters also have a delayed response to quick changes in the acoustic feedback path, such as when a phone is brought to the ear, causing brief howling which is of great annoyance to the user.

We revisit the method of frequency shifting as a tool for feedback management in hearing aids. Our approach, called Freping (frequency + warping), uses variable frequency shifting to prevent stable oscillations in the feedback loop. We used standard audiological procedures to validate the effectiveness of our approach using our OSP hearing aid, worn by an acoustic manikin simulating the human head and torso. Our frequency warping technique is successful at eliminating feedback in tested conditions, and can be used both as a standalone tool, and in conjunction with adaptive methods.

1.4 Thesis Contributions

This thesis describes an end-to-end WDRC system which addresses current problems in subband decomposition, amplification, and feedback control in hearing aids. Central to the design of our amplification system is multirate signal processing, which allows the system to process various frequency components of sound at multiple sampling rates proportional to the frequency of each band. We also present a solution to the problem of feedback management in hearing aids. The following discussion summarizes the contributions and outlines the rest of the thesis:

- Chapter 2 presents a multirate filter bank for audiometric frequency decomposition. We designed a logarithmic filter bank which separates an audio signal into half-octave frequency channels, where each octave is processed at a proportional sampling rate. We further improve the processing efficiency of our filter bank by utilizing polyphase filtering. This approach allows us to achieve about $14\times$ improvement in complexity over an equivalent single-rate implementation. Moreover, multirate processing allows us to design extremely sharp and narrowband filters, which occupy only a fraction of the entire signal bandwidth. This improved frequency resolution allows finer adjustment of the hearing aid at low frequencies than other available alternatives. The filter bank operates at a very small latency of 5.4 ms, which is sufficient for real-time operation on a hearing aid.

- Chapter 3 presents a magnitude estimation algorithm and an AGC loop for compressive amplification of the sound signal. Our Hilbert Transform-based magnitude estimation algorithm yields a highly accurate estimate of the signal envelope compared to a classic peak-detector approach. We also present a closed-form derivation of temporal parameters (attack and release time) as a function of hearing aid gain parameters. Based on this derivation, we designed a highly accurate AGC algorithm which satisfies the target temporal performance of WDRC with 0.5 ms accuracy or better. Compressive amplification is performed with high efficiency using multirate processing.

- Chapter 4 presents a novel amplification rule for WDRC. Central to WDRC is the amplification transfer function, which governs the relationship between input magnitude and output magnitude of the compressed signal within the AGC mechanism. We redesign the input-output transfer function of WDRC to leverage the capabilities of modern processors and improve upon the signal distortion introduced by WDRC. Our redesigned compression transfer function improves output SNR of compressed audio by up to 30%, and sound quality by up to 6.7% as measured by the HASQI index. We also discuss an application of the novel amplification rule to over-the-counter hearing aids.

- Chapter 5 presents an algorithm called Freping for acoustic feedback control. Adaptive filters are highly successful at cancelling the feedback path of an acoustic feedback loop, however, they may also cancel desired components of the desired signal, reducing sound quality, and have a delayed response to changes in acoustic conditions. We designed an algorithm which utilizes variable frequency shifting to prevent the constructive summation of the amplified signal. We validated the effectiveness Freping in real-world conditions using an acoustically realistic manikin. Experimental results show that Freping offers up to 11 dB of added stable gain, which measures the additional usable gain of a hearing aid offered by feedback cancellation.

- Chapter 6 presents a summary of the dissertation and discusses possible future research directions in the field of compressive amplification and feedback management in hearing aids.

Chapter 2

Multirate Audiometric Filter Bank for Frequency Decomposition in Hearing Aids

2.1 Introduction

Studies have shown that only about one-third of individuals who have hearing loss utilize a hearing aid. Among those individuals, around one-third do not use their hearing aids regularly. The main reason for this disuse is often the dissatisfaction with the speech quality offered by modern hearing aids, especially in noisy environments where hearing-impaired individuals need them the most [6]. Achieving music appreciation with hearing aids is an even greater challenge [20]. A wide variety of signal processing algorithms have been proposed to improve the sound quality of hearing aids, such as Speech Enhancement, Dynamic Range Compression, Frequency Warping, and more [21, 22, 23, 24]. However, because human hearing is inherently frequency-dependent, many of these algorithms require dividing the input signal into frequency sub-bands [21, 22, 23, 24], a process called sub-band decomposition.

Sub-band decomposition is prevalent in commercial and open-source hearing aid devices. Among open source hearing aids, one of the most widely employed designs is that of James Kates [25]. However, Kates's baseline hearing aid algorithm, which contains six frequency bands, is starting to become insufficient for the advancing needs to audiometric research. Work in [7] determined that finer frequency subdivision can provide researchers with increased flexibility for hearing loss compensation, especially for unusual hearing loss patterns.

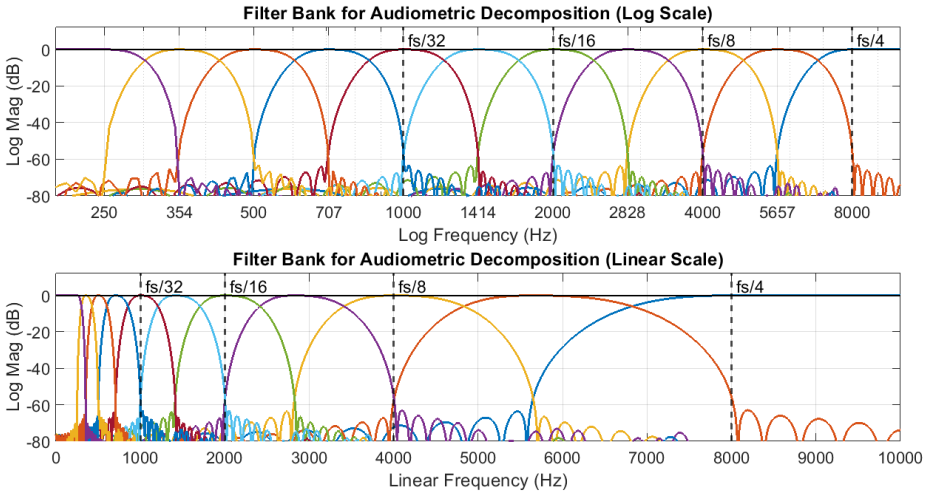

Figure 2.1. The magnitude response of our multirate audiometric filter bank, shown on the logarithmic and linear scale. Vertical dotted lines represent fractions of the input sampling rate.

In this chapter, we present a multirate real-time filter bank, also known as a channelizer, for audiometric research. Our proposed filter bank, shown in Figure 2.1, is uniformly distributed on the logarithmic scale, and reflects the behavior of the human cochlea. The cochlea inherently has frequency-dependent spectral resolution, meaning it can discern finer pitch variations at lower frequencies than at higher frequencies. Our proposed filter bank replicates this property by offering very narrow bandpass filters at lower frequencies, and wider filters at higher frequencies. Our channelizer also offers high sidelobe suppression and perfect signal reconstruction within ± 0.15 dB.

We implemented our filter bank on the Open Speech Platform (OSP) [10, 1, 11]. OSP is an open-source suite of software and hardware tools for performing research in audiology, which includes a wearable hearing aid, a wireless interface, and a set of hearing enhancement algorithms.

2.2 Filter Bank

2.2.1 Overview

Figure 2.1 shows the Eleven Band Multirate Filter Bank, also known as a channelizer, for subband decomposition. Subband decomposition is the process of separating a signal into multiple frequency bands or channels, and is used in many applications, including hearing aids [21, 22, 23, 24]. The multirate filter bank possesses the following properties:

Center Frequencies

The structure of an audiometric filter bank reflects the spectral nature of the human cochlea, which is inherently logarithmic. The American Speech-Language-Hearing Association (ASHA) defines a set of ten audiometric frequencies used for pure-tone audiometry, which are 0.25, 0.5, 1, 1.5, 2, 3, 4, 6, and 8 kHz [17]. These frequencies closely resemble a half-octave logarithmic sequence, and are commonly targeted for audiometric filter banks. However, every other frequency is not a true half-octave frequency, but rather a simplified integer approximation. The audiometric filter bank is a true half-octave channelizer, making it uniformly distributed on the logarithmic scale, as seen from Figure 2.1. It spans a range of 0.25 to 8 kHz, which produces eleven bands. Although the true half octave center frequencies diverge from the rounded ASHA approximations, they are functionally the same, and for the sake of simplicity we will be referring to each individual band by it's approximate audiometric frequency.

Attenuation and Ripple

The American National Standards Institute (ANSI) defines specifications for Half-Octave Acoustic filters [26]. The standard includes three classes of filters – class 0, 1, and 2, where class 0 has the strictest tolerances and class 2 has the most lax tolerances. The filter bank meets class 0 standards – the highest of the three. Accordingly, each band of the filter bank has -75 dB sidelobe attenuation, and the in-band ripple is within ± 0.15 dB. The ripple of the composite response of the channelizer is also within ± 0.15 dB.

Filter Shape and Composite Response

Figure 2.1 shows the audiometric filter bank on both the logarithmic and the linear scale. As seen from Figure 2.1, filters which are symmetrical on the logarithmic scale are asymmetrical on the linear scale. We designed asymmetrical bandpass filters by convolving a lowpass and a highpass filter for each band.

A more difficult challenge, though, is achieving signal reconstruction. A filter bank has perfect reconstruction if the sum of all outputs is equal to the original input signal. In the frequency domain, this means the composite frequency response of the filter bank is a flat line spanning all frequencies, as shown in Figure 2.1.

We ensure that our filter bank has perfect reconstruction by employing complementary filter design. Complementary filters are two filters the sum of which is an all-pass filter. For any highpass or lowpass filter, its complement can be found by subtracting it from an all-pass filter, which is simply an impulse in the time domain. We designed all neighboring filter edges to be complements of each other, ensuring that their sum is an all-pass filter, which guarantees signal reconstruction. The channelizer offers perfect reconstruction within ± 0.15 dB.

Multirate Processing and Latency

It is well known in the signal processing community that the sharper a digital filter is, the more coefficients it requires. As seen from Figure 2.1, the audiometric channelizer requires very narrow and sharp filters – the lowest center frequency (0.25 kHz) is $32\times$ smaller than the highest center frequency (8 kHz), and at a 32 kHz sampling rate, the width of the narrowest filter is only $1/64$ of the entire signal bandwidth. A conventional implementation of such narrow filters would result in too much latency to meet real-time processing deadlines, and would require excessive processing power. Our filter bank dramatically reduces both power consumption and latency by employing multirate signal processing. Compared to a single-rate implementation, multirate processing reduces the power consumption by a factor of $13.7\times$, and reduces latency from 32 ms down to 5.4 ms.

Table 2.1. A comparison of the number of coefficients in each filter of the Audiometric Filter Bank, with and without multirate processing

Filter Band:	Filter Taps		Sampling rate
	Single-rate	Multirate	
8 kHz	53	53	1
6 kHz	77	77	1
4 kHz	154	77	1/2
3 kHz	154	77	1/2
2 kHz	308	77	1/4
1.5 kHz	308	77	1/4
1 kHz	616	77	1/8
0.75 kHz	616	77	1/8
0.5 kHz	1232	77	1/16
0.375 kHz	1232	77	1/16
0.25 kHz	1232	77	1/16

2.2.2 Multirate Signal Processing

The motivation behind multirate processing is to decrease the complexity of a filter by reducing the sampling rate. Table 2.1 lists the number of taps needed to implement the filters shown in Figure 2.1 at a single sampling rate of 32 kHz. As the filters becomes narrower and sharper, they require an exponentially increasing number of taps, reaching impractical values at the lowest frequencies.

However, the complexity of a filter can be decreased by reducing the sampling rate. For a given bandpass filter, the relative bandwidth is narrower at a higher sampling rate and wider at a lower sampling rate. Thus, a filter spanning a fixed range of frequencies becomes relatively wider as the sampling rate decreases. As the relative filter bandwidth increases, the numbers of taps proportionately decreases. For example, when the sampling rate of a filter is decreased by half, the relative bandwidth of the filter doubles, and the number of taps needed to implement it is also halved.

We exploit the unique structure of the audiometric filter bank to map each frequency octave to a sampling rate. The audiometric channelizer is a half-octave filter bank spanning a frequency range of about 5 octaves, from 250 Hz to 8000 Hz. An octave is a logarithmic unit

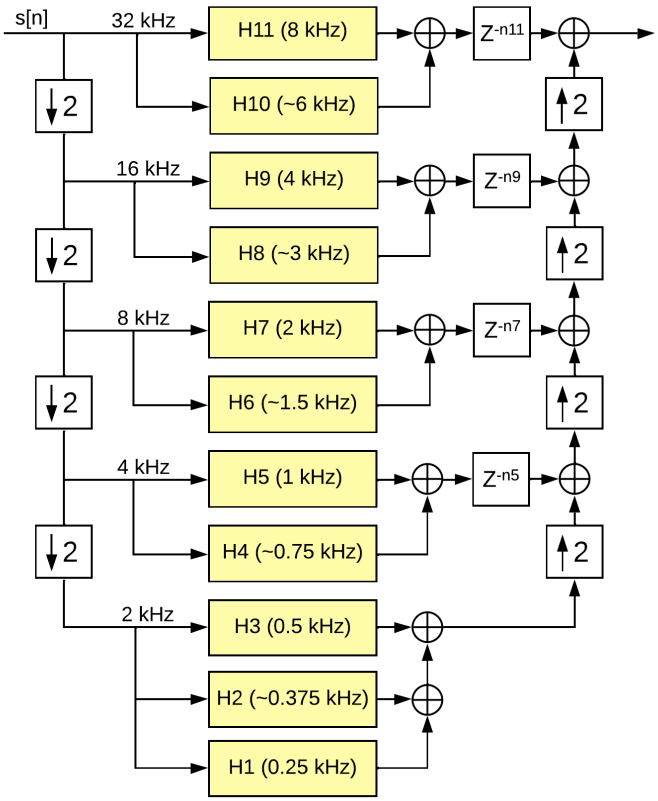

Figure 2.2. Multirate filter bank structure

defined as the difference between two frequencies separated by a factor of two, and a half-octave is the difference between two frequencies separated by a factor of $\sqrt{2}$. Thus, a half-octave filter bank is binary logarithmic, and the bandwidth of any two filters an octave apart differs by a factor of two.

As such, we are able to map each octave of the channelizer to a different sampling rate. We start by designing two bandpass filters at the original sampling rate that span one octave. The next two filters are one octave below, are half as wide, and would require double the number of taps. However, if we lower the sampling rate of the lower octave, the number of taps would decrease by half, resulting in filters of the same length as the ones we started with. Following this pattern, we are able to design all the filters in the audiometric channelizer using the same number of coefficients for each filter.

Table 2.1 compares a single-rate versus a multirate implementation of the channelizer. In the single-rate case, as the bandwidth of the filters is halved for every octave, the number of filter coefficients doubles for every octave. However, in the multirate implementation, we retain constant filter complexity because the decrease in a filter's bandwidth is compensated by a decrease in the sampling rate. (The 8 kHz band is an exception because it is a highpass rather than a bandpass filter.)

Figure 2.2 shows the conceptual block diagram of the audiometric filter bank. First the input signal is split into five sampling rates. Then the inputs are passed through the bandpass filters. Lastly, the outputs are restored to the original sampling rate using upsamplers. Delays following the upsamplers are used to compensate for latency disparity among bands.

The five different sampling rates used in the channelizer are represented with dotted vertical lines in Figure 2.1. According to the Nyquist Theorem, for any given sampling rate f_s, the only frequencies that can be observed are those lying between $-f_s/2$ and $+f_s/2$. Thus, each line represents the frequency limit of each different sampling rate. For the purposes of space, however, the original sampling rate, spanning $-f_s/2$ to $+f_s/2$, is not explicitly shown in Figure 2.1.

According to the Nyquist theorem, any frequency band which lies to the left of a dotted line can be processed at that respective sampling rate without aliasing distortion. However, resamplers are not ideal, and require constraints on overlapping transition bandwidths.

2.2.3 Resampling

Conventionally, downsampling is performed by passing a signal through an antialiasing filter, and then decimating it. Similarly, conventional upsampling is performed by zero-packing a signal, and then passing it through an interpolating filter. These operations have suboptimal efficiency, because a conventional downsampler discards filtered samples at the output of the antialiasing filter, and an interpolating filter processes samples with known zero values.

We reduce the complexity of the resamplers by employing polyphase filtering [27].

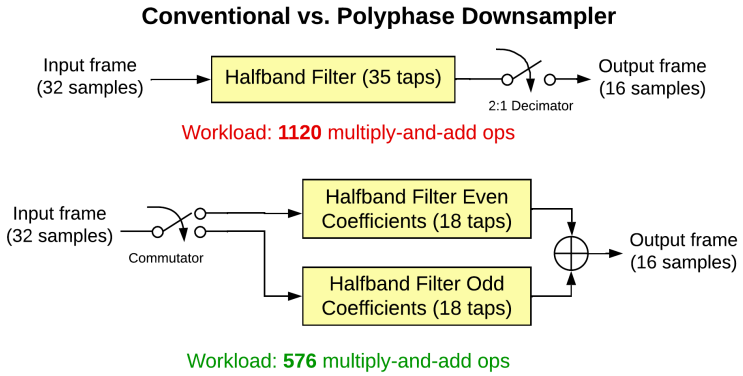

Figure 2.3. A comparison between a conventional 1:2 upsampler and an equivalent polyphase implementation.

Conventional resamplers perform many redundant computations, such as computing samples which will be discarded, or computing samples which are known to be zero. Polyphase filtering eliminates these redundant computations by splitting a single filter into multiple paths and employing the Noble identity to rearrange filtering and resampling. Figure 2.3 compares a conventional and a polyphase 2:1 downsampler. Polyphase resamplers always perform filtering at the lower of their input/output rate, and reduce the complexity of resampling by approximately a factor of M, where M is the resampling ratio. The use of polyphase resamplers reduces the computational costs of sampling rate adjustment by nearly half.

2.2.4 Power

We estimate the cumulative power consumption of the filter bank by computing the total number of multiply-and-accumulate operations per one output sample. For a filter running at a single sampling rate, the number of operations per sample is simply equal to the number of filter taps. However, in a multirate system, samples are continuously removed and added, which makes it impossible to match an input sample to a single output sample. As such, we compute the number of operations per sample of the multirate channelizer by calculating the total number of operations per input frame, and then normalizing by the input frame size. For each stage of the filter bank, we track the current frame size and the cumulative operations count. Due to the

Table 2.2. The cumulative number of multiply-and-add operations per sample of the audiometric filter bank, with and without multirate processing

Filter Band:	**Operations per sample:**		**Ratio:**
	Single-rate	Multirate	
8 kHz	53	53	1x
6 kHz	77	77	1x
4 kHz	154	74.5	2.07x
3 kHz	154	56.5	2.73x
2 kHz	308	43.25	7.12x
1.5 kHz	308	34.25	8.99x
1 kHz	616	26.63	23.14x
0.75 kHz	616	22.13	27.84x
0.5 kHz	1232	18.31	67.28x
0.375 kHz	1232	16.06	76.7x
0.25 kHz	1232	16.06	76.7x
Total:	**5982**	**437.69**	**13.67x**

multirate structure of the channelizer, normalization by frame size results in a fractional number of operations per sample.

Table 2.2 compares the total number of multiply-and-accumulate operations per sample for a single-rate and multirate implementation of the channelizer. The multirate operations estimate accounts for all filters and resamplers. Our evaluations show that compared to a conventional approach, the multirate filter bank offers 13.7$\times$ improvement in complexity. For a wearable battery-operated system, power consumption and processing capabilities are of critical importance. Reducing the number of operations improves battery-life and frees processing power for other tasks.

2.2.5 Latency

As seen from Figure 2.2, different frequency bands follow different signal paths and as such, experience varying amounts of delay. Because of the resamplers and lower sampling rates, lower frequency bands incur more delay than higher frequencies. The highest frequency bands (8 kHz and 6 kHz) experience only a few milliseconds of delay. However, the 0.5 kHz, 0.375 kHz, and the 0.25 kHz bands experience over 30 milliseconds of latency. This disparity causes a

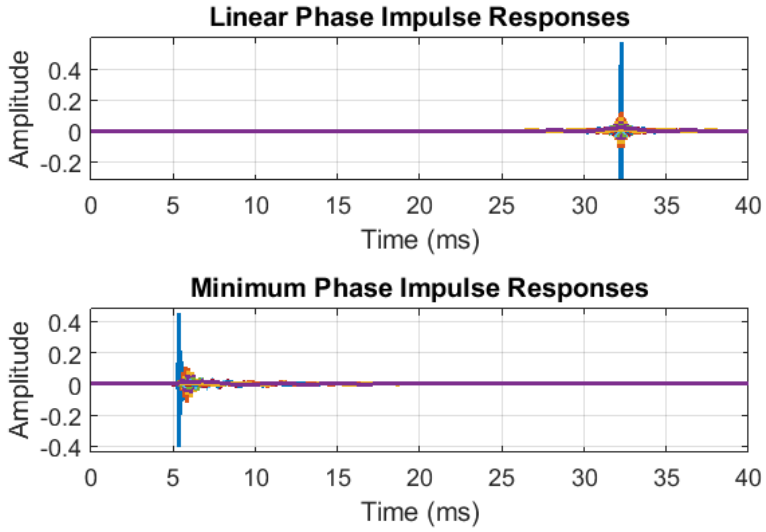

Figure 2.4. A comparison between a linear phase implementation (top) and a minimum phase implementation (bottom) of the Audiometric Filter Bank.

phase offset among the eleven bands, and causes distortion in the composite frequency response. To certain listeners, this phase disparity sounds like an echo or a distorted sound timbre.

In order to eliminate this latency disparity, we realign the bands by inserting delays into the signals paths, as seen in Figure 2.2, such that higher frequency bands are delayed until the lowest frequency bands arrive. Figure 2.4 (top) shows the aligned impulse responses of the filter bank. Although the solution above preserves perfect reconstruction, the latency far exceeds real-time operation requirements. Conventionally, the latency limit for a real-time hearing aid is considered to be 10 milliseconds [16]. As seen from Figure 2.4 (top), the latency of the aligned channelizer is about 32 milliseconds. We resolve this issue by converting the filters from linear phase to minimum phase. A minimum phase filter has the same magnitude response as a linear phase filter, but the lowest possible delay. A filter can be converted from linear phase to minimum phase by reflecting all roots which lie outside the unit circle.

Figure 2.4 (bottom) shows the aligned impulse responses of the minimum phase filter bank. As seen from Figure 2.4, converting the filters from linear to minimum phase dramatically decreases the delay of each band. While retaining the same functionality as a linear phase filter

bank, the minimum phase filter bank has a latency of only 5.4 ms, compared to 32 ms, which makes it suitable for real-time applications.

2.3 Experimental Results

2.3.1 Implementation Testbed

We have integrated our proposed filter bank into the Open Speech Platform (OSP) developed at University of California San Diego [1, 11, 12]. OSP is an open source suite of hardware and software tools for conducting research into many aspects of hearing loss both in the lab and the field. The hardware system consists of a battery operated wearable device running a Qualcomm 410c processor, similar to those in cellphones, with two ear-level assemblies attached – one for each ear. More details about the hardware systems can be found in [1].

At the core of OSP software is the real-time Master Hearing Aid (RT-MHA) reference design. Initially, the incoming audio signal from the microphones is sampled at 48 kHz, and is then downsampled to 32 kHz (not to be confused with the resamplers present in the channelizer). The audio signal is then routed to the channelizer. The outputs of the channelizer then pass through the Wide Dynamic Range Compression (WDRC) unit to compensate for the user's hearing loss. Then the outputs of the WDRC are recombined and passed through a Global Maximum Power Output (MPO) controller in order to limit the power outputted by the speaker. Finally, the audio is upsampled from 32 kHz back to 48 kHz and outputted through the speakers. Additionally, the RT-MHA reference design contains Adaptive Feedback Cancellation (AFC) in order to compensate for the feedback arising from the close proximity of the microphone and the speaker. More detailed explanations of the RT-MHA components can be found in [1, 11].

We evaluated the performance of our proposed filter bank by running the OSP RT-MHA software on the hardware system for 15 minutes. This test was repeated twice – once using the proposed eleven band multirate filter bank, and the second time using an older 6-band filter bank implementation [1]. For each design, we measured the amount of time it took to process 1

Table 2.3. Real-time processing performance statistics for the 6 Band filter bank implementation [1] and 11 Band proposed implementation. (Total time taken to perform each processing step on a 1 ms audio buffer.)

	6 Band [1] (μs)		11 Band (μs)	
	Mean	STD	Mean	STD
Down Sampling	14.77	1.23	14.64	1.11
Filter Bank	175.51	3.49	108.55	4.55
WDRC	60.12	2.09	100.80	3.21
Global MPO	11.48	0.71	11.49	0.80
AFC	131.72	2.78	131.98	2.97
Up Sampling	14.57	1.03	14.73	1.21
Overall HA*	510.82	64.91	487.81	64.92

*The measured total time includes overheads like audio processing callback, including sending work to the threads for the left and right channels and waiting for them to complete.

millisecond of audio data. As long as all the processing is completed within 1 millisecond before the next frame of data arrives, the system can operate in real-time.

Table 2.3 shows the mean and standard deviation of the processing times of each component for both MHA designs. Our experiments show that OSP running with our proposed filter bank meets the real-time processing deadline of 1 millisecond. Moreover, the 11-band channelizer outperforms the 6-band channelizer by about 65 microseconds, offering more frequency bands for less processing time. Most other stages of the MHA reference design were unaffected by the transition from 6 bands to 11 bands, with the exception of WDRC. However, despite the increased workload for the WDRC unit due to an increased number of channels, the overall processing time for the 11-band design is statistically similar to that of the 6-band design.

2.3.2 Verifit Verification Toolbox

We evaluated the design using the widely accepted Audioscan Verifit 2 Professional Verification system [28]. Verifit 2 is a verification tool consisting of a soundproof binaural audio chamber, a display unit, and a set of powerful testing procedures, such as speech map, ANSI tests, and distortion.

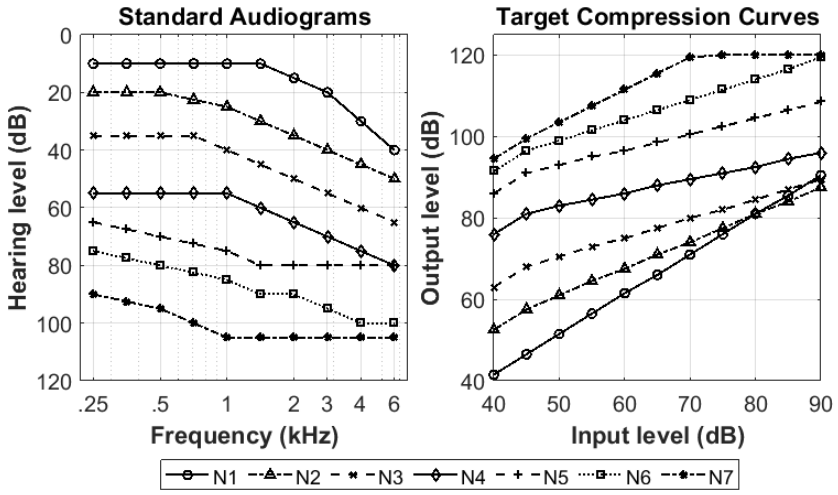

Figure 2.5. Standard Audiograms for hearing aid testing developed by the ISMADHA group (left); The corresponding target compression curves at 1000 Hz (right).

We conducted steady state input-output measurements to evaluate the multirate amplification system running on Open Speech Platform hardware. The purpose of this test is to compare the experimentally measured input-output curve of our device to the ideal target curve specified by a hearing loss prescription. In this experiment, the hearing aid device is placed into the soundproof audio chamber. The Verifit's reference speaker plays calibrated audio signals with known acoustical properties into the hearing aid microphone, which becomes the input signal for the hearing aid. The processed output signal of the hearing aid is then collected by the Verifit's coupler microphone and is compared to the input signal to identify the measured gain.

We verified our system using seven standard pure tone audiograms developed by the International Standard for Measuring Advanced Digital Hearing Aids (ISMADHA) group [29], which represent a broad class of hearing loss patterns, from very mild to profound. We obtained compression parameters by passing a subset of ISMADHA profiles through the NAL-NL2 Prescription Procedure [30], which is a widely accepted algorithm for generating hearing aid prescriptions from pure tone audiograms. Figure 2.5 shows the ISMADHA standard pure tone audiograms, and an example of the obtained target input/output amplification curves for each audiogram at 1 kHz.

Table 2.4. Maximum and average wide dynamic range compression steady state error for seven standard hearing loss profiles

No.	Category	Max (dB)	Average (dB)									
			250	**354**	**500**	**707**	**1000**	**1414**	**2000**	**2828**	**4000**	**5657**	**8000**
N1	Very Mild	**0.5**	0.0	-0.2	0.0	-0.2	0.0	0.0	-0.1	-0.1	-0.3	0.0	-0.2
N2	Mild	**1.0**	-0.1	-0.4	0.1	0.0	-0.1	0.0	-0.1	-0.1	-0.3	0.2	0.1
N3	Moderate	**1.5**	-0.1	-0.3	0.1	0.0	0.1	0.1	0.0	0.0	-0.1	0.0	0.3
N4	Mod/Severe	**2.0**	0.0	0.5	0.1	0.5	0.5	0.5	0.4	-0.1	0.0	0.0	0.5
N5	Severe	**1.5**	-0.1	-0.1	0.1	0.0	0.0	0.3	0.1	-0.1	0.0	-0.1	0.2
N6	Severe	**2.0**	-0.1	-0.1	0.2	0.0	0.2	0.3	0.2	-0.2	-0.1	-0.2	0.4
N7	Profound	**3.0**	0.2	-0.3	0.0	0.1	0.0	0.2	-0.1	-0.4	0.2	-0.1	0.3

We performed steady state measurements at the eleven half-octave frequencies offered by the audiometric filter bank. For each frequency, we obtained the target compression curves, such as the ones shown in Figure 2.5. We then took measurements for each combination of audiogram, frequency, and input level, resulting in 847 data points. Table 2.4 shows the maximum and average errors we obtained for each audiogram as a function of frequency. Our results show that the compressed output values closely match the desired target values, often with 0 dB average error. The maximum error (usually found in the MPO region) is also small, and never exceeds 3 dB, which was shown to be the threshold of just noticeable difference in speech-to-noise ratio [31].

2.3.3 Comparison with Prior Work

We compared our proposed audiometric filter bank against the filter bank from the widely popular Kates digital hearing aid [25]. Kates' hearing aid is the common go-to system for hearing aid research. Our proposed filter bank meets and improves upon the capabilities of the Kates system.

Figure 2.6 presents our proposed channelizer and the Kates filter bank on a logarithmic scale. Our proposed filter bank offers four more bands than Kates' filter bank. These additional bands provide our channelizer with higher spectral resolution, which makes it possible to more accurately fit a hearing aid to a particular hearing loss prescription.

Within each band, our filter bank also offers higher band integrity, especially at lower

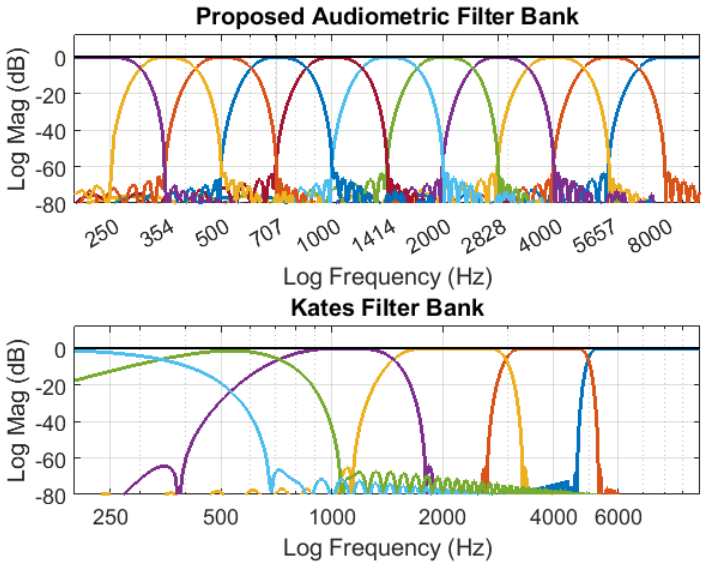

Figure 2.6. The magnitude responses of the multirate audiometric filter bank versus the Kates Filter Bank.

frequencies. The lower frequency bands of Kates's filter bank have very wide side-lobes, which introduces cross-talk between bands. According to the ANSI Standard on Half-Octave Filters [26], the 500 Hz band of the 6-Band filter bank does not meet class 2 criteria. All the filters in our proposed channelizer meet class 0 criteria.

Across-band cross-talk plays a detrimental role in the ability of a hearing aids algorithm in faithfully fulfilling a hearing aid prescription. Based on a user's hearing loss pattern, determined using pure-tone audiometry [17], a hearing aid prescription determines the desired output levels for various signal input levels at different audiometric frequencies. When cross-talk from neighboring bands interferes with any given band, the hearing aid will no longer match the desired output levels outlined in the prescription.

We used the Verifit's input-output curve feature to compare the prescription accuracy of the multirate eleven-band system versus the Kates system. The Verifit tool generates tones of increasing magnitude, indicated on the x-axis in units of dB SPL. It then records and plots the steady state output on the y-axis, forming the WDRC input-output curves for both systems at 250, 500, 1000, 2000, and 4000 Hz. Figure 2.7 shows the static behavior of the six band and

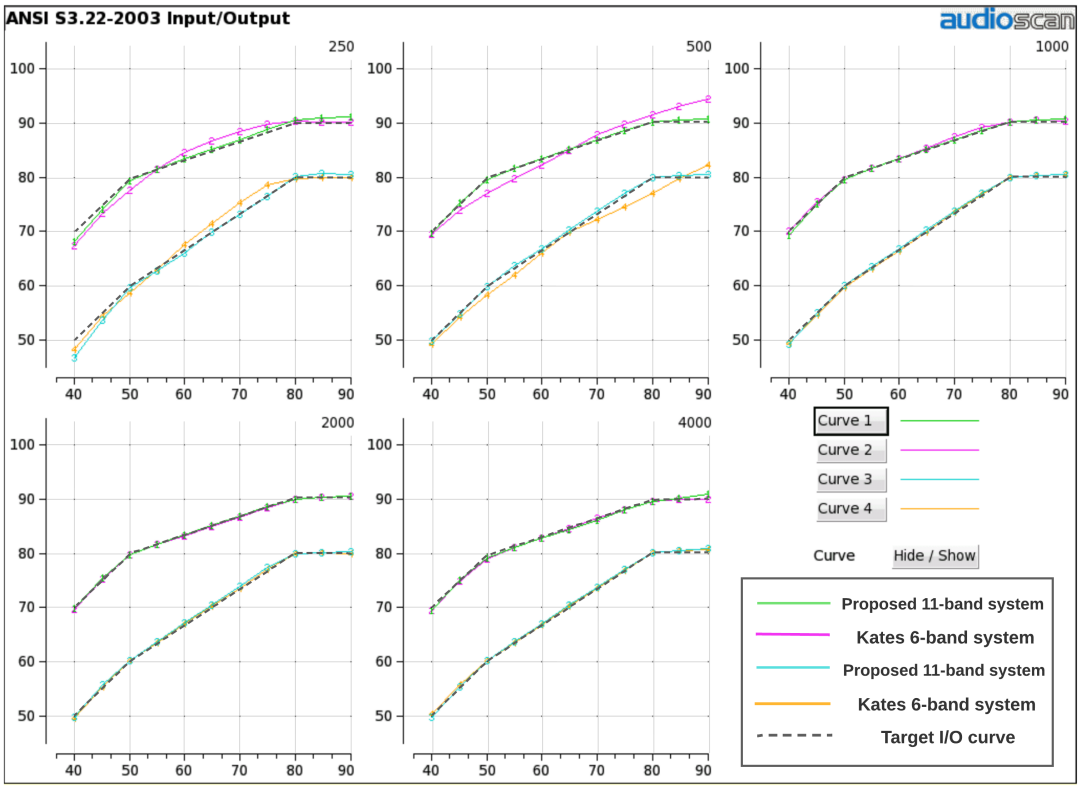

Figure 2.7. Verifit Verification Toolbox measurements comparing the steady state behavior of the multirate 11-band system and the Kates 6-band system.

eleven band realizations. The black dotted lines show the target output magnitudes, while the colored lines show real measured output magnitudes for the 6-band and 11-band systems. At higher frequencies, both realizations accurately fulfill the target prescription. However, starting from 1000 Hz, the Kates implementation begins to diverge from the target curve, and both the 250 and 500 Hz bands lose their shape integrity. This is due to the high side lobes of the of low frequency bands seen in Figure 2.6.

We also compared the Kates filter bank and our multirate filter bank in terms of computational complexity, estimated by the number of operations per output sample, and latency. Table 2.5 compares the complexity and latency of the Kates filter bank and the eleven band filter bank. In addition to offering almost twice the number of bands compared to Kates's filter bank, the proposed filter bank achieves about 3.5$\times$ improvement in complexity, with a comparable algorithmic latency of 5.43 ms.

Table 2.5. Complexity and latency comparison between the Multirate Audiometric Filter Bank and Kates Filter Bank

Filter Bank	Bands	Operations per sample	Latency
Proposed OSP Filter Bank	11	437.69	5.43 ms
Kates Filter Bank	6	1542	4.03 ms

2.4 Conclusion

In this chapter, we presented an eleven band multirate filter bank for audiology and hearing aids research. Our filter bank is tailored to the needs of the audiology community, which requires a channelizer that reflects the nature of the human cochlea, which has frequency-dependent spectral resolution. Our proposed filter bank targets the audiometric frequencies used for characterizing hearing loss in pure tone audiometry [17]. The individual bandpass filters in our channelizer offer at least 75 dB of sidelobe attenuation, and the composite response offers near perfect signal reconstruction with no more than 0.01 dB ripple in the pass-band. In order to lower the power consumption of the very narrow filters present in our channelizer, we employed multirate signal processing, which resulted in a 14$\times$ reduction in complexity compared to single-rate processing. We integrated our channelizer into Open Speech Platform [10], and verified that the Master Hearing Aid algorithm running our filter bank meets the real-time processing deadline and outperforms the previous channelizer used in the Open Speech Platform.

2.5 Acknowledgements

This work is supported by the National Institute on Deafness and Other Communication Disorders (NIH/NIDCD) grants R21DC015046, R33DC015046, "Self-fitting of Amplification: Methodology and Candidacy," and R01DC015436, "A Real-time, Open, Portable, Extensible Speech Lab," awards to University of California, San Diego; grant IIS-1838830 from the Division of Information & Intelligent Systems, "A Framework for Optimizing Hearing Aids In Situ Based on Patient Feedback, Auditory Context, and Audiologist Input."; San Diego State University

“fred harris Excellence Scholarship; Qualcomm Institute at UCSD; Halıcıoğlu Data Science Institute at UCSD, and Wrethinking, the Foundation.

Chapter 2, in part, is a reprint of the material as it appears in A. Sokolova, D. Sengupta, M. Hunt, R. Gupta, B. Aksanli, f. harris, and H. Garudadri, “Real-Time Multirate Multiband Amplification for Hearing Aids”. IEEE Access, 2022, and A. Sokolova, D. Sengupta, K.L. Chen, R. Gupta, B. Aksanli, f. harris, H. Garudadri. ”Multirate Audiometric Filter Bank for Hearing Aid Devices”, Asilomar Conference on Systems, Signals, and Computers, 2021. The dissertation author was the primary investigator and author of this material.

Chapter 3

Automatic Gain Control for Wide Dynamic Range Compression

3.1 Introduction

Wide Dynamic Range Compression (WDRC), the process of reducing the dynamic range of an audio signal, is used in nearly all modern hearing aids [7]. However, there are numerous issues with compressive hearing aids leading to user dissatisfaction [6], which include complaints about speech intelligibility, sound pleasantness, distortion, and excessive amplification of background noise.

The functionality of WDRC is to sense the input magnitude of the signal, compute the desired gain based on the input-output transfer function, and to apply the gain to the input signal. The time-varying amplification of a signal based on a desired reference level is known as Automatic Gain Control (AGC). The performance of WDRC can be evaluated in terms of static and dynamic behavior, where static behavior is measured as the steady-state response of the AGC system to a constant input, and dynamic behavior measures the system's response to a change of input. The static behavior of WDRC is defined by parameters if the input-output transfer function (ex. $g65$, CR, $knee_{low}$), while dynamic behavior is defined by the attack and release time parameters.

Compression in hearing aids is typically categorized qualitatively as fast-acting (less than 10 ms attack time and 100 ms release time), slow-acting (more than 10 ms attack time and

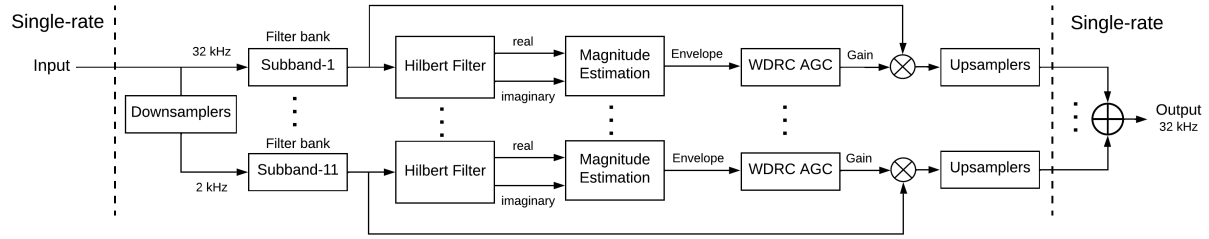

Figure 3.1. A block diagram of the multirate multiband hearing aid amplification system.

over 1000 ms release time), or in-between, and there is a great deal of inconsistency in literature about the benefits of each approach [7, 8]. One reason for the lack of consensus in setting the response times of WDRC is the inconsistent performance of hearing aids produced by different manufacturers. ANSI standard S3.22-2009 [18] provides a definition for measuring attack and release times in hearing aids, but current hearing aids do not accurately satisfy ANSI standards.

In this chapter, we present an Automatic Gain Control system for WDRC which addresses the need for finer, more precise gain control in a hearing aid device. Our design provides the audiology research community with tools which offer higher flexibility and accuracy than currently available on open-source platforms. Our Multirate Automatic Gain Control system for WDRC accurately fulfills the static and dynamic properties specified by audiologists, which include steady state gain, as well as attack and release times.

The block diagram in Figure 3.1 shows an overview of the proposed subband amplification system. This system accepts an audio signal sampled at 32 kHz, performs frequency decomposition on the signal, and transitions from single to multirate processing. The system then computes the gains necessary for Wide Dynamic Range Compression in each band. The final stage converts all multirate outputs back to the original sampling rate and combines the bands into a final output. Multirate processing is a key feature of our design, and is instrumental in ensuring real-time operation of the system and reducing power consumption.

The multirate amplification system is implemented and tested on the Open Speech Platform (OSP) – an open source suite of software and hardware tools for performing research on

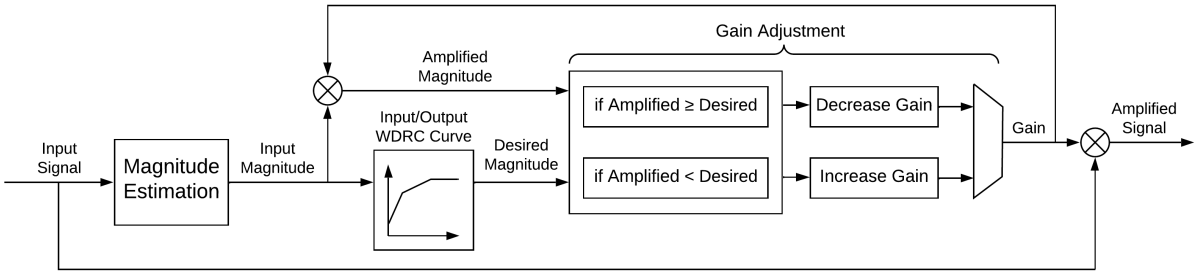

Figure 3.2. A block diagram of automatic gain control for WDRC in hearing aids.

emerging hearing aids and hearables. The OSP suite includes a wearable hearing aid, a wireless interface, and a set of hearing enhancement algorithms [1, 10, 11, 12].

3.2 Wide Dynamic Range Compression

3.2.1 Overview

WDRC is a type of automatic gain control (AGC) system which reduces the dynamic range of audio by applying varying gain to a signal depending on the input magnitude. For any instantaneous input magnitude, the WDRC input-output transfer function determines the desired instantaneous output magnitude. Based on the desired output magnitude, the Automatic Gain Control algorithm adjusts the time-varying gain of the signal. A block diagram of the AGC system is shown in Figure 3.2.

It is insufficient, however, to set the gain of each audio sample independently. Studies in acoustics and speech intelligibility have shown that the rate of change of WDRC gain has a strong effect on speech clarity and legibility [32, 33]. The rate of change of gain is measured using the attack and release times, which play a key role in the performance of WDRC. However, to the best of our knowledge, currently available open-source hearing aids do not have an accurate mechanism for setting attack and release times independently of other parameters. For example, the attack and release times of the Kates system [25] depend on the user-defined compression ratio, which gives rise to major inaccuracies.

In this chapter, we explore the complex relationship between the attack and release

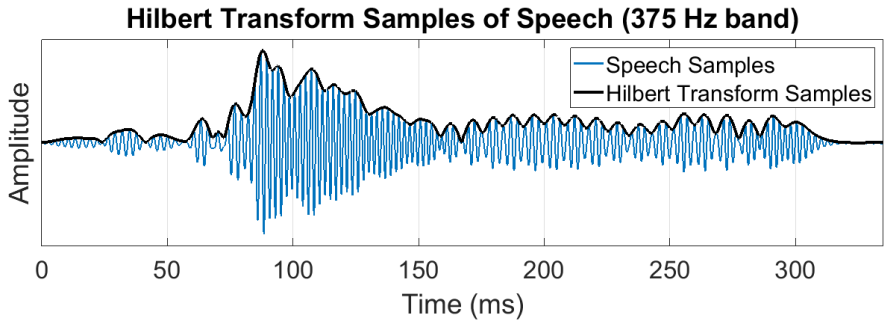

Figure 3.3. The waveform and computed envelope of the word ”please” in the 375 Hz band, spoken by a female voice.

times of WDRC and the parameters defining a WDRC curve. We also propose a multirate compression algorithm which yields precise response times in accordance with ANSI standards for any user-defined WDRC parameters.

3.2.2 Magnitude Estimation

Wide Dynamic Range Compression calculates compression gains based on input magnitude. However, sound is a modulating signal, meaning the magnitude of the signal is contained in the envelope. Common approaches to finding the envelope of a modulating signal include peak detection [25], per-frame total power [34], sliding RMS windows, and more. However, all of these approaches introduce inaccuracies into the envelope estimate, such as ripple or excessive smoothing.

We estimate the signal envelope by employing the Hilbert Transform. The Hilbert Transform accepts a real signal, and computes a 90-degree phase shifted imaginary component. The magnitude of the input signal is then found as the absolute value of the real and imaginary components.

The accuracy of the Hilbert Transform depends on the accuracy of the underlying Hilbert Filter, which is a filter that cuts off the negative frequencies of the signal spectrum. If the transition bandwidth of the Hilbert Filter overlaps with signal content, then the computed envelope becomes distorted.

The low-frequency channels used in WDRC are very close to DC (see Chapter 2 for more details), and preserving these frequencies would require an unrealistically sharp Hilbert Filter. However, we prevent distortion in the low-frequency bands by performing magnitude estimation and amplification in the multirate domain, as shown in Figure 3.1. As we discussed earlier, reducing the sampling rate of a filter increases its relative width. However, for a given center frequency, reducing the sampling rate of the signal also moves said center frequency relatively farther from DC. As such, the channel is no longer affected by the Hilbert Filter's transition bandwidth.

The multirate Hilbert Transform produces highly accurate signal envelopes for all frequency channels of the filter bank. Figure 3.3 shows the 0.375 kHz band of the word "please" spoken by a female voice from the TIMIT database [35], as well as the envelope of the waveform computed using the Hilbert Transform.

3.2.3 Proposed Automatic Gain Control

The ANSI Specification of Hearing Aid Characteristics defines the attack and release times for hearing aid devices [18]. Given a step input which changes magnitude from 55 dB to 90 dB, as shown in Figure 3.5, the attack time is defined as the time elapsed between the step change and the time the output remains within 3 dB of its steady state value, notated as A_2 in Figure 3.5. Release time is similarly defined as the time elapsed between a step change from 90 dB to 55 dB, and the time the output remains within 4 dB of steady state, notated as A_1. The steady-state values are obtained from the WDRC curve, shown in Figure 3.4, and as such, depend on compression parameters.

The general concept of Automatic Gain Control for WDRC, illustrated in Figure 3.2, is to decrease the gain when the output overshoots, and increase the gain when the output undershoots. However, since the steady state values A_1 and A_2 shown in Figure 3.5 depend on user parameters, the overshoot and undershoot also depend on user compression parameters. Thus, there is a relationship between user input parameters and the response speed of an AGC loop which is not

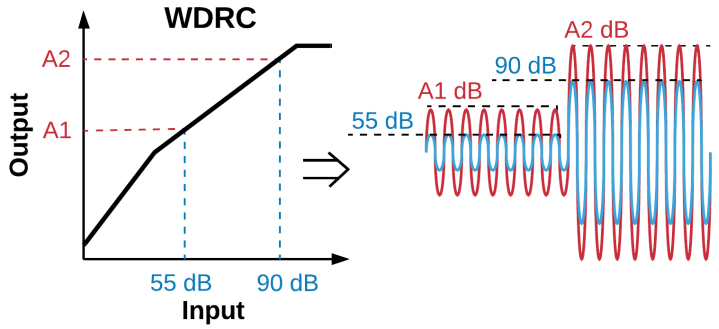

Figure 3.4. Specifications for measuring ANSI attack and release times (steady-state response).

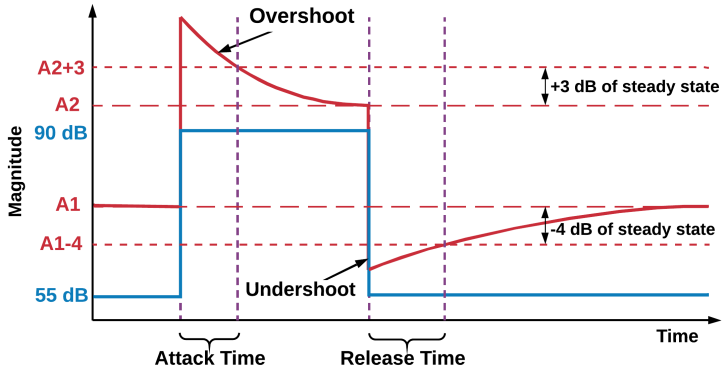

Figure 3.5. Specifications for measuring ANSI attack and release times (step response).

well explored in modern hearing aids and leads to significant error in actual attack and release times compared to desired values.

We derived a closed-form relationship between static compression parameters and the attack and release times of a hearing aid, and designed an Automatic Gain Control (AGC) loop which yields exact attack and release values for any user-defined compression parameters. Our design builds upon work in [36] by adapting radio AGC to Wide Dynamic Range Compression. The block diagram of the proposed AGC algorithm is shown in Figure 3.6. For each input sample, the gain of the previous sample is added to the current sample. The sum is then compared to the desired output level based on the WDRC curve. The scaled difference between the desired and the actual output levels is then used to modify the gain of the next sample. In the AGC loop, alpha (α) is an important scaling parameter which determines how quickly the system reacts to changes. As such, α is the only parameter determining the attack and release times of the AGC

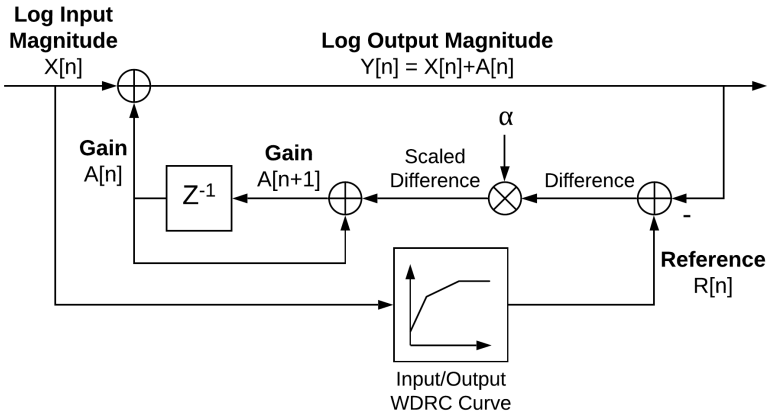

Figure 3.6. The proposed algorithm for automatic gain control. Response times are controlled by parameter α.

loop. Since WDRC must respond differently to rising and falling input levels, the AGC loop requires two distinct values of α – one for attack time, one for release time.

In this section, we derive the relationship between α and WDRC parameters such that the system yields exact attack and release times in any configuration. The behavior of the system above is described by the equation below.

$$\begin{aligned} A[n+1] &= A[n] + \alpha \times (R[n] - Y[n]) \\ &= A[n] + \alpha \times (R[n] - (X[n] + A[n])) \\ &= A[n] \times (1-\alpha) + \alpha \times (R[n] - X[n]) \end{aligned} \tag{3.1}$$

Consider the ANSI test signal, which is a step input which changes magnitude from 55 dB to 90 dB at time $n = 0$. Let us define G_0 as the initial steady state gain before the step change. For $n < 0$, $R[n] = A1$, $X[n] = 55$, so $G_0 = R[n] - X[n] = A1 - 55$.

Let us define G_∞ as the final steady state gain after the step change. For all times $n \geq 0$, $R[n] = A2$, $X[n] = 90$, so $G_\infty = R[n] - X[n] = A2 - 90$. Using these definitions, for all $n \geq 0$, equation 1 can be rewritten as:

$$A[n+1] = A[n] \times (1-\alpha) + \alpha \times G_\infty \tag{3.2}$$

In order to gain insight into the behavior of the system, let us write out the gains of the first few samples:

$$A[0] = G_0 \tag{3.3a}$$

$$A[1] = G_0 \cdot (1-\alpha) + \alpha \cdot G_\infty \tag{3.3b}$$

$$A[2] = G_0 \cdot (1-\alpha)^2 + \alpha \cdot G_\infty \cdot (1-\alpha) + \alpha \cdot G_\infty \tag{3.3c}$$

$$\begin{aligned} A[3] = {} & G_0 \cdot (1-\alpha)^3 + \alpha \cdot G_\infty \cdot (1-\alpha)^2 \\ & + \alpha \cdot G_\infty \cdot (1-\alpha) + \alpha \cdot G_\infty \end{aligned} \tag{3.3d}$$

As seen from the pattern formed in equation 3, the gain of the n'th sample is found as a geometric series, shown in equation 4a and simplified in equation 4b.

$$\begin{aligned} A[n] = {} & G_0 \times (1-\alpha)^n + \alpha \times G_\infty \times (1 + (1-\alpha) \\ & + (1-\alpha)^2 + \ldots + (1-\alpha)^{n-1}) \end{aligned} \tag{3.4a}$$

$$\begin{aligned} A[n] &= G_0 \times (1-\alpha)^n + \alpha \times G_\infty \times \left(\frac{1-(1-\alpha)^n}{\alpha}\right) \\ &= G_0 \times (1-\alpha)^n + G_\infty \times (1-(1-\alpha)^n) \\ &= (G_0 - G_\infty) \times (1-\alpha)^n + G_\infty \end{aligned} \tag{3.4b}$$

This important result provides us with an equation for gain as a function of time and α. As expected, at time $n = 0$ the gain is equal to G_0, and as n reaches infinity the gain approaches G_∞.

Using the equation above, we can use known values of n to solve for α. As explained

earlier, α is the only parameter which sets the attack and release times of the AGC system. Let AT represent the attack time. From the ANSI definition of attack time, we know that at time $n = AT$, the gain needs to be within 3 dB of steady state, which is $G_\infty + 3$. Substituting these values into equation 4b yields:

$$G_\infty + 3 = (G_0 - G_\infty) \times (1-\alpha)^{AT} + G_\infty \tag{3.5}$$

The equation above contains only one unknown variable, allowing us to solve for α_{attack}:

$$\begin{aligned} \alpha_{attack} &= 1 - (\frac{3}{G_0 - G_\infty})^{\frac{1}{AT}} \\ &= 1 - (\frac{3}{A1 - A2 + 35})^{\frac{1}{AT}} \end{aligned} \tag{3.6}$$

Following similar steps and using the ANSI definition for release time, we can find a similar expression for $\alpha_{release}$:

$$\begin{aligned} \alpha_{release} &= 1 - (\frac{4}{G_0 - G_\infty})^{\frac{1}{RT}} \\ &= 1 - (\frac{4}{A1 - A2 + 35})^{\frac{1}{RT}} \end{aligned} \tag{3.7}$$

Equations 6 and 7 provide us with values for α_{attack} and $\alpha_{release}$ that guarantee exact attack and release times for the AGC loop. It is important to note that in equation 6 and 7, the units for AT and RT are samples. Samples and milliseconds are related to each other through sampling rates which, as described earlier, varies between the different subbands.

It can be noted that the difference $G_0 - G_\infty$ is none other than the Overshoot pictured in Figure 3.5. The Overshoot is a variable which depends on the parameters setting the WDRC curve. By deriving the relationship between α and Overshoot, we account for all WDRC parameters, including compression ratio, in our calculations for attack and release times.

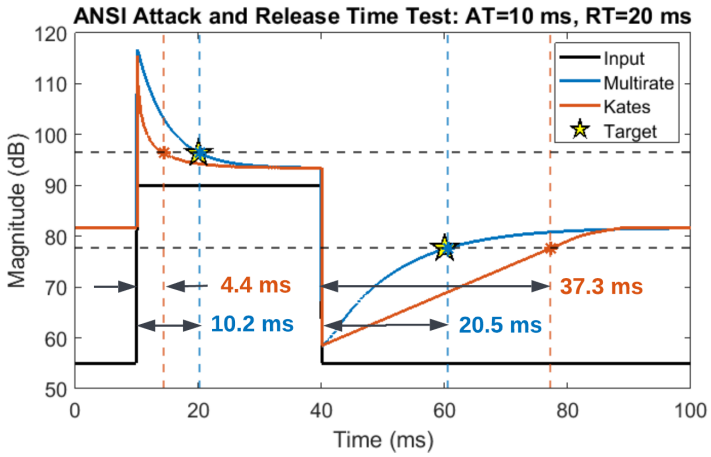

Figure 3.7. ANSI attack and release time test for the Proposed Multirate and Kates's automatic gain control. Asterisks mark the measured attack and release times, and stars mark the ANSI S3.22 target values.

Another feature of the AGC loop, shown in Figure 3.6, is that the reference signal $R[n]$ needs not be a piecewise curve, as shown in Figure 3.4. The piecewise input-output WDRC curve benefits from simplicity, but our system can accept any function for the input-output curve, including smooth continuous functions and 'S' curves. This flexibility allows the user to employ other input-output curves, which may be more appropriate for the user.

3.3 Experimental Results

We compared our Multirate Multiband Automatic Gain Control with Kates's approach. As described Section 3.2.3, the relationship between WDRC parameters and AGC response times are not explored in previous works. In the Kates approach, the AGC response times are controlled by the coefficients of the peak detector used to estimate the signal magnitude. The resulting coefficients are approximated to meet ANSI attack and release time standards, but diverge from target values significantly.

As a test case, Figure 3.7 compares the dynamic responses of the proposed multirate system and the Kates system. The input is a gated sinusoid test signal stepping between 55 and 90 dB, as defined by the ANSI S3.22 standard [18], centered at 2000 Hz. Both systems were

configured to have a compression ratio of 3:1, and the attack and release times were set to 10 ms and 20 ms respectively. The dynamic responses of the two systems are shown in Figure 3.7.

In this experiment, the measured attack and release times of the proposed Multirate system are 10.2 ms and 20.5 ms respectively, which deviate from the target values by 0.2 ms (2%) and 0.5 ms (2.5%). On the other hand, the measured attack and release times of the Kates system are 4.4 ms and 37.3 ms respectively, which is a 5.6 ms (45%) and 17.3 ms (87%) deviation from the target values. This experiment shows that the proposed Multirate system satisfies attack and release times within 0.5 ms of the target value. However, the Kates system yields attack and release time values that significantly diverge from the target. Furthermore, this error is unpredictable because the internal coefficients responsible for attack and release times of the Kates system are designed to be ”fudge” factors.

The proposed Multirate systems offers very accurate fulfillment of user designated attack and release times. However, neither the current standards [18] nor popular HA prescription tools, e.g., [30] provide guidance for the dynamic aspects of dynamic range compression. There is a need for the signal processing and audiology research community to address this important gap and investigate the role that response times play in speech legibility and perception.

3.4 Conclusion

In this chapter, we presented a real-time multirate, multiband Automatic Gain Control (AGC) algorithm for Wide Dynamic Range Compression in hearing aids. Our algorithm offers highly accurate prescription accuracy for both steady state and dynamic conditions. We use the Hilbert Transform to find the instantaneous signal magnitude, which provides higher accuracy than conventional instantaneous power estimation methods. Furthermore, we derived a closed-form relationship between the compression parameters of WDRC, and the attack and release times of the AGC system. Based on this derivation, we designed an AGC feedback loop which yields precise attack and release times as specified by ANSI standards. The accurate fulfilment of

attack and release times in dynamic range compression opens new opportunities for exploring the relationship between response times and hearing impaired users' satisfaction. We implemented the Multirate Multiband Amplification System on Open Speech Platform - an open source suite of hardware and software tools for hearing loss research. The system runs in real-time on a wearable device, and is suited for hearing loss research both in the lab and in the field.

3.5 Acknowledgements

This work is supported by the National Institute on Deafness and Other Communication Disorders (NIH/NIDCD) grants R21DC015046, R33DC015046, "Self-fitting of Amplification: Methodology and Candidacy," and R01DC015436, "A Real-time, Open, Portable, Extensible Speech Lab," awards to University of California, San Diego; grant IIS-1838897 from the Division of Information & Intelligent Systems, "A Framework for Optimizing Hearing Aids In Situ Based on Patient Feedback, Auditory Context, and Audiologist Input."; Qualcomm Institute at UCSD; Halıcıoğlu Data Science Institute at UCSD, and Wrethinking, the Foundation."

Chapter 3, in part, is a reprint of the material as it appears in A. Sokolova, D. Sengupta, M. Hunt, R. Gupta, B. Aksanli, f. harris, and H. Garudadri, "Real-Time Multirate Multiband Amplification for Hearing Aids". IEEE Access, 2022. The dissertation author was the primary investigator and author of this material.

Chapter 4

A Curvilinear Transfer Function for Wide Dynamic Range Compression with Expansion

4.1 Introduction

Wide Dynamic Range Compression (WDRC), the process of reducing the dynamic range of an audio signal, is a staple of most modern digital hearing aids. Studies have shown that compressive hearing aids provide greater loudness comfort, greater audibility of low-level speech signals and, in certain configurations, greater intelligibility [7, 8]. However, many users continue to be dissatisfied with many aspects of hearing aids [6, 5], and hearing aid adoption has seen little change over twenty years of observation, with only about one in four individuals with hearing loss adopting hearing aids [4].

One important adverse effect of compression on noisy signals is the reduction of long-term output Signal-to-Noise Ratio (SNR) relative to the input SNR. This reduction is caused by the property that low-level noise receives greater amplification than the higher-level signal of interest, resulting in unequal amplification. This effect has been well-documented empirically [37, 38, 39, 40], as well as analytically [41]. An additional side effect of unequal amplification can be the excessive audibility of microphone, circuit, and low-level ambient noise, as observed in [6, 5, 4].

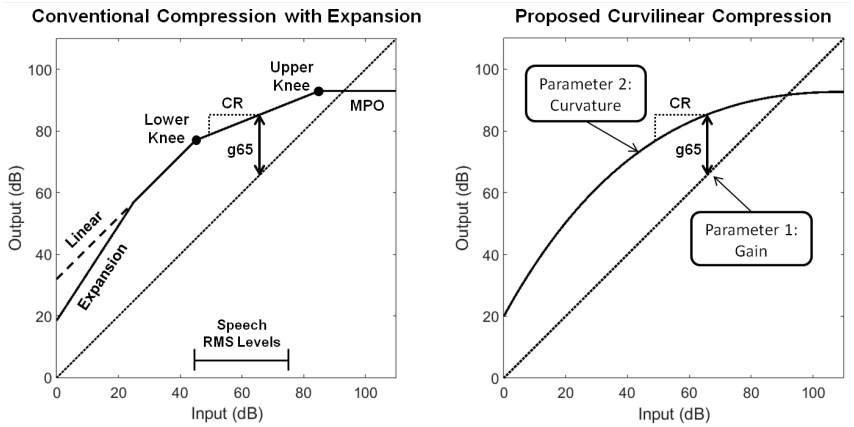

Figure 4.1. (Left) The conventional piece-wise WDRC transfer function with governing parameters; (Right) The reparameterized curvilinear compression transfer function.

One solution for improving the output SNR and reducing the noisiness of a hearing aid is the use of expansion [42, 43, 44]. Expansion, as shown in Fig. 4.1 (left), is the counterpart to compression, wherein gain increases with input levels. Expansion was proposed as a method of improving patient satisfaction by reducing the amplification of ”noisy” low-intensity sounds [45], however, research in this area has been hindered by the historical lack of expansion features on commercial hearing aids. Some early studies in compression included experiments in expansion [42, 43], but due to technology limitations, both yielded mostly negative results. More recent studies in [44] found that expansion can significantly improve perceptual sound quality, but at the same time can also degrade speech intelligibility due to lesser amplification of quiet sounds. Expansion is a common feature on modern commercial hearing aids, but there is a lack of research on how to best to utilize this approach.

This work builds upon an alternative input-output transfer function for Wide Dynamic Range Compression [46]. The transfer function of WDRC has remained largely unchanged since analog amplifiers, and is typically implemented as a piece-wise curve shown in Fig. 4.1 (left), defined by linear segments adjoining at the upper and lower knee points. However, in modern digital hearing aids it is possible to define any arbitrary transfer function to achieve the desired amplification properties of a hearing aid. In this work, we present an alternative input-output

transfer function for WDRC with the following properties:

- The input-output transfer function is a continuous curve
- The curvature of the continuous transfer function provides expansion for low-level signals
- The function is defined by only two parameters – gain and curvature (compression ratio)

The curvilinear transfer function, shown in Figure 4.1 (right), is a continuous function with logarithm-like behavior, offering compression in the regions of speech, and expansion for low-level signals. The elimination of knee points reduces the perceptual noticeability of compression and improves the fidelity of the signal envelope.

The properties of the re-parameterized transfer function are highly applicable for self-fitted and Over-the-Counter (OTC) hearing aids. Various studies, which are well summarized in [7, 47], describe diverse and often conflicting observations of the effect of WDRC on hearing aid outcomes, such as speech intelligibility, perceived sound quality, listening comfort, and overall satisfaction. Amidst a lack of consensus on how to best configure a hearing aid, most studies agree that the benefits of compression are highly individualistic [8, 47]. In an effort to achieve the best individual fit, modern hearing aids offer a growing number of tunable and adjustable parameters [48], which poses the growing challenge of optimal hearing aid fitting. In response to this challenge, the concept of self-fitted hearing aids is seeing growing acceptance in the audiology community [13, 49, 50], and over-the-counter hearing aids are already appearing on the market.

The structure of this chapter is as follows: The first section discusses the theory and analytical basis of our method. The following section presents experimental results, comparing the proposed method to conventional WDRC in terms of output SNR and HASQI, the Hearing Aids Speech Quality Index [51]. The third section discusses a proposed application of our method to self-fitted and OTC hearing aids. Our experiments show that curvilinear compression improves output SNR by up to 1.0 dB, which has the impact of up to 25%, and HASQI by up to

6.7% compared to conventional compression. The greatest benefits are seen with more severe hearing loss profiles, which tend to have higher compression ratios.

4.2 Theory

Wide dynamic range compression (WDRC) is conventionally implemented as a piece-wise segmented curve, such as the one pictured in Figure 4.1 (left), with a region of hard-limiting above the upper knee point, a region of compression, and a region of linear (constant) gain below the lower knee point. The curve is defined by a set of four parameters – $g65$, Compression Ratio (CR), Maximum Power Output (MPO), and $knee_{low}$, where $g65$ is the gain at an input level of 65 dB SPL. Alternatively, parameters $g65$ and CR can be replaced with $g50$ and $g80$. For hearing aids with expansion capabilities, the expansion threshold and the expansion ratio, XR, are also included.

The functionality of WDRC can be broken down into two distinct operations – amplification of the signal and reduction of dynamic range, a.k.a. compression. We define a continuous function which replicates the intended behavior of WDRC with separable parameters for controlling volume (vertical displacement) and compression (curvature). The function yields linear amplification when compression is disabled, and a smooth logarithm-like curve when compression is enabled. The proposed amplification curve is shown in Figure 4.1 (right), and has the following mathematical form:

$$y = Cxe^{-\alpha x} + b \tag{4.1}$$

Where x is the input level in dB and y is the output level in dB, and the controlling parameters are:

α : Controls the curvature of the function. When α is set to zero, the curve is a straight line.

b : Sets the vertical displacement of the curve, which corresponds to overall volume.

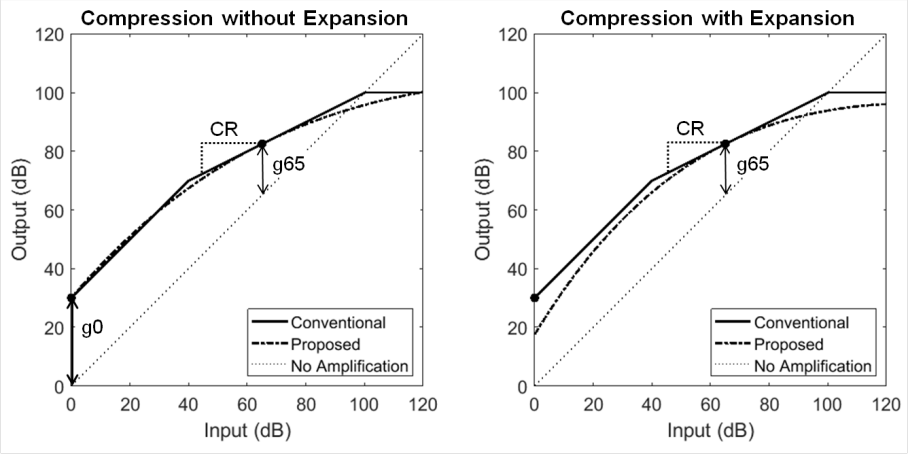

Figure 4.2. (Left) The proposed curve matched to a conventional piece-wise transfer function; (Right) The curvilinear transfer function with expansion.

C : Sets the slope of the curve at the origin, which affects the tilt of the curve.

The proposed amplification function in Eg. (4.1) can be directly substituted for a conventional piece-wise WDRC curve when three parameters are matched – $g65$, CR, and linear gain, which is equivalent to the y-intercept, and which we will call $g0$. Since Eg. (4.1) has three unknowns, we can construct a system of three equations using three data points, $g65$ and $g0$, as well as the slope $1/CR$. Substituting two data points directly into Eg. (4.1), and taking the derivative of Eg. (4.1), we can construct the following system of equations:

$$65 + g65 = C \times 65 \times e^{-65\alpha} + b \tag{4.2a}$$

$$g0 = b \tag{4.2b}$$

$$\frac{1}{CR} = Ce^{-65\alpha}(1 - 65\alpha) \tag{4.2c}$$

The solution to this system of equations results in a curve which closely matches the performance of conventional WDRC, as seen in Fig. 4.2 (left).

Although the proposed amplification curve can be substituted directly for a conventional curve, the purpose of this work is to rethink the way compression is addressed in hearing aids. As

shown in Fig. 4.1 (right), we propose a function which supplies constant linear gain to a signal, and then balloons outward to provide compression in the region where the signal of interest lies. Contrary to conventional WDRC, the gain does not remain constant for all low-level signals, and instead, falls back to the initial linear value. This reduction of gain for quiet signals is known as expansion.

To accomplish this goal, let us simplify the system of equations by substituting the expression $1/e^{-65\alpha}$ for parameter C in Eg. (4.2c), such that parameter C becomes a function of α:

$$\frac{1}{CR} = \frac{1}{e^{-65\alpha}} e^{-65\alpha}(1-65\alpha) = 1-65\alpha \tag{4.3}$$

By substituting C with a function of α, only one unknown variable remains in Eg. (4.3), yielding a closed form solution for α. Subsequently, all other unknown parameters can be solved for, yielding the following solutions for parameters α, b, and C:

$$\alpha = \frac{1}{65}(1-\frac{1}{CR}) \tag{4.4a}$$

$$b = g65 \tag{4.4b}$$

$$C = e^{65\alpha} \tag{4.4c}$$

Egs. (4.4a-c) define the reparameterized continuous compression curve, pictured in Fig. 4.1 (right). The curve is entirely defined by two audiological parameters – $g65$ and CR, where $g65$ sets the overall gain and CR sets the curvature of compression. The curve resembles a logarithm, and has a curvature dependent on CR. A comparison between the proposed amplification curve and a conventional WDRC curve without expansion is shown in Fig. 4.2 (right).

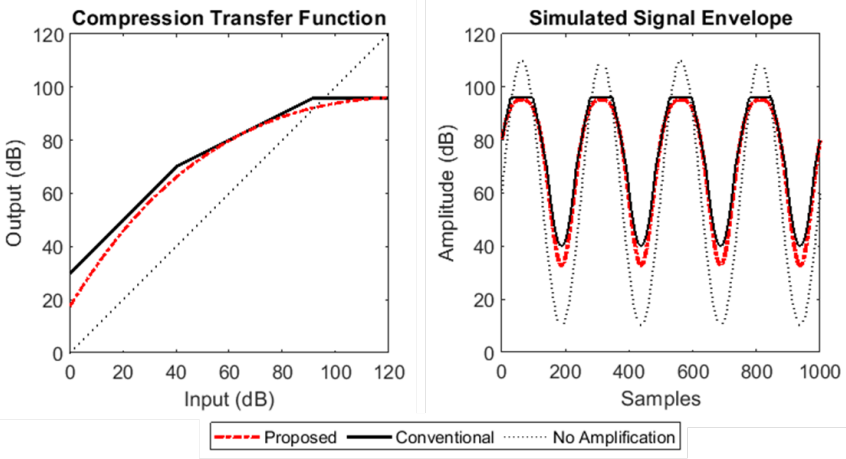

Figure 4.3. (Left) Conventional and curvilinear compression transfer functions with matched parameters; (Right) Effects of conventional and curvilinear compression on a sinusoidal signal envelope.

4.2.1 Signal Envelope Fidelity

Signal envelope is one of the most important perceptual cues for speech intelligibility. In WDRC, the signal envelope determines the instantaneous target gain of the signal. However, WDRC introduces distortions which may negatively affect perceived sound quality [7, 52].

One of the qualities of the curvilinear amplification rule is the removal of hard knee points, which reduces harmonic distortion in the envelope. To demonstrate this, let us simulate a modulated signal which has a sinusoidal envelope on the logarithmic scale, as shown in Figure 4.3 (right). Let us then apply the amplification transfer function of both conventional piece-wise compression and curvilinear compression to the simulated envelope. (For the purposes of this analysis, dynamic aspects of compression, such as attack and release time, are not included.) Both the conventional and curvilinear transfer functions are pictured in Figure 4.3 (left), and both curves are set to equal compression ratios (2:1), and equal maximum power output.

The output envelopes for both conventional and curvilinear compression are shown in Figure 4.3 (right). Both envelopes are compressed sinusoids, however, in the case of the conventionally amplified envelope, the peaks are clipped due to the hard-limiting MPO threshold, and the slopes of the sinusoid have a hard angle at the compression knee point. These hard

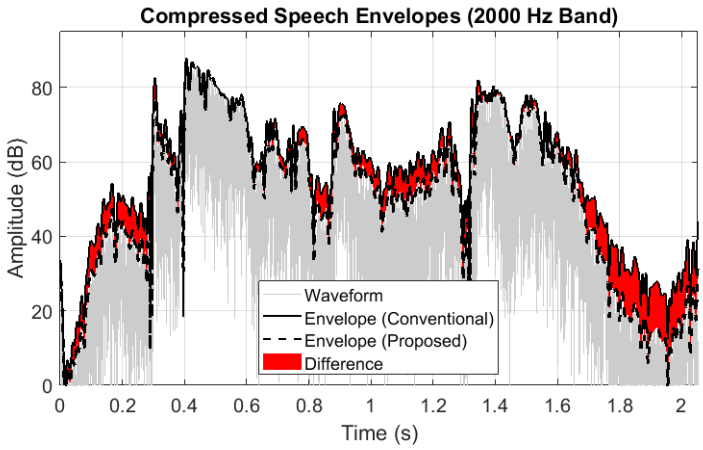

Figure 4.4. Speech envelopes amplified with conventional and curvilinear compression.

angles introduce harmonic distortion throughout the entire spectrum of the signal. In the case of the curvilinear amplification rule, compression is distributed across the entire dynamic range of the envelope, resulting in a smooth envelope and few harmonics. Because of this property, it is common to use compressors with soft knee points in music production.

Perhaps an even more important benefit of the proposed curve on the signal envelope is its expansion properties. It has been shown that compression distorts the envelope of a signal by equalizing amplitude variation across time, which distorts the temporal fine structure of the signal and disrupts important temporal cues, which may result in decreased speech intelligibility [8, 32, 53]. The envelope equalizing effect is more prominent for higher compression ratios, faster release times, and a greater number of frequency channels [8, 53].

The expansion properties of the curvilinear amplification function help preserve modulation depth of the signal envelope. Since the proposed curvilinear compression applies less gain to low-level signal components than conventional compression, the valleys of a waveform receive less amplification, as shown in Figure 4.3 (right), reducing the flattening of the envelope and preserving more temporal fine structure. Figure 4.4 shows the envelope of the sentence ”Please call Stella” spoken by a female voice amplified using both conventional and curvilinear compression with the same parameters as shown in Fig. 4.3. As seen in Fig. 4.4, higher-level portions of the signal receive equal amplification by both methods, but the proposed system

Table 4.1. Noise levels, speech levels, and signal-to-noise ratios of generalized listening environments for hearing aid users

Environment	Noise Level (dB)	Speech Level (dB)	SNR (dB)
Quiet	50	65	15
Moderate	60	68	8
Loud	70	70	0

yields deeper valleys, dependent on signal level. The impact of this change in temporal fine structure is evaluated using the HASQI index in Section 4.3.4.

4.3 Experimental Results

4.3.1 Experimental Setup

To validate the effectiveness of curvilinear compression, we measured the output SNR and HASQI [51] scores of recorded speech signals mixed with noise, processed using both curvilinear and conventional compression. The hearing aid software used for these experiments is the Open Speech Platform digital hearing aid [54]. Open Speech Platform is an open source suite of software and hardware tools for hearing aid research. The open source nature of the hearing aid software makes it possible to implement any custom amplification transfer function and access internal signals for direct SNR computations. The hearing aid performs multichannel compression, with subband decomposition performed in eleven frequency bands, corresponding to the standard audiometric frequencies [17].

A number of studies have been conducted to explore and quantify the real-world listening environments encountered by hearing aid users. One of the earliest studies by Pearsons et al. [55] collected speech, noise, and SNR data in various environments with normal hearing participants. Later, similar studies for hearing impaired users were conducted by Smeds et al. [56] and later by Wu et al. [57]. Summarizing the data found in [55, 56, 57], we determined a set of representative listening situations, summarized in Table 4.1, which we used in our experiments. (In some of the listed studies, sound levels are reported both in units of dB, and in weighted units of dBA. Since

the hearing aid operates on unweighted signals, we will be using dB units throughout this work)

Speech data was obtained from the Clarity Speech Corpus [58] – a public dataset created for the Clarity Project, which includes short high quality recordings of speech in the English language by 40 speakers of both genders. Noise recordings were obtained through the FSD50k Freesound Dataset Project [59] – a public dataset of crowd sourced human annotated audio recordings spanning a wide variety of sound categories.

4.3.2 Signal-to-Noise Ratio with Modulated Speech-Shaped Noise

Long-term signal-to-noise ratio (SNR) is a valuable indicator of perceptual performance, though it does not account for temporal dynamics of the signal [38, 39]. One well-documented effect of WDRC is a reduction of long-term output SNR [37, 38, 40, 41]. In normal listening situations, during segments of a signal dominated by speech, the compressor applies a lower gain determined mostly by the speech level. However, during gaps in speech dominated by noise, the compressor applies a higher gain determined mostly by the noise level. As such, noise receives greater amplification than the signal of interest, reducing the long-term SNR. This effect is more pronounced for higher input SNR values, as the separation between signal and noise increases, resulting in a SNR reduction which is roughly proportional to input SNR.

In the first experiment, we measured the effect of fast-acting compression on long-term output SNR as a function of input SNR, using speech signals mixed with modulated speech-shaped noise. The speech signal was generated by randomly selecting twenty speech passages from different speakers from the Clarity Speech Corpus. The speech passages were then concatenated to form a single running speech passage of 60s duration, with excessive silence removed. The noise signal is a speech-shaped modulated noise signal, derived using the methods specified by the International Collegium of Rehabilitative Audiology (ICRA) [60], and available courtesy of University of Oldenburg. ICRA noise is shaped to match the long-term spectrum of speech in order to isolate the effects of signal processing techniques, such as compression, from the effects of differing spectra. ICRA noise is also modulated to have a speech-like envelope for

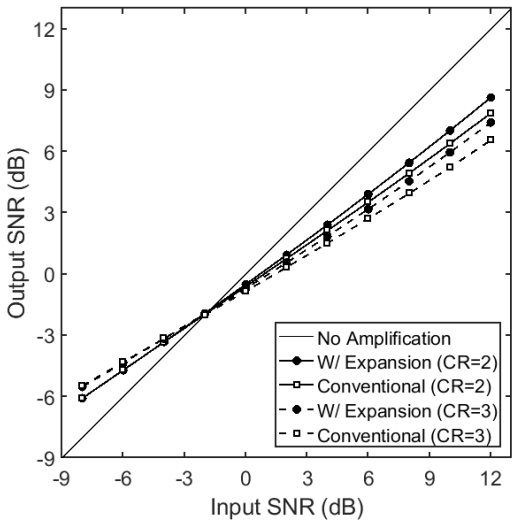

Figure 4.5. Output SNR as a function of input SNR and compression ratio for fast-acting compression using conventional and curvilinear WDRC.

a more realistic representation of the dynamic properties of real-world noise.

Similarly to work in [37, 38], we set the hearing aid to fast acting compression, with an attack time of 5 ms, and a release time of 50 ms. The compression threshold was set to 45 dB, and compression ratios of 2:1 and 3:1 were tested. The root mean square (RMS) level of speech was fixed at 60 dB, similarly to [38], and the noise level was varied based on desired input SNR.

Fig. 4.5 shows the long-term output SNR of curvilinear and conventional WDRC. In our experiments it was not necessary to use the phase-inversion technique of Hagerman and Olofsson [61], since the internal gain of the hearing aid is available for direct SNR computations. Consistent with the findings of similar experiments conducted in [37, 38, 40], long-term output SNR of fast-acting WDRC decreases roughly linearly with input SNR. Curvilinear compression follows a similar pattern, but output SNR levels are higher than with conventional compression due to the expansion properties of curvilinear transfer function, which reduce amplification of noise-dominated signal segments. Increased benefits of expansion are seen at higher input SNR values and higher compression ratios.

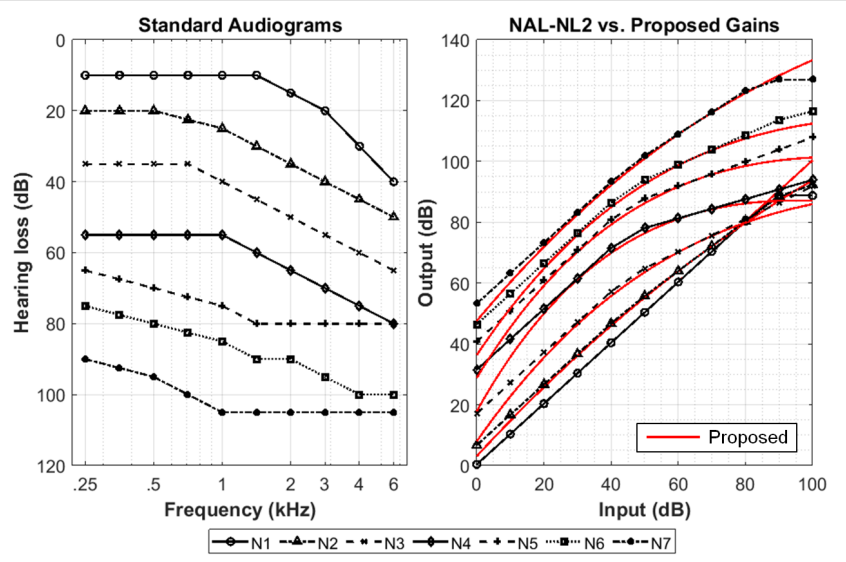

Figure 4.6. (Left) Standard hearing loss audiograms N1-N7; (Right) Respective NAL-NL2 prescription gains, with overlayed curvilinear counterparts.

4.3.3 Signal-to-Noise Ratio with NAL-NL2 Prescription Gains

In our second experiment, we compare long-term output SNR of curvilinear and conventional WDRC in more realistic hearing aid conditions. In this experiment, the hearing aid is programmed with NAL-NL2 prescription gains [30] for ten standard hearing loss audiograms specified by the International Standard for Measuring Advanced Digital Hearing Aids (ISMADHA) [29]. For the conventional WDRC prescriptions, the compression threshold is set to 45 dB. For the proposed compression rule, the curve is specified with the NAL-NL2 values for $g65$ and CR. The compression speed was set to moderately fast compression, with an attack time of 5 ms and a release time of 100 ms, in accordance with common practice [7, 8]. The standard audiograms, their corresponding NAL-NL2 prescription gains, and the respective curvilinear amplification curves with matching $g65$ and CR are pictured in Fig. 4.6.

Noisy speech signals were generated by combining speech passages with recordings of common noise sources. We selected three classes of sounds from the FSD50k dataset which represent three common types of listening situations: indoor noise, traffic noise, and babble noise [56, 57]. The selected sound classes are ”dishwasher”, ”car”, and ”chatter”, which we will

Table 4.2. Output SNR levels in dB of speech in noise compressed with conventional NAL-NL2 prescription gains and with the proposed curvilinear gains.

	Speech in Quiet (15 dB SNR)							**Speech in Moderate Noise (8 dB SNR)**							**Speech in Loud Noise (0 dB SNR)**						
	Dishwasher		**Car**		**Babble**			**Dishwasher**		**Car**		**Babble**			**Dishwasher**		**Car**		**Babble**		
	NAL.	Prop.	NAL.	Prop.	NAL	Prop.	**Avg Δ (%)**	NAL.	Prop.	NAL.	Prop.	NAL	Prop.	**Avg Δ (%)**	NAL.	Prop.	NAL.	Prop.	NAL	Prop.	**Avg Δ (%)**
N1	13.1	13.7	12.8	13.3	12	12.5	**0.5 (23%)**	6.3	6.7	6	6.3	5.2	5.5	**0.3 (15%)**	-0.8	-0.7	-1.2	-1.1	-1.8	-1.7	**0.1 (8%)**
N2	11.4	12.4	11.2	12	9.7	10.5	**0.9 (20%)**	4.6	5.2	4.7	5.2	3.3	3.8	**0.5 (14%)**	-1.9	-1.6	-1.7	-1.6	-2.8	-2.6	**0.2 (9%)**
N3	10.9	11.8	10.7	11.5	9	9.9	**0.9 (18%)**	4	4.7	4.2	4.8	2.8	3.3	**0.6 (14%)**	-2.3	-2	-2	-1.8	-3.1	-2.9	**0.2 (9%)**
N4	10.7	11.7	10.3	11.2	8.6	9.4	**0.9 (18%)**	3.9	4.6	3.9	4.6	2.5	3.1	**0.7 (15%)**	-2.3	-2	-2.1	-1.9	-3.2	-3	**0.2 (9%)**
N5	11.2	12.2	10.4	11.2	8.5	9.4	**0.9 (18%)**	4.4	5.1	3.9	4.6	2.4	3	**0.7 (15%)**	-2.1	-1.7	-2.4	-2	-3.6	-3.3	**0.4 (14%)**
N6	12	13.1	11.2	12.3	9.4	10.3	**1.0 (25%)**	5	6.1	4.4	5.5	2.8	3.7	**1.0 (26%)**	-2	-1.1	-2.2	-1.5	-3.5	-3	**0.7 (27%)**
N7	13.1	14	12.2	12.9	10.6	11.3	**0.8 (25%)**	6	7	5.2	6.1	3.6	4.4	**0.9 (29%)**	-1.3	-0.5	-2.1	-1.3	-3.4	-2.9	**0.7 (31%)**
S1	12.1	13	12.6	13.3	11.5	12.3	**0.8 (27%)**	5.2	5.8	5.8	6.2	4.8	5.3	**0.5 (18%)**	-1.4	-1.2	-1.1	-1	-1.8	-1.7	**0.1 (9%)**
S2	12.6	13.4	12.1	12.8	10.2	10.9	**0.7 (22%)**	5.8	6.3	5.5	6	4	4.4	**0.5 (16%)**	-1	-0.8	-1.1	-1	-2.4	-2.3	**0.1 (9%)**
S3	11.8	12.7	11.5	12.4	9.7	10.5	**0.9 (22%)**	4.9	5.6	4.9	5.5	3.4	3.9	**0.6 (17%)**	-1.7	-1.4	-1.7	-1.4	-2.9	-2.7	**0.3 (13%)**

call "babble" for consistency with audiology literature. Within each class we randomly selected 20 noise recordings. The 20 noise recordings from each class were mixed with 20 speech passages at three different SNR levels to represent three different listening environments, as designated in Table 4.1 – a quiet environment, a moderately noisy environment, and a loud environment. For each scenario, twenty speech passages were mixed with twenty noise recordings, for a total of 20 noisy speech passages for each listening environment. The mixed signals were processed by the hearing aid with both curvilinear and conventional compression.

The results of the experiment are summarized in Table 4.2. Output SNR levels were observed for each combination of listening scenario, hearing loss pattern, and noise type, and averaged over the twenty noisy speech passages. The average difference between NAL-NL2 output SNR and curvilinear output SNR (Δ) is obtained for each scenario and hearing loss pattern. The percentage difference $(\%)$ is calculated as the ratio of the difference (Δ) to the average SNR drop yielded by the NAL-NL2 prescription gains.

The reduction of output SNR seen in Table 4.2 for both NAL-NL2 and the proposed system is consistent with findings in [37, 38, 39]. Results show that curvilinear compression offers up to 1 dB SNR improvement for quiet and moderately noisy situations, and up to 0.7 dB in loud noise situations. For comparison, the SNR drop at moderate noise levels is in the range of 2-5 dB, and 1-4 dB for environments with loud noise. Thus, the improvement offered by the proposed method has an impact of up to 30%, as seen in Table 4.2. It is also noteworthy that the greatest improvements in SNR are seen at more severe hearing loss patterns where more intervention

Table 4.3. Output HASQI values of clean speech and speech in noise compressed with conventional NAL-NL2 prescription gains and with the proposed curvilinear compression gains.

	Clean			**SNR 15**			**SNR 8**			**SNR 0**		
	Conventional	Proposed	Difference	Conventional	Proposed	Difference	Conventional	Proposed	Difference	Conventional	Proposed	Difference
N1	0.95	0.97	**2.8%**	0.64	0.67	**2.7%**	0.49	0.51	**1.7%**	0.31	0.32	**0.8%**
N2	0.86	0.91	**5.0%**	0.67	0.71	**3.9%**	0.53	0.55	**2.5%**	0.34	0.35	**1.2%**
N3	0.79	0.84	**5.5%**	0.69	0.73	**4.6%**	0.56	0.59	**3.4%**	0.37	0.39	**1.4%**
N4	0.71	0.78	**6.7%**	0.66	0.72	**5.4%**	0.55	0.59	**4.1%**	0.38	0.40	**1.7%**
N5	0.76	0.82	**6.4%**	0.73	0.78	**5.3%**	0.65	0.69	**3.8%**	0.48	0.50	**1.7%**
N6	0.82	0.87	**4.9%**	0.80	0.84	**4.0%**	0.74	0.77	**3.0%**	0.56	0.58	**2.4%**
N7	0.92	0.94	**2.0%**	0.91	0.93	**1.6%**	0.88	0.89	**1.0%**	0.72	0.74	**1.8%**
S1	0.93	0.96	**2.8%**	0.66	0.68	**2.6%**	0.51	0.53	**1.7%**	0.32	0.33	**0.9%**
S2	0.89	0.93	**3.7%**	0.72	0.74	**2.8%**	0.57	0.59	**1.6%**	0.37	0.37	**0.6%**
S3	0.79	0.84	**5.3%**	0.69	0.73	**4.6%**	0.57	0.60	**3.0%**	0.39	0.39	**0.8%**

and higher signal quality are needed, likely due to the higher compression ratios prescribed at these hearing loss levels. Overall, this experiment shows that curvilinear compression provides equivalent or better output SNR than NAL-NL2 with simplified parameterization.

4.3.4 HASQI Score

To evaluate the perceptual effects of curvilinear compression on amplified speech, we used the Hearing Aid Speech Quality Index (HASQI) [51]. HASQI is an index designed to predict the perceived quality of speech for individuals listening through hearing aids, based on perceptual experiments involving normal-hearing and hearing-impaired listeners. It uses an auditory model of the impaired ear to measure the envelope fidelity and temporal fine structure fidelity of audio signals degraded by noise, linear, and nonlinear distortion. The HASQI index ranges from 0 to 1, where 1 represents no deterioration compared to the reference signal, and 0 represents fully degraded speech.

The HASQI index is computed by comparing a degraded audio signal to a clean reference signal. The reference signal is amplified with linear multi-channel amplification to compensate for hearing loss, while the test signal is processed with the algorithms being evaluated.

In this experiment, we evaluate the HASQI index of clean and noisy amplified speech at the output of the compressor. This test is performed for the ten standard ISMADHA hearing loss profiles with their corresponding NAL-NL2 prescription gains. The reference signal is the concatenated clean speech passage from Sec. 4.3.2, with linear $g65$ gain applied in eleven bands.

The test signal is the output of the compressor for four different listening conditions – clean speech, and the three noise conditions listed in Table 4.1.

As seen from Table 4.3, results show that our proposed curvilinear compression method shows similar, but improved performance compared to conventional WDRC, for all listening conditions. This is consistent with expectations, since the proposed compression rule aims to accomplish similar performance to conventional WDRC, with improvements. Clean speech shows the highest improvement in HASQI output, with a maximum improvement of 6.7% for profile N4, representing Moderate/Severe hearing loss. For all SNR conditions, the most severe hearing loss profiles, N4, N5, and N6, show the greatest improvement in HASQI score, since these hearing loss patterns are prescribed with higher compression ratios, which lead to greater signal distortion in conventional WDRC.

4.4 Applicability to Over-the-Counter Hearing Aids

The hearing loss research community has long recognized the need and concept for self-fitted hearing aids [13]. The population of individuals who could benefit from over-the-counter (OTC) self-fitted hearing aids includes individuals hindered by the high cost of traditional prescription hearing aids, individuals without access to reliable hearing health care, or individuals without access to professional hearing health care specialists. Commercial hearing enhancement devices, such as hearables and personal sound amplification products (PSAPs), have been influential in making hearing intervention more available to a wider population of hearing impaired individuals for over a decade, and on August 17, 2022, the FDA approved the commercial distribution of over-the-counter (OTC) hearing aids for individuals with perceived mild to moderate hearing loss.

In addition to affordability and accessibility, self-fitted OTC hearing aids also offer an avenue for greater hearing aid customization. Recent studies in self-fitted hearing aids reveal significant variability in user preferences for gain adjustments, often deviating significantly

from clinical prescription settings [49, 62]. Work in [49] observed gain changes between the NAL-NL2 prescribed gain and the self-adjusted gain as large as 24 dB in low frequencies and 37 dB in high frequencies, and describes the variability among participants as ”striking”. Despite the greater than expected between-subject variability, recent studies concur that self-adjusted amplification strategies are successful in selecting amplification gains which preserve speech intelligibility and meet hearing outcomes when compared to traditional prescription hearing aids [49, 50, 63].

In addition to tailoring hearing aids for speech amplification, there is a growing demand for higher fidelity of amplified music through a hearing aid [20, 64]. It is reported that hearing aids improve the enjoyment of listening to music for hearing impaired individuals, but there is significant dissatisfaction with sound quality and distortion. Moreover, the optimal gain settings of the hearing aid for music enjoyment may heavily depend on the musical genre, the musical instrument, and the personal preferences of the user. Thus, musicians and music enthusiasts may benefit from highly customizable self-adjustment of their hearing aids.

4.4.1 Musician-inspired hearing aid customization

While full customization of a user's hearing aid is an appealing concept, it is challenging to design a user interface which would be accessible to a user without professional training. In order to program a traditional WDRC hearing aid, a user would require individual control of $g65$, CR, $knee_{low}$, MPO, and possibly XR along with the expansion threshold, in each of the frequency channels available in the hearing aid. Moreover, the audiological terminology used in hearing aids may not be intuitive to a typical user.

We borrowed concepts from music production and sound systems to develop a graphic user interface for controlling the proposed curvilinear compression system. The interface, shown in Figure 4.7, is a set of two sliders per frequency channel – one slider for gain, and one for compression. The sliders are grouped in two rows, with one row containing the gain sliders, and the other containing compression sliders.

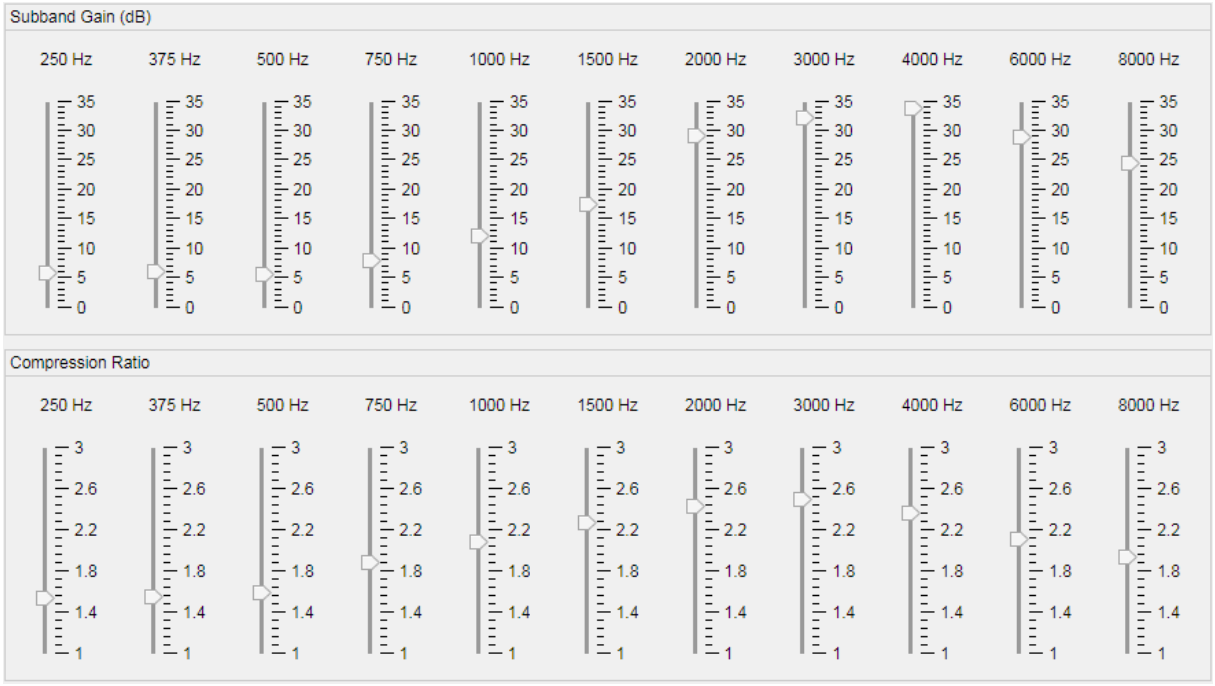

Figure 4.7. A prototype graphic user interface leveraging the proposed curvilinear amplification rule for highly customizable hearing aid user self-adjustment.

A row of gain sliders is immediately recognizable as a conventional Graphic Equalizer (EQ) system, seen in most sound systems in cars, computers, and smartphone apps. Audio equalization is one of the most indispensable tools of music production. Manipulating the EQ-balance of a signal dramatically affects sound timbre (ex. boosting bass to make audio sound ”fuller”, or boosting treble to make it sound ”brighter”), which can have a profound effect on the pleasantness and enjoyment of audio. An audio equalizer interface is likely to be familiar to many modern users.

We propose a similar interface for controlling the compression ratio of each frequency channel. Since the curvilinear transfer function consolidates $knee_{low}$, CR, and MPO, one slider controls all compressive aspects of each band. The gain and compression sliders can be adjusted in tandem to achieve a desired output sound. For example, increasing the compression ratio in all bands may cause the output to sound ”tingy”. The perceived ”tinginess” can be compensated by ”filling out” the sound by boosting low-frequency gains.

Table 4.4. Summary and characteristics of prevalent over-the-counter and self-adjusted hearing aids.

	Eargo	Lexie by Bose*	Ear Machine*	2D Touch Surface	Goldilocks	Curvilinear
Customization interface	Limited app interface	Volume slider and spectral tilt slider	Comprehensive adjustment with loudness wheel and spectral tilt wheel	2D plane interface with adjustable volume and spectral tilt	Volume, Crispness, Fullness sliders	Graphic equalizer and compression adjustment with slider interface
First fit	Proprietary	Proprietary Default	NAL-NL2	NAL-NL2	Mild-moderate generic	Linear
Gain adjustment	Linear	Profile	Profile	Linear	Linear	Any
Adjustable spectral tilt	Treble/Bass only	Yes	Yes	Yes	Yes	Yes
Individually adjustable subband gains	No	No	No	No	No	Yes
Adjustable compression	No	Profile	Profile	No	No	Yes
Individually adjustable subband compression	No	No	No	No	No	Yes

*Lexie Powered by Bose is a commercial sibling product of Ear Machine.

As a proof-of-concept, we designed a realtime WDRC signal processing app in Matlab for audio playback. The app reads an input frame from an audio file, processes the frame using the hearing aid software described in [54] and this work, and plays the continuous output stream through the device's audio output. Our Matlab app will be available for download soon at [65].

4.4.2 Comparison with self-fitting strategies

In this section, we survey a set of recent OTC and self-fitted hearing aids which have been prominent in recent audiological research. The approaches for hearing aid self-adjustment are greatly varied. Some approaches allow greater degrees of adjustment to allow for more customization, while other devices restrict the range of adjustment for a more user-friendly experience, and as a safeguard against improperly adjusted settings.

Table 4.4 shows a summary of prominent OTC and self-fitted hearing aids. Of the listed designs, Eargo offers the least customization. The user is directed to take a hearing screening test to generate an audiogram, and is automatically assigned proprietary prescription gains. Adjustment is limited to boosting treble or bass, changing volume, and selecting between preset configurations.

Ear Machine and Lexie Powered by Bose, two sibling products based on similar technology, are highly esteemed in the hearing aid community for closely replicating an audiologist's fitting. The self-fitting algorithm is governed by two controls – a coarse fit, which cycles between prescription gains derived from a large set of representative audiograms, and a fine fit, which

adjusts the spectral tilt. While the Bose self-fitting algorithm is highly successful in replicating an audiologist's fitting, it may not accommodate a user seeking non-standard fittings, such as gains tailored for music appreciation.

The 2D Touch Surface interface [66] is another strategy which seeks to consolidate a large number parameters into a smaller set of intuitive controls. The user adjusts the hearing aid parameters by moving a point on a two-dimensional surface, where movement along the x-axis changes spectral tilt, and movement along the y-axis changes overall gain. This approach functionally resembles Eargo and other published self-fitting strategies, but does not offer any customization of compression.

Goldilocks [62] differs from the other strategies listed by offering a generic starting condition and performing self-adjustment based on subjective descriptors. The user is prompted to adjust fullness, crispness, and loudness, which correspond to low-frequency cut, high-frequency boost, and overall gain respectively.

Our curvilinear graphic equalizer-inspired self-fitting strategy has a more involved user interface than other OTC devices, but offers features which are not available in other self-fitting strategies explored in this study. Per-channel gain adjustment allows for complex EQ shaping, such as a "vocal cut" (emphasize low-to-mid frequencies), a "powerful" cut (boost lower bass and upper treble), or targeting specific vocal regions such as nasal tones (around 2-3 kHz) or sibilance tones (around 5-8 kHz). The curvilinear graphic equalizer is also the only strategy which proposes individual adjustment of per-channel compression values. The authors propose an initial fitting with unity compression, and encourage an experimental approach to selecting per-channel compression ratios. Because unity compression offers the most signal fidelity, the authors argue in favor of starting low and increasing compression as necessary. However, any initial gain and compression values can be programmed into the equalizer.

The authors believe that the musician-inspired self-fitting strategy can benefit hearing impaired musicians, music enthusiasts, and users desiring high control and customization of their hearing aid.

4.5 Conclusion

In this chapter, we presented an alternative amplification transfer function for Wide Dynamic Range Compression (WDRC) in hearing aids. Our curvilinear transfer function is a continuous function with logarithm-like behavior, governed by two parameters – gain and compression, where gain applies constant amplification to a signal, while compression adjusts the curvature of the function. The proposed method reparameterizes a conventional compression curve from a piece-wise function with four or more parameters, to a smooth function with only two parameters. The proposed method also revisits dynamic range expansion, wherein gain decreases with signal level, to reduce the amplification of low-level ambient and circuit noise. The expansion properties of the curvilinear compression rule preserve greater modulation depth of the signal envelope and improve the temporal fine-structure compared to conventional WDRC. We evaluated our compression rule using speech signals mixed with modulated speech-shaped noise, as well as real-world noise recordings. Experimental results show that curvilinear compression reduces amplification of noise and improves long-term output SNR by up to 30%, and improves sound quality as measured by HASQI by up to 6.7%. We also discussed an application of curvilinear compression to over-the-counter (OTC) hearing aids, and compared the capabilities of prevalent OTC hearing aids in terms of customization and personalization.

4.6 Acknowledgements

Chapter 4, in part, is a reprint of the material as it appears in A. Sokolova, B. Aksanli, f. harris, and H. Garudadri, "A Curvilinear Transfer Function for Wide Dynamic Range Compression with Expansion," *Journal of the Audio Engineering Society*, 2023. (under review) The dissertation author was the primary investigator and author of this material.

Chapter 5

Freping (Frequency Warping) as a Tool for Feedback Control in Hearing Aids

5.1 Introduction

Acoustic feedback – an unpleasant "howling" or "squealing" caused by positive feedback in an audio path, is one of the leading causes of dissatisfaction with hearing aids. Feedback may prevent the hearing aid from supplying sufficient gain to a user with heavy hearing loss, and the growing use of open fittings in modern devices, as well as shorter distances between speaker and microphone, further complicate the problem of feedback control [19]. Many methods have been used to control feedback and increase the maximum usable gain of a hearing aid, such as reduction of gain in high frequencies, notch filtering, phase shifting, and frequency shifting, with varied success [67].

With the advent of digital hearing aids, the gold standard of feedback control is the adaptive cancellation filter, which estimates the feedback signal and subtracts the estimate from the input signal. However, because real-world speech and music are highly correlated in time, bias is introduced into the feedback path estimate, and the desired signal is partially canceled, negatively affecting the sound quality [19]. Many methods have been developed to minimize the error, such as decorrelation techniques and adaptation speed adjustments, but the problem persists.

In this work, we present experimental validation of a new approach for feedback cancel-

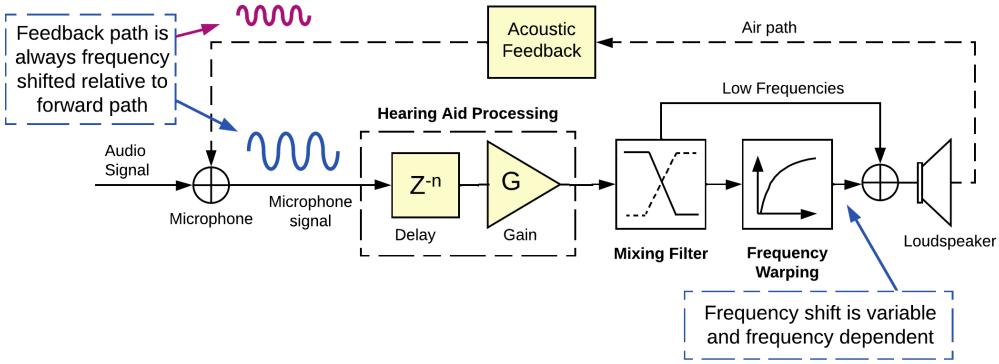

Figure 5.1. A block diagram of a hearing aid processing chain with frequency warping (Freping) for feedback mitigation.

lation called Frequency Warping (Freping). Freping is a nonlinear transformation which applies frequency-dependent spectral shift to an input signal. When applied to hearing aids, Freping inserts a variable frequency offset between the forward path and the feedback path, helping to prevent the onset of howling [68]. We tested the capabilities of Freping for feedback control in hearing aids by outfitting an acoustic manikin with a hearing aid running Freping and conducting experiments using Verifit, a standard verification tool for audiological research. Experimental results show that Freping helps to suppress feedback onset and offers up to 11 dB of Added Stable Gain (ASG) under the current test conditions.

5.2 Background and Related Work

Our feedback mitigation strategy using Freping is shown in Figure 5.1. The hearing aid signal processing chain can be simplified as a block diagram where an input signal is received by the microphone, amplified to compensate for a hearing impaired individual's hearing loss, and is outputted from the received (loudspeaker) some time later, represented by a latency block. The signal then travels through an unknown channel created by the acoustic air path from the loudspeaker back to the input microphone, resulting in a filtered version of hearing aid output being added to the received microphone signal. Thus, the signal recursively circulates through the system with diminishing amplitude. Positive feedback and squealing occurs when a copy of

the feedback signal is added to a phase-aligned version of itself in the feed forward path. The conditions of positive feedback have been defined by the Nyquist Stability Criterion [69], and requires two conditions to occur – the loop gain must exceed unity, and the phase offset between feedback and forward path must be a multiple of 2π.

In this work, we revisit frequency compression as a strategy for feedback control by leveraging the discrete transform described by Oppenheim in [70]. Frequency compression and frequency shifting are strategies which offset the frequency of the signal in the forward path such that the signal being fed back into the amplifier is different with every cycle, which prevents constructive summation. A frequency shifting circuit for feedback control was first described in [71], which inserted a 5 Hz frequency shift in the amplifier pathway of public address systems. The technique successfully stabilized the system, but introduced audible beating which limited the additional gain increase to 6 dB. Frequency shifting was first applied to hearing aids in [72], which showed that a 30 Hz shift eliminated feedback in the test conditions. However, the constant shift over all frequencies introduced audible distortions to the harmonic relationships of the original signal. A follow-up study introduced progressive frequency shifting, which increased the amount of shifting based on the input level, producing 6 to 15 dB improvement of usable gain [73]. However, this study first observed the warbling effect which occurs when low frequency signals are frequency shifted.

More recently, frequency shifting for feedback control in hearing aids has been simulated in [74]. Similarly to [72], this study found that a frequency shift of 30 Hz eliminated feedback and offered about 2 dB of additional usable gain. However, the same audible distortion was observed. The authors of [74] hypothesize that nonlinear frequency mapping can be used an alternative to constant frequency shifting, however, the theory was not explored in the study.

Some feedback cancellation strategies pair frequency shifting with adaptive filtering [75, 76]. In these approaches, the frequency shifter is used as a decorrelation method to decorrelate the hearing aid output signal from the microphone input, which would make it possible to use the hearing aid output signal as the reference signal for the adaptive filter. To avoid distortion

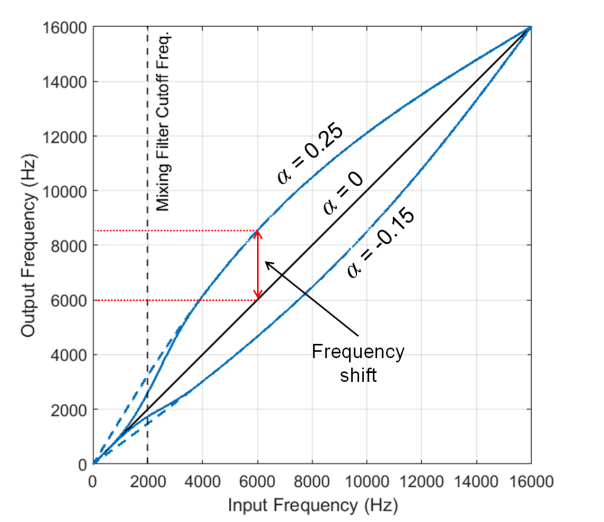

Figure 5.2. The input-output relationship of the frequency warping (Freping) transformation (large values of α pictured for visibility).

Table 5.1. Freping maximum frequency displacement

Alpha (alpha)		0.02	0.04	0.06	0.08	0.1
Max. frequency displacement	**(Hz)**	204	408	612	816	1020
	(%)	2.6%	5.2%	7.9%	10.7%	13.6%

of the signal's harmonic structure, authors in [75] used frequency compression – a frequency transformation where the frequency shift is proportional to the frequency component. The authors estimate that frequency compression of up to 6% is barely noticeable during informal listening tests.

It is also worth noting that frequency compression is also widely used to improve high frequency phonemic recognition for individuals with sloping and severe hearing losses [77, 78]. A device with nonlinear frequency compression proportionally lowers the frequencies of signals above a certain cutoff threshold in order to transpose signal components into a region where the individual has higher audibility. This technique has been shown to significantly improve recognition of */s/* and */z/* phonemes.

In this work, we validate the use frequency warping as a method of feedback management, first presented in [68]. The frequency warping method is based on Oppenheim's all-pass network

transformation which represents a digital signal by another digital signal with the same spectrum, but with a nonlinearly transformed frequency axis [70]. One way of viewing this transformation is as a frequency warping function, where the extent of warping is controlled by a parameter α. Figure 5.2 shows the mapping of input frequency to output frequency of the transformation as a function of α. Real-time operation of the Oppenheim transformation is implemented using frame-based processing with overlapping windowed frames.

Prior work on frequency shifting for acoustic feedback control identified two major drawbacks — the distortion of harmonic relationships in sound, and a warbling sound at lower frequencies. To address these drawbacks, we incorporated an adjustable mixing filter which separates lower pitch carrying frequencies from higher timbral frequencies. Similarly to frequency compressing hearing aids for sloping hearing loss, frequency warping is only applied to higher frequencies. By excluding low frequency components from the frequency transformation, critical pitch and harmonic information related to timbre and pitch of the voice is preserved and warbling is prevented.

5.3 Method

In this chapter, we explore how Freping and its associated parameters affect acoustic feedback. We used Added Stable Gain (ASG) as a metric for objective evaluation. ASG is defined as the additional gain margin available to the hearing aid when feedback suppression is enabled. It is commonly measured by first identifying the maximum stable gain (MSG) of the hearing aid without any feedback cancellation features enabled, and then comparing this value to the maximum stable gain obtained with feedback suppression enabled.

Freping is governed by two parameters – alpha (α) and the warping threshold.

Alpha

Parameter α determines the amount of frequency deviation yielded by Freping. The frequency input-output transfer function is illustrated in Figure 5.2 for various values of α.

Positive values of α shift the frequency upwards (higher frequencies), while negative values of α shift the frequency downward. (For the purposes of this work, only positive values of α are explored) As seen from Figure 5.2, the Oppenheim frequency transfer function produces smaller frequency shifts at lower and higher frequencies, and a larger shift for mid frequencies. The benefit of this property is that at a sampling rate of 32 kHz, the greatest shift is applied to frequencies in the 3-10 kHz region, where feedback is most likely to occur, and less shift is applied to frequencies outside that range to preserve more of the original signal. The maximum displacement occurs at approximately 8000 Hz. Maximum frequency displacement in Hz, as well as the percentage relative to the frequency of occurrence, are listed in Table 5.1. It is expected that the higher the value of α, the more frequency suppression, but with the trade-off of higher signal distortion.

Warping threshold

An important parameter in preserving the quality of sound is the warping threshold. The warping threshold divides the signal spectrum into low-frequency components and high-frequency components. Since feedback occurs in high frequencies in range of 3-10 kHz, only the high frequencies are warped with Freping, while low-frequencies are unaltered. The separation of high and low frequencies is implemented with a pair of complementary low-pass and high-pass filters, which intersect at the warping threshold. This provides a smooth transition between the warped and unwarped frequency regions. Figure 5.2 shows the warping threshold set at 2000 Hz, and the resulting input-output transfer function of Freping. The preservation of low-frequency components in extremely important for perceived audio quality because low frequencies contain essential pitch and harmonic information. For example, the vocal range of a soprano singer is approximately A3 to A5, which is 220 Hz to 880 Hz. A violin rarely plays above the sixth octave, which ends below 2000 Hz. The fundamental frequency of voiced speech lies approximately in the range of 80 to 450 Hz [79]. Thus, if the warping threshold is set below the region of feedback (about 3000 Hz), pitch information for conversational speech and even music is unaffected,

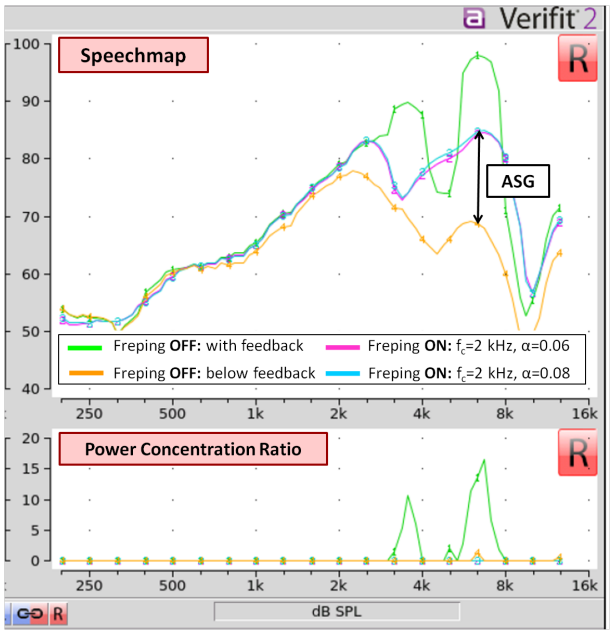

Figure 5.3. Speechmap output from Verifit, visualizing Added Stable Gain.

preserving natural voice pitch and harmonic relationships between pitches. There is an expected trade-off between frequency suppression and a higher warping threshold.

5.3.1 Experimental Setup

Experiments were conducted using the Open Speech Platform digital hearing aid [1, 54]. Open Speech Platform (OSP) is an open-source hearing aid for hearing loss research. The hardware consists of a small battery-powered wearable enclosure, containing a smart-phone chipset, which performs all signal processing and wireless communication tasks. The processing and communication device connects to custom hearing aid style audio transducers in behind-the-ear (BTE) receiver-in-canal (RIC) form factor. The audio transducers support four microphones and one receiver (loud speaker) per ear, as well as an inertial measurement unit (IMU), which can be used for a wider range of studies.

The OSP hearing aid was worn by KEMAR, an acoustic human head simulator manikin. To measure ASG, we performed on-ear instrument measurements using AudioScan Verifit 2, a standard audiology verification toolbox. The Verifit On-Ear feedback test presents a speech

signal in real-time to the hearing aid and measures the long-term average speech spectrum (LTASS) outputted by the hearing aid, and the power concentration ratio (PCR), which measures the degree to which a large amount of power is concentrated in a small frequency region.

ASG measurements were obtained using the procedure described in [80]. The hearing aid is programmed with constant gains across all frequencies. Then the gain is uniformly increased until howling begins, denoting the maximum stable gain without any feedback suppression techniques enabled. The orange curve in Figure 5.3 shows the maximum speech power attainable by the hearing aid without feedback suppression, while the green curve shows the hearing aid with stable feedback. The energy of the feedback oscillations can be seen in the Power Concentration Ratio plot, which measures the distribution of power as a function of frequency in the signal. After recording maximum output gain of the baseline hearing aid, Freping is enabled, and the hearing aid gain is once again uniformly increased, until the new howling threshold is reached with feedback suppression enabled. The new speech output power is recorded from the speechmap, which is exemplified by the blue and magenta curves in Figure 5.3. The average difference between the speech output map with Freping disabled and enabled is considered as the average ASG measurement.

5.4 Experimental Results

In prior work, Freping has been tested in conjunction with adaptive feedback cancellation [68]. Matlab simulations show that Freping with adaptive feedback cancellation eliminates howling artifacts which are otherwise not eliminated by adaptive feedback cancellation alone. The suppression of feedback artifacts has been shown to improve the Perceptual Evaluation of Speech Quality (PESQ) [81] score and Hearing Aid Speech Quality Index (HASQI) [51] score of sample speech recordings used in the experiment.

In this work, we preform On-Ear measurements to evaluate the capabilities of Freping as a standalone strategy for feedback mitigation in real-world conditions. We performed two

experiments: In the first experiment, we configure the hearing aid to have constant (flat) gain across all frequency channels. We then measure the ASG as a function of Freping α and mixing frequency by consecutively sweeping α with different mixing frequency values. In the second experiment, we program the hearing aid with real-world prescription gain for representative sample audiograms, and measure the value of α at which audible feedback is eliminated.

5.4.1 Added Stable Gain

In the first experiment, we measured the ASG of Freping as a function of warping parameter α and the warping threshold. We tested four warping threshold frequencies – 500 Hz, 1000 Hz, 1500 Hz, and 2000 Hz. For each warping threshold, we swept α from 0 to 0.10 in increments of 0.02. Figure 5.3 shows sample ASG measurements of Freping for a warping threshold of 2000 Hz and α values of 0.06, indicated by the magenta line, and 0.08, indicated by the blue line. The average measured ASG as a function of α and warping threshold frequency is shown in Figure 5.4.

As seen from the speechmap in Figure 5.3, high frequency regions have significantly higher output power when Freping is enabled, compared to the speech output levels with Freping disabled. The low-frequency regions below the warping threshold are unchanged, as explained in Section 5.3. The additional ASG in the high-frequency region is impactful because hearing loss generally affects higher frequencies greater than lower frequencies, and greater high frequency amplification is required for improved speech intelligibility. However, these frequency regions are also more susceptible to feedback, and as a result, some hearing aids cannot output sufficient high-frequency gain without incurring feedback [67].

Consistent with expectations and simulations, average ASG increases with increasing values of α. In all configurations, Freping offers at least 3 dB of average ASG at α equal to 0.06, and at least 9 dB at α equal to 0.08. Results in Figure 5.4 suggest that average ASG may not increase rapidly past α values of 0.1. Contrary to expectations, results in Figure 5.4 do not show increased average ASG as a function of a lower warping threshold. It was hypothesized that as

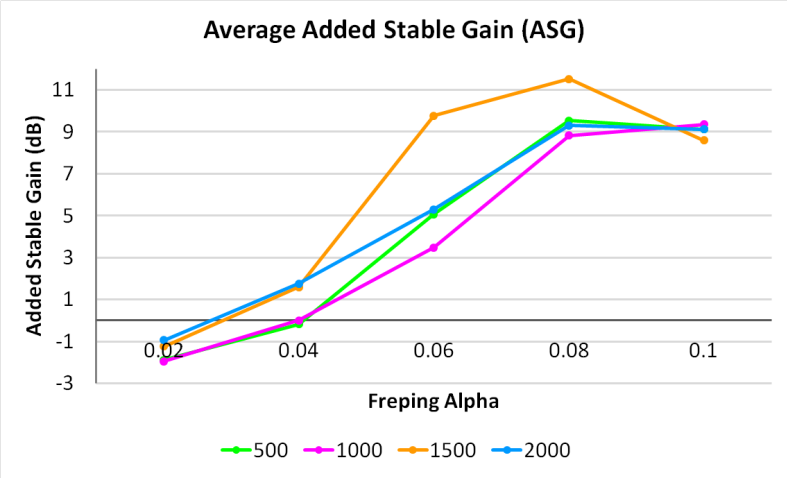

Figure 5.4. Added Stable Gain (dB) averaged over frequency as a function of α (x-axis) and warping threshold frequency in Hz (stacked curves).

the warping threshold is lowered, more of the signal undergoes frequency shifting, resulting in greater feedback suppression. However, warping thresholds of 500 Hz, 1000 Hz, and 2000 Hz show comparable average ASG values, with slightly better performance at 2000 Hz. However, Freping with a warping threshold of 1500 Hz showed significantly better average ASG than other configurations, yielding 10 dB of average ASG at α equal to 0.06, and 11 dB at α equal to 0.08. It can be concluded that suitable values for the warping threshold lie between 1500 and 2000 Hz.

5.4.2 Freping for Representative Audiograms

In the second experiment, we programmed the hearing aid with real world prescription gains for representative hearing loss profiles. The International Standard for Measuring Advanced Digital Hearing Aids (ISMADHA) group defines a set of sample audiograms which represent a broad class of hearing loss patterns, from mild to profound hearing loss [29]. For each hearing loss profile, we programmed the hearing aid with standard prescription gains generated from the ISMADHA audiograms using the commonly accepted NAL-NL2 prescription algorithm [30]. The hearing loss profiles are numbered in ascending order of hearing loss severity, where 'N' stands for normal hearing loss, and 'S' stands for sloping hearing loss. More severe hearing losses have higher prescription gains, and are more likely to cause feedback.

Table 5.2. Configurations of Freping (α) which suppress feedback for standard ISMADHA hearing loss profiles.

ISMADHA Profile	N1	N2	N3	N4	N5	S1	S2	S3
Alpha (α)	0	0	0.02	0.09	0.10	0	0	0.09

We tested eight hearing loss profiles, ranging from very mild to severe hearing loss. For a given profile, if the gain settings caused the hearing aid to enter feedback, we enabled Freping and incrementally increased α until audible feedback was eliminated. Freping successfully suppressed feedback for all profiles tested. Table 5.2 shows the value of α which suppressed feedback for each hearing loss profile. Consistent with expectations, more severe hearing loss profiles with higher prescription gains required higher values of α to eliminate feedback. A maximum value of α equal to 0.10 was required to suppressed feedback for severe hearing loss.

The perceptual user evaluation of Freping for hearing aids is a topic for future clinical research. However, informal listening experiments show that distortions introduced by Freping are fairly benign, consistently with other research in frequency shifting [75]. Because the warping threshold is set above the tonal frequencies of the voice, the pitch of the voice is unaffected. Rather, Freping affects the harmonics and overtones of the voice, which lie in higher frequencies, and convey information about voice timbre. Thus, Freping sounds like a minor change in voice timbre, where positive values of α make the voice sound ”brighter”, and negative values of α make the voice sound ”darker” (Not to be confused with the voice sounding higher or deeper). There is a trade-off between the degree of feedback suppression, determined by α and the warping threshold, and the deviation from the original signal.

This work explored the capabilities of Freping as a standalone algorithm for acoustic feedback management. Based on prior work [68], we believe optimal performance can be achieved by pairing Freping with adaptive filtering techniques.

5.5 Conclusion

Frequency Warping, which we term ”Freping”, is a discrete transformation which uses variable frequency shifting to suppress acoustic feedback in hearing aids. Freping is an adjustable algorithm which warps the frequency of a signal while retaining its magnitude, which hinders stable oscillations in the feedback loop. To preserve harmonic relationships within the signal and prevent warbling, pitch frequencies are separated from timbral frequencies, and warping is only applied to timbral frequencies. We evaluate the efficacy of Freping for feedback control in real world conditions by performing On-Ear Feedback tests on an acoustic KEMAR manikin. Experimental results show that Freping offers a trade-off between the extent of signal warping and feedback suppression, and offers up to 11 dB of Added Stable Gain (ASG) in current test conditions. Experimental results also show that Freping successfully suppresses feedback for standard hearing loss profiles ranging from mild to severe hearing loss. Results thus far show that Freping is an effective method for feedback management in hearing aids which can be used independently, or in conjunction with adaptive filtering methods.

5.6 Acknowledgements

This works is partially sponsored by: NIH R33-DC015046: ”Self-fitting of Amplification: Methodology and Candidacy”; NIH R01-DC015436: ”A Real-time, Open, Portable, Extensible Speech Lab”; NSF IIS-1838830: Division of Information & Intelligent Systems, “A Framework for Optimizing Hearing Aids In Situ Based on Patient Feedback, Auditory Context, and Audiologist Input”; NIH R44 DC020406: ”An Open-source Speech Processing Platform (OSP) for Research on Hearing Loss and Related Disorders”; Wrethinking Foundation; The Qualcomm Institute, UC San Diego.

Chapter 5, in part, is a reprint of the material as it appears in A. Sokolova, V. Rallapalli, A. Yellamsetty, M. Hunt, B. Aksanli, f. harris, and H. Garudadri, “Validation of Frequency Warping (Freping) as a new tool for feedback control in hearing aids”. *Asilomar Conference on*

Systems, Signals, and Computers, 2023. The dissertation author was the primary investigator and author of this material.

Chapter 6

Summary and Future Work

Wide Dynamic Range Compression (WDRC) is a staple of modern hearing aids. WDRC is used to compress the dynamic range of audio signals to accommodate the reduced dynamic range of hearing impaired persons, and has been shown to improve audibility, loudness comfort, and intelligibility. However, it is challenging to adjust WDRC parameters for an individual's unique hearing loss pattern, and there exists significant dissatisfaction with the performance and sound quality of WDRC among hearing aid users. Some prevalent issues in WDRC design include: (i) the accuracy of fitting hearing aid prescription gains to an individual's unique hearing loss; (ii) the improvement of hearing aid sound quality; (iii) the management of signal interference, such as noise and acoustic feedback.

6.1 Thesis Summary

This dissertation presents novel solutions for ongoing design challenges in digital hearing aids. We address issues of accuracy, processing power, and latency in WDRC. We also challenge the conventional method of performing compression in hearing aids and propose a novel approach for the compressive amplification of signals. Our approach aims at improving sound quality and pleasantness of listening through hearing aids, and offers greater possibilities for personalization of the hearing aid. We also address the issue of acoustic feedback in hearing aids, which is one of the most prominent sources of dissatisfaction for hearing impaired users.

6.1.1 Multirate Audiometric Filter Bank for Frequency Decomposition in Hearing Aids

We designed a multirate filter bank for audiometric frequency decomposition in hearing aids. The multirate processing structure of our frequency bank enables us to provide very steep narrowband filters with very low complexity. We use polyphase filtering to further reduce the computational complexity of our channelizer. We also used minimum phase filtering to reduce the latency of our channelizer for real-time operation. Our multirate filter bank offers a $14\times$ improvement in complexity compared to a singlerate implementation, with 5.4 ms of latency. Our filter bank offers more channels with higher frequency resolution than prior work, improving the accuracy of a hearing aid prescription.

6.1.2 Automatic Gain Control for Wide Dynamic Range Compression

We present a novel Automatic Gain Control (AGC) system for WDRC. Our approach offers very accurate signal envelope estimation and highly precise AGC. We explored the relationship between the dynamic and static parameters of WDRC, and derived a closed-form solution for the attack and release times of AGC as a function of WDRC parameters. Based on this derivation we designed an AGC algorithm which accurately fulfills user-defined response time specifications within 0.5 ms. Our AGC system utilizes multirate processing to reduce the computational complexity of compressive amplification.

6.1.3 A Curvilinear Transfer Function for Wide Dynamic Range Compression with Expansion

We present a novel approach for performing compressive amplification in hearing aids. We propose to reparameterize the underlying input-output magnitude transfer function of WDRC. Our proposed transfer function is a curvilinear continuous function, offering compression in the regions of speech, and expansion for low-level signals. The expansion properties of our transfer function help to preserve the modulation depth of speech and improve longterm output

SNR. Our experiments show that the curvilinear transfer function improves output SNR by up to 30%, and sound quality by up to 6.7% as measured by the HASQI index. Our alternative reparameterization of WDRC is highly applicable to over-the-counter hearing aids, and offers greater opportunities for user customization than conventional WDRC.

6.1.4 Freping (Frequency Warping) as a Tool for Feedback Control in Hearing Aids

We designed an algorithm called Freping for acoustic feedback control as an alternative for adaptive filtering. Freping utilizes variable frequency shifting to break the Nyquist Stability Criterion and prevent stable feedback oscillations. We preserve sound quality and harmonic relationships within the signal by separating pitch frequencies from timbral frequencies, and only applying frequency warping to the latter. We validated the effectiveness Freping in real-world conditions using an acoustically realistic manikin and representative hearing loss profiles. Experimental results show that Freping offers up to 11 dB of added stable gain and prevents feedback for all conditions tested.

6.2 Future Work

6.2.1 Temporal Aspects of Compression

Various fitting algorithms, such as NAL-NL2 [30], are widely used in the audiology community to generate prescriptive gains from an audiogram. Although the algorithms may vary, there is a general consensus on the different ways hearing aid gains are determined. However, there is much less certainty in setting the response speed of compression, which is determined by attack and release time. Compression speeds are generally qualitatively described as ”slow”, ”fast”, or in-between, and are adjusted mostly experimentally. There is a need for better understanding of the effect of compression response times on speech intelligibility. Hearing aid design may benefit from signal dependent variable response times, determined by signal properties such as crest factor.

6.2.2 Perceptual Validation of Alternative Compression Methods

We validated a novel curvilinear transfer function for WDRC using objective metrics, such as SNR and HASQI [51]. However, formal perceptual experiments are necessary to determine the true validity of our design. To this purpose, we are planning perceptual studies to compare the sound quality of conventional compression and curvilinear compression. For this study, we will recruit normal-hearing participants. Experiments will be conducted under the guidance of our audiology collaborators at Northwestern University.

6.2.3 Customization of Hearing Aids

The curvilinear compression rule described in this dissertation holds compelling potential for self-adjusted hearing aids. We introduce a novel idea of offering a user control over the per-band compression ratios of their hearing aids. However, this is a demanding task for a user without training in audiology. We hypothesize that offering a user some guidance in choosing suitable parameters would improve a user's experience with self-adjustment. We propose labeling different channelizer bands with subjective descriptors, such as ”warm”, ”tinny”, or ”hissy”, to aid the user in targeting different signal properties. Existing research shows that subjective descriptors can be used to describe the coloration of sound with a fair degree of consensus [82]. We expect that such descriptors can be combined with our self-adjustment user interface for a much improved user experience.

6.2.4 Perceptual Validation of Alternative Feedback Management

In this dissertation, we described a trade-off between frequency warping and feedback cancellation, where larger amounts of frequency warping reduce feedback, but distort the signal more strongly. Although our algorithm, Freping, preserves low frequency content, it remains to be determined how the distortions of the timbral frequencies affect user perception and satisfaction with the hearing aid output with frequency warping.

Bibliography

[1] L. Pisha, J. Warchall, T. Zubatiy, S. Hamilton, C.-H. Lee, G. Chockalingam, P. P. Mercier, R. Gupta, B. D. Rao, and H. Garudadri, "A wearable, extensible, open-source platform for hearing healthcare research," *IEEE Access*, vol. 7, pp. 162083–162101, 2019.

[2] D. L. Blackwell, J. W. Lucas, and T. C. Clarke, "Summary health statistics for us adults: national health interview survey, 2012.," *Vital and health statistics. Series 10, Data from the National Health Survey*, no. 260, pp. 1–161, 2014.

[3] National Institute on Deafness and Other Communication Disorders, "Quick statistics about hearing," *Statistics and Epidemiology*, 2016.

[4] S. Kochkin, "Marketrak VIII: Consumer satisfaction with hearing aids is slowly increasing," *The Hearing Journal*, vol. 63, no. 1, pp. 19–20, 2010.

[5] A. McCormack and H. Fortnum, "Why do people fitted with hearing aids not wear them?," *International journal of audiology*, vol. 52, no. 5, pp. 360–368, 2013.

[6] R. J. Bennett, A. Laplante-Lévesque, C. J. Meyer, and R. H. Eikelboom, "Exploring hearing aid problems: Perspectives of hearing aid owners and clinicians," *Ear and hearing*, vol. 39, no. 1, pp. 172–187, 2018.

[7] P. E. Souza, "Effects of compression on speech acoustics, intelligibility, and sound quality," *Trends in amplification*, vol. 6, no. 4, pp. 131–165, 2002.

[8] J. M. Kates, "Understanding compression: Modeling the effects of dynamic-range compression in hearing aids," *International journal of audiology*, vol. 49, no. 6, pp. 395–409, 2010.

[9] R. L. Miller and A. Donahue, "Open speech signal processing platform workshop," *Nat. Inst. Health, Bethesda, MD, USA, Tech. Rep*, 2014.

[10] "THE lab at UC San Diego," *Open Speech Platform*, 2019. Available: http://openspeechplatform.ucsd.edu/.

[11] D. Sengupta, T. Zubatiy, S. K. Hamilton, A. Boothroyd, C. Yalcin, D. Hong, R. Gupta, and H. Garudadri, "Open speech platform: Democratizing hearing aid research," in *Proceedings of the 14th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2020.

[12] H. Garudadri, A. Boothroyd, C.-H. Lee, S. Gadiyaram, J. Bell, D. Sengupta, S. Hamilton, K. C. Vastare, R. Gupta, and B. D. Rao, "A realtime, open-source speech-processing platform for research in hearing loss compensation," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pp. 1900–1904, 2017.

[13] E. Convery, G. Keidser, H. Dillon, and L. Hartley, "A self-fitting hearing aid: Need and concept," *Trends in amplification*, vol. 15, no. 4, pp. 157–166, 2011.

[14] "Medical devices; ear, nose, and throat devices; establishing over-the-counter hearing aids." Food and Drug Administration Rule 87 FR 50698, 2022.

[15] J. M. Kates, "Principles of digital dynamic-range compression," *Trends in amplification*, vol. 9, no. 2, pp. 45–76, 2005.

[16] M. A. Stone and B. C. Moore, "Tolerable hearing aid delays. I. estimation of limits imposed by the auditory path alone using simulated hearing losses," *Ear and Hearing*, vol. 20, no. 3, pp. 182–192, 1999.

[17] American Speech-Language-Hearing Association and others, "Guidelines for manual pure-tone threshold audiometry [guidelines]," 2005.

[18] *Specification of Hearing Aid Characteristics*. ANSI Standard S3.22-2009.

[19] A. Spriet, S. Doclo, M. Moonen, and J. Wouters, "Feedback control in hearing aids," *Springer Handbook of Speech Processing*, pp. 979–1000, 2008.

[20] M. Chasin and F. A. Russo, "Hearing aids and music," *Trends in Amplification*, vol. 8, no. 2, pp. 35–47, 2004.

[21] Y. Hu and P. C. Loizou, "Subjective comparison and evaluation of speech enhancement algorithms," *Speech communication*, vol. 49, no. 7-8, pp. 588–601, 2007.

[22] Y. Ghanbari, M. Karami, and B. Amelifard, "Improved multi-band spectral subtraction method for speech enhancement," in *Proc. 6th IASTED internat. conf. on signal image process*, pp. 225–230, 2004.

[23] J. M. Kates, "Dynamic range compression using digital frequency warping," Sept. 6 2011. US Patent 8,014,549.

[24] L. Lee and R. Rose, "A frequency warping approach to speaker normalization," *IEEE Transactions on speech and audio processing*, vol. 6, no. 1, pp. 49–60, 1998.

[25] J. M. Kates, *Digital hearing aids*. Plural publishing, 2008.

[26] *Specification for Octave-Band and Fractional-Octave-Band Analog and Digital Filters*. ANSI Standard S1.11-2004.

[27] R. E. Crochiere and L. R. Rabiner, *Multirate digital signal processing*. Prentice-Hall, 1983.

[28] Audioscan, "Audioscan verifit user's guide 4.28," 2022.

[29] N. Bisgaard, M. S. Vlaming, and M. Dahlquist, "Standard audiograms for the iec 60118-15 measurement procedure," *Trends in amplification*, vol. 14, no. 2, pp. 113–120, 2010.

[30] G. Keidser, H. Dillon, M. Flax, T. Ching, and S. Brewer, "The NAL-NL2 prescription procedure," *Audiology research*, vol. 1, no. 1, pp. 88–90, 2011.

[31] D. McShefferty, W. M. Whitmer, and M. A. Akeroyd, "The just-noticeable difference in speech-to-noise ratio," *Trends in hearing*, vol. 19.

[32] L. M. Jenstad and P. E. Souza, "Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility," 2005.

[33] A. C. Neuman, M. H. Bakke, C. Mackersie, S. Hellman, and H. Levitt, "The effect of compression ratio and release time on the categorical rating of sound quality," *The Journal of the Acoustical Society of America*, vol. 103, no. 5, pp. 2273–2281, 1998.

[34] G. Grimm, T. Herzke, S. Ewert, and V. Hohmann, "Implementation and evaluation of an experimental hearing aid dynamic range compressor," *Threshold*, vol. 80, no. 90, p. 100, 2015.

[35] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, N. L. Dahlgren, and V. Zue, "Timit acoustic phonetic continuous speech corpus," *Linguistic Data Consortium, 1993*, 1993.

[36] F. Harris and G. Smith, "On the design, implementation, and perfor-mance of a microprocessor controlled agc system for a digital receiver," 1988.

[37] P. E. Souza, L. M. Jenstad, and K. T. Boike, "Measuring the acoustic effects of compression amplification on speech in noise," *The Journal of the Acoustical Society of America*, vol. 119, no. 1, pp. 41–44, 2006.

[38] G. Naylor and R. B. Johannesson, "Long-term signal-to-noise ratio at the input and output of amplitude-compression systems," *Journal of the American Academy of Audiology*, vol. 20, no. 03, pp. 161–171, 2009.

[39] K. S. Rhebergen, T. H. Maalderink, and W. A. Dreschler, "Characterizing speech intelligibility in noise after wide dynamic range compression," *Ear and Hearing*, vol. 38, no. 2, pp. 194–204, 2017.

[40] J. M. Alexander and K. Masterson, "Effects of wdrc release time and number of channels on output snr and speech recognition," *Ear and Hearing*, vol. 36, no. 2, p. e35, 2015.

[41] R. M. Corey and A. C. Singer, "Modeling the effects of dynamic range compression on signals in noise," *The Journal of the Acoustical Society of America*, vol. 150, no. 1, pp. 159–170, 2021.

[42] G. Walker, D. Byrne, and H. Dillon, "The effects of multichannel compression/expansion amplification on the intelligibility of nonsense syllables in noise," *The Journal of the Acoustical Society of America*, vol. 76, no. 3, pp. 746–757, 1984.

[43] R. A. van Buuren, J. M. Festen, and T. Houtgast, "Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality," *The Journal of the Acoustical Society of America*, vol. 105, no. 5, pp. 2903–2913, 1999.

[44] P. N. Plyler, A. B. Hill, and T. D. Trine, "The effects of expansion on the objective and subjective performance of hearing instrument users," *Journal of the American Academy of Audiology*, vol. 16, no. 02, pp. 101–113, 2005.

[45] G. H. Mueller, "Experts debate key fitting issues: Ldls, kneepoints, and high frequencies," *The Hearing Journal*, vol. 52, no. 10, pp. 21–26, 1999.

[46] A. Sokolova, B. Aksanli, F. Harris, and H. Garudadri, "Consolidating compression and revisiting expansion: an alternative amplification rule for wide dynamic range compression," in *2023 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 1–5, IEEE, 2023.

[47] S. Gatehouse, G. Naylor, and C. Elberling, "Linear and nonlinear hearing aid fittings–1. patterns of benefit: adaptación de auxiliares auditivos lineales y no lineales–1. patrones de beneficio," *International Journal of Audiology*, vol. 45, no. 3, pp. 130–152, 2006.

[48] H. Dillon, J. A. Zakis, H. McDermott, G. Keidser, W. Dreschler, and E. Convery, "The trainable hearing aid: What will it do for clients and clinicians?," *The Hearing Journal*, vol. 59, no. 4, p. 30, 2006.

[49] P. B. Nelson, T. T. Perry, M. Gregan, and D. VanTasell, "Self-adjusted amplification parameters produce large between-subject variability and preserve speech intelligibility," *Trends in Hearing*, vol. 22, p. 2331216518798264, 2018.

[50] A. T. Sabin, D. J. Van Tasell, B. Rabinowitz, and S. Dhar, "Validation of a self-fitting method for over-the-counter hearing aids," *Trends in Hearing*, vol. 24, p. 2331216519900589, 2020.

[51] J. M. Kates and K. H. Arehart, "The hearing-aid speech quality index (hasqi) version 2," *Journal of the Audio Engineering Society*, vol. 62, no. 3, pp. 99–117, 2014.

[52] K. H. Arehart, J. M. Kates, and M. C. Anderson, "Effects of noise, nonlinear processing, and linear filtering on perceived speech quality," *Ear and hearing*, vol. 31, no. 3, pp. 420–436, 2010.

[53] M. A. Stone and B. C. Moore, "Quantifying the effects of fast-acting compression on the envelope of speech," *The Journal of the Acoustical Society of America*, vol. 121, no. 3, pp. 1654–1664, 2007.

[54] A. Sokolova, D. Sengupta, M. Hunt, R. Gupta, B. Aksanli, F. Harris, and H. Garudadri, "Real-time multirate multiband amplification for hearing aids," *IEEE Access*, vol. 10, pp. 54301–54312, 2022.

[55] K. S. Pearsons, R. L. Bennett, and S. A. Fidell, *Speech levels in various noise environments*. Office of Health and Ecological Effects, Office of Research and Development, 1977.

[56] K. Smeds, F. Wolters, and M. Rung, "Estimation of signal-to-noise ratios in realistic sound scenarios," *Journal of the American Academy of Audiology*, vol. 26, no. 02, pp. 183–196, 2015.

[57] Y.-H. Wu, E. Stangl, O. Chipara, S. S. Hasan, A. Welhaven, and J. Oleson, "Characteristics of real-world signal-to-noise ratios and speech listening situations of older adults with mild-to-moderate hearing loss," *Ear and hearing*, vol. 39, no. 2, p. 293, 2018.

[58] S. Graetzer, M. A. Akeroyd, J. Barker, T. J. Cox, J. F. Culling, G. Naylor, E. Porter, and R. Viveros-Muñoz, "Dataset of british english speech recordings for psychoacoustics and speech processing research: The clarity speech corpus," *Data in Brief*, vol. 41, p. 107951, 2022.

[59] E. Fonseca, X. Favory, J. Pons, F. Font, and X. Serra, "FSD50K: an open dataset of human-labeled sound events," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 829–852, 2022.

[60] W. A. Dreschler, H. Verschuure, C. Ludvigsen, and S. Westermann, "ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment: Ruidos ICRA: Señates de ruido artificial con espectro similar al habla y propiedades temporales para pruebas de instrumentos auditivos," *Audiology*, vol. 40, no. 3, pp. 148–157, 2001.

[61] B. Hagerman and Å. Olofsson, "A method to measure the effect of noise reduction algorithms using simultaneous speech and noise," *Acta Acustica United with Acustica*, vol. 90, no. 2, pp. 356–361, 2004.

[62] A. Boothroyd and C. Mackersie, "A "Goldilocks" approach to hearing-aid self-fitting: User interactions," *American journal of audiology*, vol. 26, no. 3S, pp. 430–435, 2017.

[63] D. W. Swanepoel, I. Oosthuizen, M. A. Graham, and V. Manchaiah, "Comparing hearing aid outcomes in adults using over-the-counter and hearing care professional service delivery models," *American Journal of Audiology*, vol. 32, no. 2, pp. 314–322, 2023.

[64] S. M. Madsen and B. C. Moore, "Music and hearing aids," *Trends in Hearing*, vol. 18, p. 2331216514558271, 2014.

[65] "Open speech platform (GitHub repository)." https://github.com/nihospr01/OpenSpeechPlatform.

[66] J. A. Gößwein, J. Rennies, R. Huber, T. Bruns, A. Hildebrandt, and B. Kollmeier, "Evaluation of a semi-supervised self-adjustment fine-tuning procedure for hearing aids," *International journal of audiology*, vol. 62, no. 2, pp. 159–171, 2023.

[67] F. Kuk, C. Ludvigsen, and T. Kaulberg, "Understanding feedback and digital feedback cancellation strategies," *Hearing Review*, vol. 9, no. 2, pp. 36–41, 2002.

[68] C.-H. Lee, K.-L. Chen, B. D. Rao, H. Garudadri, *et al.*, "On mitigating acoustic feedback in hearing aids with frequency warping by all-pass networks," in *Interspeech*, vol. 2019, p. 4245, NIH Public Access, 2019.

[69] H. Nyquist, "Regeneration theory," *The Bell System Technical Journal*, vol. 11, no. 1, pp. 126–147, 1932.

[70] A. V. Oppenheim and D. H. Johnson, "Discrete representation of signals," *Proceedings of the IEEE*, vol. 60, no. 6, pp. 681–691, 1972.

[71] M. R. Schroeder, "Improvement of acoustic-feedback stability by frequency shifting," *The Journal of the Acoustical Society of America*, vol. 36, no. 9, pp. 1718–1724, 1964.

[72] M. Bennett, S. Srikandan, and L. Browne, "A controlled feedback hearing aid," *Hearing Aid Journal*, vol. 33, no. 5, pp. 42–43, 1980.

[73] D. P. Egolf, "Review of the acoustic feedback literature for a control systems point of view," *The Vanderbilt Hearing-Aid Report*, pp. 94–103, 1982.

[74] E. Berdahl and D. Harris, "Frequency shifting for acoustic howling suppression," in *Proceedings of the 13th International Conference on Digital Audio Effects, Graz, Austria*, vol. 610, 2010.

[75] H. A. L. Joson, F. Asano, Y. Suzuki, and T. Sone, "Adaptive feedback cancellation with frequency compression for hearing aids," *The Journal of the Acoustical Society of America*, vol. 94, no. 6, pp. 3248–3254, 1993.

[76] M. Guo and B. Kuenzle, "On the periodically time-varying bias in adaptive feedback cancellation systems with frequency shifting," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 539–543, IEEE, 2016.

[77] A. Simpson, A. A. Hersbach, and H. J. McDermott, "Improvements in speech perception with an experimental nonlinear frequency compression hearing device," *International journal of audiology*, vol. 44, no. 5, pp. 281–292, 2005.

[78] Y. Mao, J. Yang, E. Hahn, and L. Xu, "Auditory perceptual efficacy of nonlinear frequency compression used in hearing aids: A review," *Journal of Otology*, vol. 12, no. 3, pp. 97–111, 2017.

[79] T. Bäckström, O. Räsänen, A. Zewoudie, P. P. Zarazaga, L. Koivusalo, S. Das, E. G. Mellado, M. B. Mansali, D. Ramos, S. Kadiri, and P. Alku, *Introduction to Speech Processing*. 2 ed., 2022.

[80] S. C. Marcrum, E. M. Picou, C. Bohr, and T. Steffens, "Feedback reduction system influence on additional gain before feedback and maximum stable gain in open-fitted hearing aids," *International Journal of Audiology*, vol. 57, no. 10, pp. 737–745, 2018.

[81] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, "Perceptual evaluation of speech quality (pesq)-a new method for speech quality assessment of telephone networks and codecs," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221)*, vol. 2, pp. 749–752, IEEE, 2001.

[82] A. T. Sabin, Z. Rafii, and B. Pardo, "Weighted-function-based rapid mapping of descriptors to audio processing parameters," *Journal of the Audio Engineering Society*, vol. 59, no. 6, pp. 419–430, 2011.