Tampere University

Phan Vu Thien Quang

LITERATURE REVIEW ON STATE-OF-THE-ART ALGORITHMS ON LOW POWER HEARING-AID RELATED AUDIO PROCESSING

Bachelor of Science Thesis
Faculty of Information Technology and Communications Sciences
Supervisor: Prof. Pekka Jääskeläinen, Matti Haavisto
April 2021

ABSTRACT

Phan Vu Thien Quang: Literature Review on State-of-the-Art Algorithms on Low Power Hearing-Aid Related Audio Processing
Bachelor of Science Thesis
Tampere University
Bachelor of Science and Engineering
April 2021

In today world, hearing-aid devices are crucial for people with defective hearing ability. In the heart of such device lies the digital signal processor (DSP), which takes responsibility for processing the incoming speech and producing the undistorted audio signal to the user. The hearing-aid devices must be small enough to be worn in human ears, yet powerful to compute a large amount of data with low latency and high accuracy. Last but not least, hearing-aid batteries should last for several days to meet the customer satisfaction. Therefore, they should be optimized in terms of performance and energy consumption.

In this thesis, the state-of-the-art audio processing algorithms that are widely utilized in digital hearing-aid devices are identified. The potential algorithms are divided into different parts according to the DSP functionality. For each part, the fundamentals of the most applicable methods and some of their revised versions are explained in detail.

Based on the selected domains and algorithms, this thesis work explores some efficient low power implementations in real digital hearing aids. The collection is again separated into subsections corresponding to the DSP processing segments. Subsequently, the cutting-edge low-power implementations of the whole DSP for hearing-aid purposes are classified based on the hardware architecture.

In summary, most of the findings belong to feedback cancellation and noise reduction categories because those sections determine the quality of the audio output. In addition to the signal characteristics, a high output signal amplitude is desired, which is why beamforming algorithms play an important role in hearing-aid processors. As regards the technology, Application-Specific Instruction-set Processor (ASIP) with numerous hardware accelerators dominates the recent low-power hearing-aid implementations because of its performance and power efficiency.

Keywords: digital hearing-aid, low power, feedback cancellation, noise reduction, beamforming, DSP, ASIP, hardware accelerators

PREFACE

This work was carried out at Tampere University Finland from January 2021 to April 2021 as the Bachelor's thesis of Science and Engineering.

First and foremost, I would like to send my deepest gratitude to my supervisor Professor Pekka Jääskeläinen for his valuable guidance and comments from the beginning. I would also like to thank my supervisor Matti Haavisto for his continuous feedback and advice for the writing process of my thesis work.

Finally, I really appreciate the support and encouragement from my beloved family and friends which motivates me to finish this Bachelor's degree.

Tampere, 26th April 2021

Phan Vu Thien Quang

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

ADC	Analog-to-Digital Converter
APA	Affine Projection Algorithms
ASIC	Application-Specific Integrated Circuit
ASIP	Application-Specific Instruction-set Processor
DAC	Digital-to-Analog Converter
DRC	Dynamic Range Compression
DSP	Digital Signal Processor
FB	Filter Bank
FBC	Feedback Cancellation
FIR	Finite Impulse Response
FPGA	Field-Programmable Gate Array
GPP	General-Purpose Processor
GSC	Generalized Side-lobe Canceller
HA	Hearing Aid
LCMV	Linearly Constrained Minimum Variance
LMS	Least Mean Square
MOS	Metal-Oxide Semiconductor
MSE	Mean-Square Error
NLMS	Normalized Least Mean Square
NR	Noise Reduction
SNR	Signal-to-Noise Ratio
SOTA	State-Of-The-Art
WF	Wiener Filter

1. INTRODUCTION

Hearing impairment is one of the most universal chronic diseases for American residents aged 65 years or older [1]. For people with hearing loss, traditional communication is almost impossible, especially with significant background noise, because some or all essential parts of the speech are inaudible. Hearing loss is also proved to be involved in incident all-cause dementia [2]. Hearing-aid devices are invented to help people with hearing loss overcome these obstacles.

Hearing-aid systems were first popularized as analog, which introduced the elementary approach of hearing-loss indemnity [3]. The improvement from analog to digital hearing aids made them more indispensable for people with hearing disability. Indeed, when performing operations with similar complexity, digital hearing aids reduces the required energy and volume compared to analog devices [4]. Consequently, digital hearing aids are able to adapt digital signal processing techniques, which further improves its computing performance and output audio quality. However, all systems must preserve a low power consumption to prolong the limited battery life of the devices [5]. Therefore, the outstanding power-optimized algorithms and implementations in digital hearing-aid applications are most interested.

In this thesis, state-of-the-art (SOTA) audio processing methods in digital hearing-aid systems are classified based on their application domains: beamforming, noise reduction and acoustic feedback cancellation. All aforementioned processing domains help to improve the audio quality. The algorithms are collected based on popularity, performance and applicability. In the scope of this thesis, only major algorithms for each domain and their upgraded versions, if exist, are introduced. A few outstanding low-power platforms in terms of area and power consumption of the analyzed algorithms are presented next. Finally, some SOTA hardware implementations of the digital signal processors (DSP) are illustrated and compared with others that are not covered in this work.

The remainder is organized as follows. Chapter 2 provides a brief picture of modern digital hearing-aid systems. Chapter 3 presents the observed audio processing algorithms, whilst chapter 4 depicts the procured implementations of these methods. SOTA implementations of the DSP based on the hardware architecture are discussed next in chapter 5. Chapter 6 concludes this thesis work and proposes future improvements.

2. DIGITAL HEARING-AID SYSTEMS

The simplified structure of a digital hearing-aid system and its key electronic components are introduced in this chapter. Figure 2.1 shows an elementary block diagram of a typical modern hearing aid (HA).

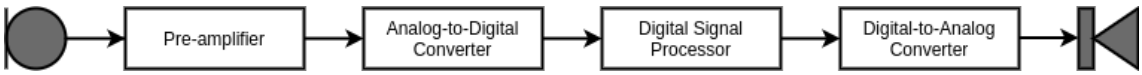

Figure 2.1. *Block diagram of a digital hearing aid*

The incoming speech arrives at the hearing aids through microphones, where it is first amplified before processing. In the real world, sound signals are analog, whereas advanced audio processing techniques require digital format (bits and symbols) to maximize their performance and reduce the computing resources. Therefore, Analog-to-Digital Converters (ADCs) are used to convert analog waves to digital form. Taking these digitized signals as input, Digital Signal Processor (DSP) is the core of HAs, where the signal processing algorithms are executed to create the best audio output without noise or echo. Digital-to-Analog Converters (DACs) are placed after the processing phase to convert the signals back to analog and deliver to users through receivers.

2.1 Microphone

Microphones are must-have items of any hearing-aid systems, where speech sound waves are converted into analog electric currents with similar wave properties. Invented by E. Berliner in 1877 [6], microphones have evolved drastically with various types such as pressure, condenser, laser and the latest MEMS (MicroElectrical-Mechanical System).

Directionality shows the orientations to which the microphone can distinguish the desired sound from background noise. Traditionally, microphones are divided into two categories: omni-directional or directional. The omni-directional microphones collect sound waves from any direction. This type of microphone is well suited for quiet environments, where the noise amplitude is insignificant compared to the speech's strength. On the other hand, directional microphones are designed to preserve good receptiveness to sound from one particular direction, typically the front of the wearers in terms of hearing aids, while suppressing the noise from some other sources [4]. This feature allows the users

to concentrate on desired sound locations without the interference of background noise, thus is most effective in noisy environments.

Based on the conventional directional microphones, adaptive directional microphones enable the wearers to continuously change the focus area according to speech and sound signals by adjusting the filter parameters. Ultimately, a set of microphones, named *microphone array*, is used to leverage the advantages of both omni-directional and directional microphones. The targeted directional response is produced by accommodating the results of all microphones in the array [6]. The microphone array has unlocked a new era of portable and accurate HAs by their applications in digital audio processing techniques.

2.2 Converter

In the outside world, speech signals are analog, which is continuous in time domain and amplitude. However, complex audio processing techniques in HAs require digital representations of them for better performance and lower energy consumption. Therefore, incoming analog waves are converted into digital forms before infiltrating the digital circuits, and then changed back to analog for hearing purposes [4]. This is done by analog-to-digital and digital-to-analog converters. The basic operation principle of ADCs is illustrated in figure 2.2 and explained in the following, while DACs can be obtained by overturning the procedure.

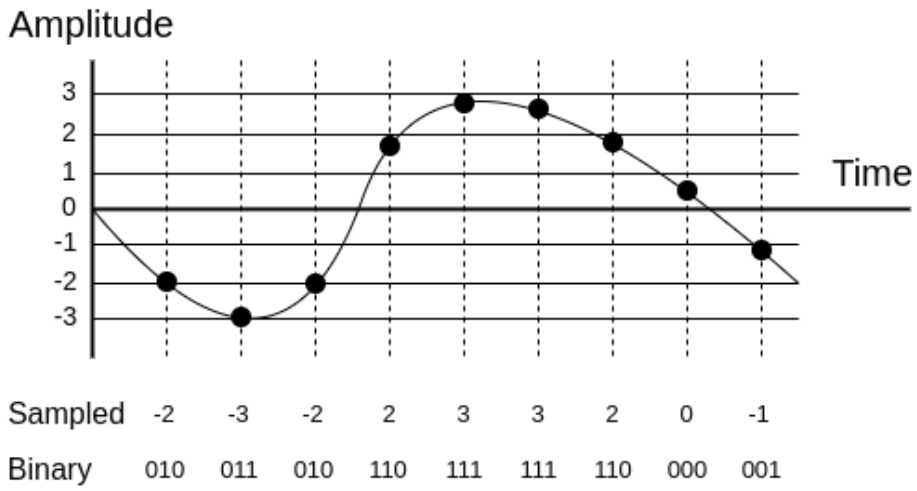

Figure 2.2. *Analog-to-digital conversion*

In general, the digital format of the signals is actually a sequence of numbers that are taken from the analog waves once for a particular time interval. This process is referred to as *sampling*. In order to preserve all essential information, the duration of this interval must be small enough. Following the *Nyquist* theorem, its inverse, called *sampling frequency*, needs to be larger than twice of the highest frequency in the speech combination [7]. The converter employs some low-pass filters to exclude all components that have a

higher frequency than half of the sampling rate to prevent *aliasing* [4]. The next step is to map those sampled values into a specific range of numbers, which is referred to as *quantization*. The algorithm matches each sample with the closest number in the alphabet. The waveform is finally digitized after transforming the mapped result into binary format with user-defined coding rules [4].

2.3 Digital Signal Processor

Lying at the center of HAs, Digital Signal Processor (DSP) is responsible for executing several complex signal processing techniques to enhance the speech quality as well as suppress the unwanted noise. The DSP structure varies in terms of implementation platforms and dedicated algorithms. Nowadays, there are two main approaches to design a hearing aid DSP [8]:

- **General-purpose processors** (GPPs) offer great flexibility to adopt sophisticated processing methods, but consume more power and silicon area. These are capable of serving all objectives, but not optimized for any specific function, which might be extravagant in terms of computation power and energy consumption for DSPs.
- **Fixed-function processors** (FFPs) are specially designed for some particular purposes, which bring a better performance with a lower power expense. However, one main drawback of this architecture is the fixed functional behavior which demands a re-implementation for new algorithms [8].

A rising trend for processor designs is to combine the design flexibility of GPPs and computing power of FFPs. This is where Application-Specific Instruction-set Processor (ASIP) comes into play. They originate from GPPs, but the instruction-set architecture has been specialized for a target application [9]. This customization allows the processors to operate at a lower clock frequency, which reduces the power consumption at the same processing rate [8]. ASIPs, which still retain poorer performance characteristics of the GPPs compared to FFPs, can be further boosted by dividing computational intensive segments into different dedicated FFP modules, also known as *hardware accelerators* [10]. There are different ASIP designing tools such as the commercial ASIP Designer from Synopsys, or the open-source TCE tool developed at Tampere University [11].

Application-Specific Integrated Circuit (ASIC) is the actual hardware implementation platform for ASIPs and FFPs. It consists of three main categories, specifically full-custom, semi-custom and programmable [12]. Full-custom ASICs refer to those integrated circuits where all components such as resistors, transistors, capacitors, etc are hardwired on the printed circuit layout and cost million of dollars for designing and fabrication [13]. Semi-custom ASICs are somewhat similar, but some blocks are pre-designed in ASIC standard libraries for reusing. Such components are actually designed as full-custom

ASICs and then integrated into a larger ASIC implementation. Field-Programmable Gate Array (FPGA) is a programmable ASIC which allows users to re-configure the functionality by changing the interconnects between logic elements [14]. FPGAs are now commonly used for small/medium prototyping, while semi-custom ASICs are the most cost-effective choice for large-scale projects.

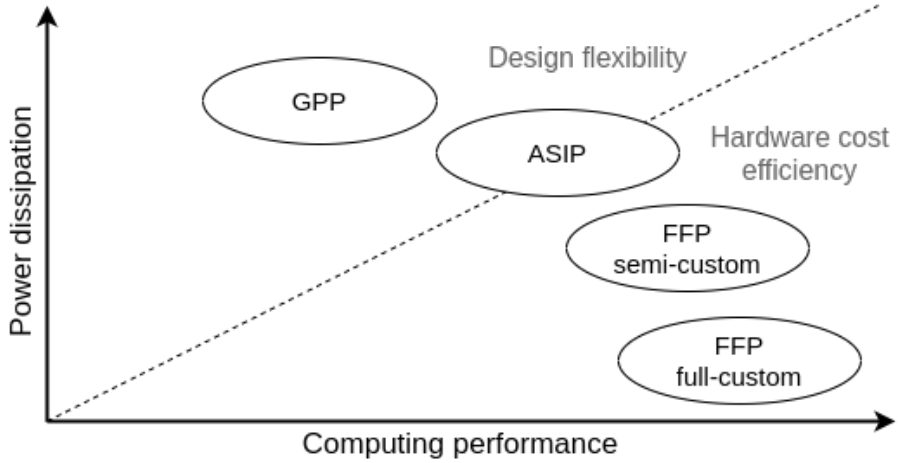

Figure 2.3. *Comparison of different hardware architectures*

Figure 2.3 displays the comparison between different DSP layouts. In particular, they are GPP, ASIP, FFP with semi- and full-custom ASIC architecture. It is shown that GPP remains the most flexible design at the expense of low computational capability and high power consumption. FFPs with ASIC implementation have an extreme processing speed with low power dissipation. The cost for this feature is the non-tunable nature which prevents FFPs from adapting complex algorithms. ASIP is a great trade-off between programmability and computing performance [8], which is well-suited for low-power hearing-aid applications.

3. STATE-OF-THE-ART ALGORITHMS IN DIGITAL HEARING-AID SYSTEMS

This chapter describes some of the most prevalent and effective signal processing algorithms that are widely used in digital hearing-aid systems. They are categorized based on their application parts in typical hearing aids, specifically Noise Reduction and Acoustic Feedback Cancellation.

3.1 Signal Models

In this thesis work, the input signal, also known as the *regressor*, $y[k]$ at time *k* of n-th microphone consists of a distorted version of the clean speech $s[k]$ and undesired noise [15]

$$y_n[k] = h_n[k] \circledast s_n[k] + v_n[k] = x_n[k] + v_n[k], \tag{3.1}$$

in which

$h_n[k]$	the impulse response between the sound source and microphone;
$x_n[k]$ and $v_n[k]$	the speech and noise component at the microphone, respectively;
$\circledast$	convolution operation.

Using the analysis and synthesis window technique that bases on overlapping frames, we can derive the frequency response of the input signal with Short-time Fourier Transform (STFT) [16] as

$$Y_n = H_n S_n + V_n. \tag{3.2}$$

Thus, in general, the lower case x denotes the signal in the time-domain, whereas the upper case X is used for frequency-domain. Vector $\hat{u}$ represents the single response

$$\hat{u} = \begin{bmatrix} 1 & 0 & ... & 0 \end{bmatrix}^T. \tag{3.3}$$

The superscript T implies the matrix transposition while the superscript $*$ refers to the complex conjugate operator. Their combination, the complex conjugate transposition, is indicated as the superscript H. In addition, $\mathbb{E}[.]$ intimates the expected or mean value of a property and can also be expressed using correlation.

3.2 Beamforming

In realistic situations, the desired speech signal is heavily mixed with background noises which can cause speech quality degeneration and break the acoustic scenario. Beamforming (BMF) methods enhance the desired signal by continuously steering the microphone array towards the direction of arrival and suppresses the unwanted noise or competing sounds from other directions [17]. In this section, we discuss a widely-adapted algorithm named Linearly Constrained Minimum Variance (LCMV) and its realization called the Generalized Side-lobe Canceller (GSC).

3.2.1 Linearly Constrained Minimum Variance (LCMV)

The Linearly Constrained Minimum Variance (LCMV) beamformer, proposed by Frost [18], laid the foundation of modern beamforming techniques. The algorithm forces a linear equality constraint on the array of N microphones to reproduce the non-distorted speech signal while reducing the unwanted noise and interference [19]. The architecture of the LCMV beamformer is based on the *Filter-and-sum* concept, which is introduced in figure 3.1.

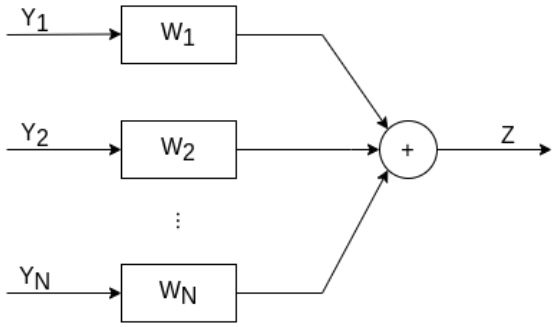

Figure 3.1. *Filter-and-sum principle*

Beamforming methods are usually associated with the directional microphone array introduced in section 2.1. The filter output in frequency domain is derived as

$$Z = \mathbf{W}^H\mathbf{Y} = \mathbf{W}^H\mathbf{H}S + \mathbf{W}^H\mathbf{V} \triangleq Z_s + Z_n, \tag{3.4}$$

in which Z_s and Z_n denote the speech and undesired components, respectively [20]. The LCMV beamformer produces an output power (PSD) of

$$\mathbb{E}\{\|Z\|^2\} = \mathbf{W}^H\mathbf{R}_{YY}\mathbf{W}, \tag{3.5}$$

where $\mathbf{R}_{YY}$ implies the correlation matrix of the collected input signals [20]. The PSD is

minimized with respect to the linear constraint matrix C

$$\mathbf{C}^H\mathbf{W} = \hat{u}, \tag{3.6}$$

where $\hat{u}$ is the response vector defined in section 3.1 [21]. Utilizing the complex Lagrangian function [22], the LCMV optimal solution is obtained as

$$\mathbf{W}_{LCMV} = \frac{\mathbf{R}_{YY}^{-1}\mathbf{C}}{\mathbf{C}^H\mathbf{R}_{YY}^{-1}\mathbf{C}}\hat{u}. \tag{3.7}$$

The Minimum Variance Distortion Response (MVDR) algorithm, first proposed by Capon [23] and upgraded to the adaptive framework by Frost [18], is a particular case of the LCMV, where the constraint matrix is reduced to just a scalar operator on the desired speech signal

$$\mathbf{W}_{MVDR} = \frac{\mathbf{R}_{YY}^{-1}\mathbf{H}}{\mathbf{H}^H\mathbf{R}_{YY}^{-1}\mathbf{H}}. \tag{3.8}$$

Both LCMV and MVDR beamformers are well-recognized in the field of noise reduction and speech enhancement. When assistive hearing is needed for both ears, LCMV is usually extended to the Binaural version (BLCMV) [19], which is proved to conserve the *binaural cues* and hence protect the auditory scenes for the wearers. Moreover, in the context of low-power, the BLCMV seems to be a strong candidate, since it is designed to optimize the output power while producing undistorted speech [24] and can be further simplified by relaxing the constraint matrix with the cost of higher inaccuracies. However, although it shows a great performance in noise reduction and interference cancellation while preserving the speech quality, the BLCMV requires prior knowledge about the desired direction of arrival and the absence of reverberation effect , which might not be achievable in real life. Furthermore, with a restricted amount of microphones in HAs, the noise elimination is less effective due to the lack of applicable constraints [25]. To sum up, the LCMV is a promising technique for the desired low-power hearing-aid applications.

3.2.2 Generalized Side-lobe Canceller (GSC)

The Generalized Side-lobe Canceller (GSC) is a widely-used NR technique for hearing aid systems, proposed by Griffiths and Jim in 1982 [26]. Perceiving the impulse response of the acoustic channel, the GSC divides the minimization problem in the LCMV into two orthogonal phases [27]. The first phase, named *spatially pre-processing*, contains a fixed beamformer $\mathbf{A}(z)$ and a blocking matrix $\mathbf{B}(z)$ [28] and aims to solve the distortionless speech constraint. The output noise power is minimized in the later phase by a multi-channel adaptive noise canceller (ANC) [29]. The GSC structure is shown in figure 3.2.

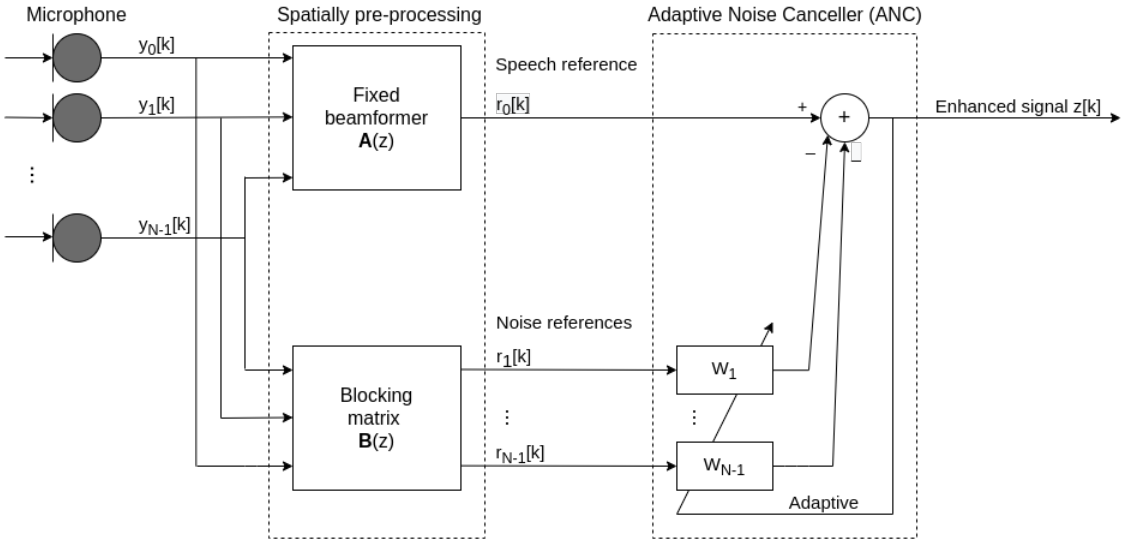

Figure 3.2. *Generalized Side-lobe Canceller structure*

The so-called *speech reference* $r_0[k]$ is collected in *speech-periods* when both speech and noise are presented by steering the beamformer $\mathbf{A}(z)$ towards the source direction and contains both clean speech and noise components [28]. Meanwhile, the blocking matrix $\mathbf{B}(\mathbf{z})$ produces the *noise references* $r_{1:N-1}[k]$ for each secondary channel in *noise-only-periods* where only the noise components are received [30]. The noise is assumed to be short-term stationary in speech pauses and uncorrelated with the speech component so that they can be estimated and used in *speech-periods* [15].

As proved by Breed and Strauss [31], the GSC and LCMV in section 3.2.1 are equivalent. However, since the speech distortion and noise reduction are split into two separate components, the filtering part (ANC) is unconstrained, which allows designers to employ rudimentary noise cancellation techniques to reduce the computational cost [32]. On the other hand, because of external effects such as speech reverberation, microphone mismatch or source localization error, unwanted speech parts might leak into the *noise references*, also known as *speech leakage* [33], resulting in the speech distortion and even signal cancellation [30]. This make the GSC vulnerable to noisy environments and hence not preferable for the low-power hearing-aid domain. Several methods have been developed to preserve the speech integrity and enhance the system's robustness against signal errors, including the *quadratic inequality constraint* (QIC), hence the name QIC-GSC (for example in [34]). Another important and effective technique based on the GSC that takes the speech distortion into consideration will be introduced in the upcoming section. Additionally, the computation complexity of GSC can be saved by estimating the filter coefficients with adaptive method such as LMS [35], which will be described in the Feedback Cancellation section.

3.3 Noise Reduction

Before processing, low-power incoming speech signals with noise added from the surroundings or from the hearing devices themselves are amplified. Such unwanted noise may deteriorate the quality of the speech and even break its intelligibility. Therefore, hearing-aid systems require some noise reduction (NR) procedures to minimize the interference of noise to the desired speech signal. In an ideal case, only the clean speech signal is preserved, whereas additive noise are completely annihilated [4]. This is not achievable in real life; however, the main goal of NR algorithms is to amplify the speech as much as possible while maintaining low noise level, in other words to maximize the signal-to-noise ratio (SNR). Among all, the most popular methods for noise reduction in digital HAs nowadays are Wiener Filter & its derivations.

3.3.1 Wiener Filter

For audio processing applications, Wiener Filters, named after Norbert Wiener, are used to diminish the error between input and output signals, or reduce the effect of unpleasant noise on the coveted speech in detailed. Small devices such as hearing aids need a compact and simple NR technique, thus finite impulse response (FIR) Wiener Filter is preferred. Compared to infinite impulse response (IIR) implementations, FIR filters are more stable and practical, and do not require heavy computation resources [36]. Figure 3.3 illustrates the mechanism of a FIR Wiener filter of length L.

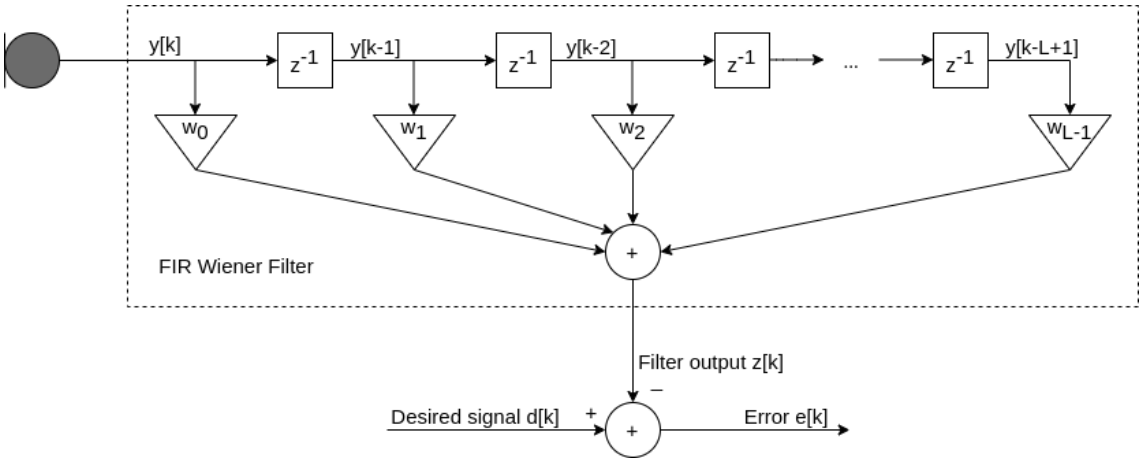

Figure 3.3. *FIR Wiener Filter principle*

The Wiener filter stores the current input as well as its $L-1$ delay versions, producing an output signal of

$$z[k] = \sum_{l=0}^{L-1} w_l^*[k]y[k-l] = \mathbf{w}^H[k]\mathbf{y}[k], \tag{3.9}$$

where $\mathbf{w}[k]$ is the filter coefficient vector, also known as weight [29]. The difference

between the desired signal $d[k]$ and the filter output $z[k]$ defines the error signal $e[k]$ of FIR Wiener filter [36]

$$e[k] = d[k] - z[k] = d[k] - \mathbf{w}^H[k]\mathbf{y}[k]. \tag{3.10}$$

The Wiener filter emphasizes to obtain the minimum of quadratic cost function called the minimum mean-square error (MMSE) criterion [37]

$$J_{MSE}(\mathbf{w}) = \mathbb{E}\{e^2[k]\} = \mathbb{E}\{d^2[k]\} - 2\mathbf{w}^H\mathbb{E}\{d^*[k]\mathbf{y}[k]\} + \mathbf{w}^H[k]\mathbb{E}\{\mathbf{y}[k]\mathbf{y}^H[k]\}\mathbf{w}. \tag{3.11}$$

$\mathbf{R_{yy}}$ and $\mathbf{r_{yd}}$ denote the auto-correlation matrix of the input and the cross-correlation vector of the input & desired signals, respectively [37]. Then, (3.11) can be given by

$$J_{MSE}(\mathbf{w}) = \sigma_d^2 - 2\mathbf{w}^H[k]\mathbf{r_{yd}} + \mathbf{w}^H[k]\mathbf{R_{yy}}\mathbf{w}[k], \tag{3.12}$$

where σ_d^2 indicates the variance of the desired response $d[k]$ [38]. Differentiating this equation with respect to $\mathbf{w}^H[k]$ yields

$$\frac{\partial J_{MSE}(\mathbf{w})}{\partial \mathbf{w}^H[k]} = -2\mathbf{r_{yd}} + 2\mathbf{R_{yy}}\mathbf{w}[k]. \tag{3.13}$$

With the assumption that $\mathbf{R_{yy}}$ is non-singular, the famous Wiener-Hopf optimal solution for filter designing is derived by setting (3.13) to zero

$$\mathbf{w}[k] = \mathbf{R_{yy}}^{-1}\mathbf{r_{yd}}. \tag{3.14}$$

3.3.1.1 Multi-channel Wiener Filter

In a multi-channel fashion, Multi-channel Wiener Filter (MWF) leverages a microphone array containing *N* elementary microphones. Each microphone acts as an individual channel, delivering its own observed signal to the corresponding Wiener filter. The MWF estimates the speech signal $x[k]$ by combining the outputs of all microphones $\mathbf{y}_n[k]$ [15]. Assuming all Wiener filters have a length of L, i.e. each has L taps, according to (3.9), the output signal $z[k]$ can be derived as

$$z[k] = \sum_{n=0}^{N-1} \mathbf{w}_n^*[k]\mathbf{y}_n[k] = \mathbf{w}^H[k]\mathbf{y}[k], \tag{3.15}$$

where input signals and filter coefficients are now described as stacked vectors for all secondary channels. Figure 3.4 illustrates the above-mentioned architecture of a MWF. It should be noticed that there is only one channel containing the desired signal, i.e. this particular microphone actually faces towards the direction of speech [28].

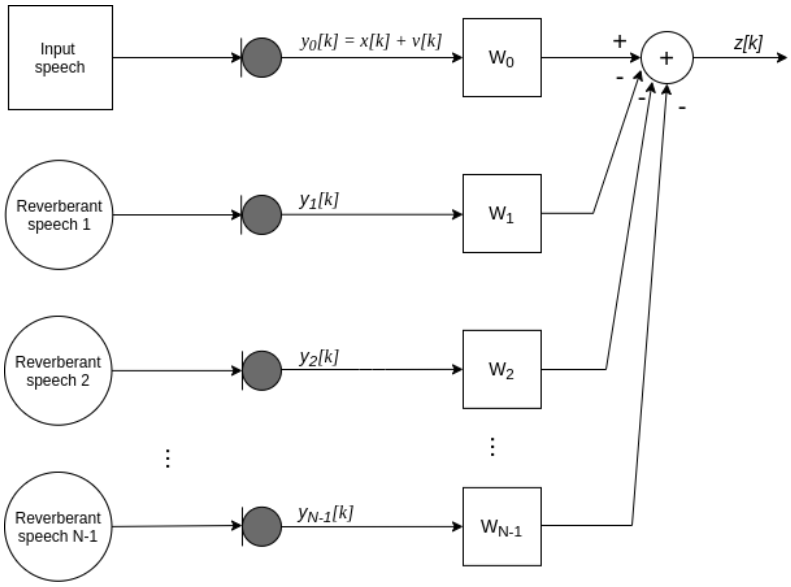

Figure 3.4. *Basic configuration of a Multi-channel Wiener Filter*

Along with the MSE optimization problem, by forcing some linear constraints on the filter weights, we can reproduce the LCMV algorithm that was described in section 3.2.1 [37]. The MWF approach assumes that the desired speech signal is always known, which is infeasible in real-life conversations. Besides, the MWF requires correct approximations of the second order statistics of the noise to generate the clean speech, which might not be achievable in those complex acoustic situations [39]. Therefore, a linearly optimized method such as BLCMV or GSC is more favorable in this case.

3.3.1.2 Spatially Pre-processed Speech Distortion Weighted Multi-channel Wiener Filter

First presented in 2004 by Spriet and her colleagues, Spatially Pre-processed Speech Distortion Weighted Multi-channel Wiener Filter (SP-SDW-MWF) is an advanced MWF-based NR approach that compromises between the speech distortion and the noise reduction [29]. Figure 3.5 shows the basic components in a SP-SDW-MWF system. The algorithm leverages the *spatially pre-processing* part of the GSC illustrated in section 3.2.2 that takes care of the *look direction* requirement [20]. The mission of the later SDW-MWF adaptive filtering stage is to reduce the speech distortion effect while estimating the noise component. Later on, the noise approximation will be subtracted from the *speech reference* to produce a purified speech signal [30].

The term Speech Distortion Weighted (SDW) indicates that instead of considering the output $z[k]$ as a whole as in section 3.3.1.1, the filter adjusts the residual noise and the speech distortion by introducing a trade-off parameter μ [40]. The MSE cost function of

the SDW-MWF part can be written as

$$\mathbf{J}_{MSE}(\mathbf{w}) = \mathbb{E}\{(\mathbf{x}_{ref}[k] - \mathbf{w}^H[k]\mathbf{x}[k])^2\} + \mu\mathbb{E}\{(\mathbf{w}^H[k]\mathbf{v}[k])^2\}, \tag{3.16}$$

where $\mathbf{x}_{ref}[k]$ describes the pure speech signal at time *k*.

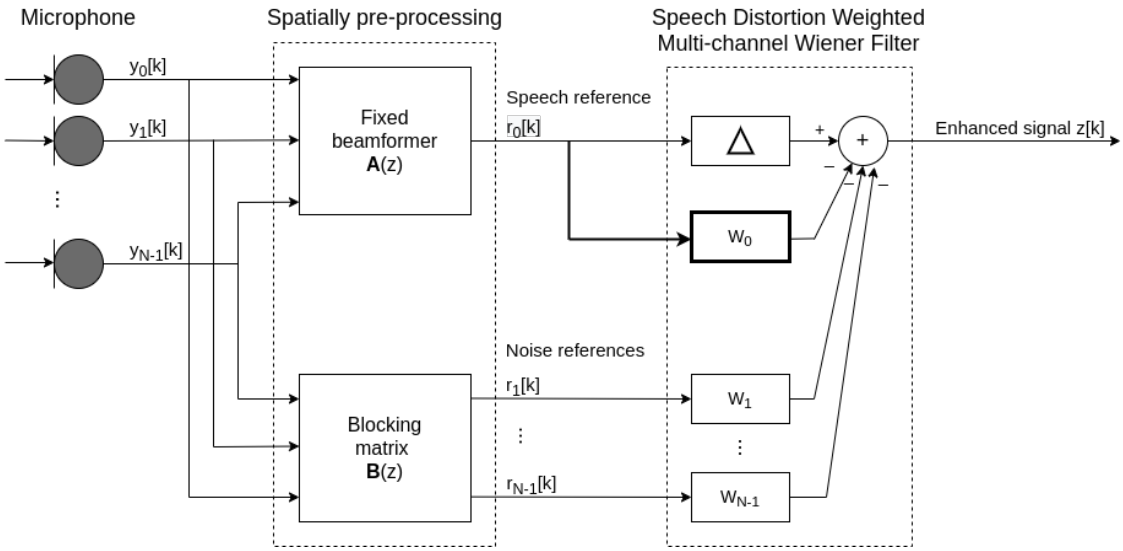

Figure 3.5. *Structure of a SP-SDW-MWF*

By adjusting $\mu > 1$, the filter emphasizes more on noise suppression at the expense of higher speech distortion level [41]. With $\mu < 1$, the algorithm emphasizes on the speech distortion and becomes the original GSC when $\mu = 0$ [29].

The SP-SDW-MWF method increases the system robustness against large signal model errors because it does not depend on prior information about the acoustic scheme [28]. It also solves the need for accurate temporal estimation of the standalone MWF (SDW-MWF in this case), which is to determine the periods when the speech or noise dominates and reduce the *speech leakage*, by introducing the pre-processing part of the GSC. Compared to the advanced QIC-QSC, the SP-SDW-MWF demonstrates a superior noise suppression capability for a known level of speech distortion [40]. Combining the merits of both the QSC and the Wiener-based filtering, the SP-SDW-MWF seems to be the perfect solution for low-power hearing aids, except that the limited resources there might curtail the noise reduction performance.

3.4 Feedback Cancellation

In any audio system, it is possible for a closed feedback loop to exist between the audio output and input when they are placed close to each other. In hearing-aid systems particularly, the desired speech signal is captured by the microphones, processed and passed to the receivers where it will then be delivered to the wearers' ear canal [42]. The

audio feedback problem, referred to as *acoustic feedback*, occurs when some parts of the output from the receiver are accidentally fed to the microphone [43]. Those feedback signals are amplified further and further when they get back to the input, and might damage the quality of the desired signal and alter the hearing-aid operations when it reaches a significant loudness [4]. This results in perceptible audio artifacts and even howling effect under certain conditions [43] . Many automatic feedback cancellation (FBC) methods are proposed, but the most auspicious and well-known strategy is the adaptive feedback cancellation (AFC) [44] using adaptive filters, some of which are the recursive implementations of the Wiener filter [36]. Under the scope of this work, the most common AFC algorithms are presented, explicitly Least Mean Square (LMS) & its derivation Normalized LMS (NLMS) and Affine Projection Algorithm (APA).

3.4.1 Least Mean Square

The Least Mean Square algorithm, invented by B. Widrow and T. Hoff in 1960, is one of the most important adaptive filtering techniques using the stochastic gradient descent method [38]. The LMS is best known for its simplicity and wide applicability range. Nevertheless, its main limitations are slow convergence speed and higher steady-state error with unknown systems [45].

To begin with, the LMS algorithm was constructed based on the combination of the *steepest-descent algorithm* (SDA) [46] and the Wiener Filter described in section 3.3.1. Apart from the Wiener solution $\mathbf{w}_{WF}$ that minimizes the total error (Least Square), the LMS solution $\mathbf{w}_{LMS}$ reduces the error of the current sample in an adaptive manner, hence the name Least Mean Square. Consequently, (3.13) can be applied to the LMS as

$$\frac{\partial J(\mathbf{w}_{LMS}[k])}{\partial \mathbf{w}_{LMS}[k]} = -2\mathbf{y}[k]e^*[k]. \tag{3.17}$$

The LMS adaptation is defined by substituting (3.17) to the SDA solution

$$\mathbf{w}_{LMS}[k+1] = \mathbf{w}_{LMS}[k] + \mu\mathbf{y}[k]e^*[k]. \tag{3.18}$$

The operating principle of the LMS filtering is displayed in figure 3.6. When the input signal $\mathbf{y}[k]$ and the LMS coefficient vector $\mathbf{w}_{LMS}[k]$ are statistically independent, the LMS algorithm *converges in the mean* [47] if

$$0 < \mu < \frac{2}{\lambda_{max}}, \tag{3.19}$$

where λ_{max} indicates the maximum value of the correlation matrix of the input $\mathbf{y}[k]$ [38].

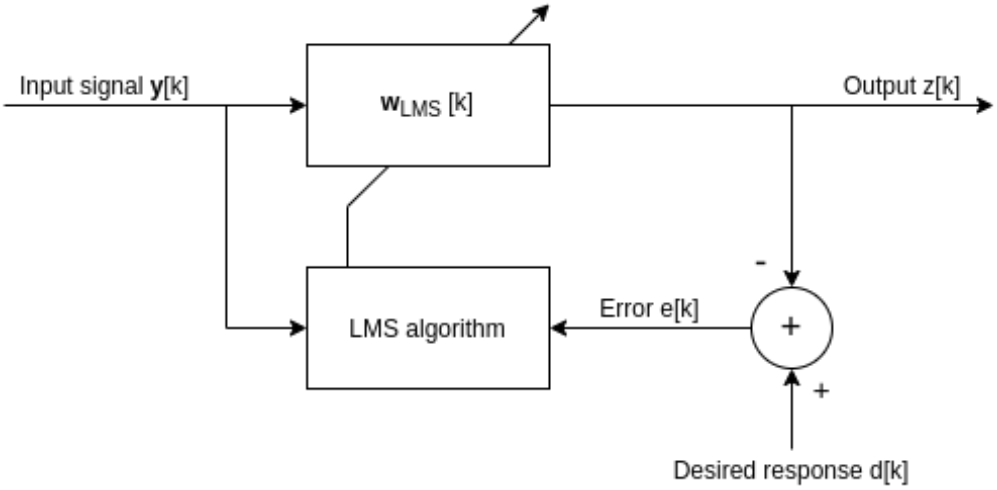

Figure 3.6. *The LMS adaptive filter*

The LMS technique does not require extra memory, as it depends only on current-time information of the signals [48]. This is a crucial criterion for the algorithm to be considered for low-power HAs. Besides the AFC, the LMS is also widely-used for other applications such as channel estimation or equalization [49].

3.4.2 Normalized Least Mean Square

According to (3.18), the LMS algorithm adjusts the weight vector $\mathbf{w}_{LMS}$ proportionally to the input data as well as the *priori* error from the previous iteration cycle. Therefore, for large input signals, it will face the *gradient noise amplification* problem [38] and may eventually corrupt the whole system. The revised version of LMS, called Normalized Least Mean Square (NLMS), tackles this problem efficiently by setting the step-size variable μ *time-variant* and *normalized* with respect to the squared Euclidean norm of the input vector $\mathbf{y}[k]$ [50] as

$$\mu[k] = \frac{\hat{\mu}}{\|\mathbf{y}[k]\|^2}, \tag{3.20}$$

where $\hat{\mu}$ is a positive scaling factor called *normalized step size* [47] to control the weight adaptation. In addition, to avoid large step sizes with very small input energy, a random small positive value α is introduced to the denominator [37], and the NLMS adaptation can finally be expressed as

$$\mathbf{w}_{NLMS}[k+1] = \mathbf{w}_{NLMS}[k] + \frac{\hat{\mu}\mathbf{y}[k]e[k]}{\alpha + \|\mathbf{y}[k]\|^2}. \tag{3.21}$$

Compared to the traditional LMS algorithm, the updated NLMS shows a potentially faster convergence speed for both uncorrelated and correlated input signals [38]. However, this comes at the expense of heavier computation for the norm $\|\mathbf{y}[k]\|^2$ [47].

3.4.3 Affine Projection Algorithm

First introduced by K. Ozeki and T. Umeda in 1984, the fundamentals of Affine Projection Algorithm (APA) class of adaptive filters showed a superior performance for correlated inputs than LMS class in terms of convergence speed at a reasonable computation expense [51]. While NLMS updates the weight vector $\mathbf{w}$ on the basis of the current-time input signal, APA also relies on its $M-1$ predecessors [52]. Apparently, the NLMS is a special case of the generalized APA algorithm with $M=1$. The *regressor block* of *M* latest inputs is formed as a $L \times M$ matrix where L is the length of each filter [46]

$$\mathcal{Y}^T[k] = \begin{bmatrix} \mathbf{y}[k] & \mathbf{y}[k-1] & \mathbf{y}[k-2] & \dots & \mathbf{y}[k-M+1] \end{bmatrix}. \tag{3.22}$$

The APA directly stores the input and $M-1$ of its previous values, whereas the MWF collects the current input and then delayed versions are produced by passing it to the filter taps. The error signal of the APA filter is now derived as a vector [45]

$$\mathbf{e}[k] = \mathbf{d}[k] - \mathcal{Y}^H[k]\mathbf{w}_{APA}[k]. \tag{3.23}$$

By means of *Affine Projection*, hence the name Affine Projection Algorithm, the APA solution for adaptive signal filtering is defined as

$$\mathbf{w}_{APA}[k+1] = \mathbf{w}_{APA}[k] + \hat{\mu}\mathcal{Y}^H[k](\mathcal{Y}[k]\mathcal{Y}^H[k])^{-1}\mathbf{e}[k]. \tag{3.24}$$

where $\mathcal{Y}[k]\mathcal{Y}^H[k]$ is a non-singular invertible Gramian matrix under the statistically independence assumption [46]. To address the numerical matrix inversion problem, a process referred to as *regularization* is executed by adding a small partition $\delta\mathbf{I}$ where δ is a small positive number called *regularization factor* and $\mathbf{I}$ denotes the identity matrix [53]. This procedure defines the so-called *Regularized APA* (R-APA) [46], in which the adaptation operation for the filter coefficients is stated as

$$\mathbf{w}_{APA}[k+1] = \mathbf{w}_{APA}[k] + \hat{\mu}\mathcal{Y}^H[k](\delta\mathbf{I} + \mathcal{Y}[k]\mathcal{Y}^H[k])^{-1}\mathbf{e}[k]. \tag{3.25}$$

The APA algorithm brings a remarkably higher performance over NLMS even with $M=2$ [45], especially for correlated inputs such as speech [48], with an acceptable computational complexity. As the filter order M rises, meaning that more input vectors are required, the convergence speed is boosted, whilst the extent of improvement decreases [52]. Therefore, Benesty [45] had recommended M to be in the range $2-5$ for the best cancellation achievement. The aforementioned criteria make the APA one of the most important AFC techniques in this field. Nonetheless, in hearing-aid domain where battery lifetime is the most important concern, APA might not be as well-suited as the LMS, since it exhausts the computational resources when storing a huge amount of signal data.

4. EFFICIENT LOW-POWER IMPLEMENTATIONS OF DIGITAL HEARING-AID TECHNIQUES

Based on the algorithms introduced in the previous chapter, this chapter is devoted for some of the low-power implementations in the digital hearing-aid domain. They are FFPs produced in different platforms such as FPGA or custom ASIC, but the latter is preferred due to its power efficiency.

4.1 Adaptive Beamforming

In 2002, Luo and her co-workers [54] presented an *adaptive null-forming* scenario for noise reduction in hearing aids. In his work, null indicates the directions from which sound waves are mostly attenuated. The main obstacle of the algorithm is to precisely steer the microphones in a way such that the null of the beamforming pattern of the systems points towards the noise. The system leverages the endfire microphone orientation with a delay of d/c, where d is the distance between the microphones and c implies the speed of the sound [54]. The polar pattern is dynamically controlled by an adaptive gain using LMS algorithm. The gain (or weight) is related to the null direction as

$$\mathbf{w}[k] = \frac{\sin\left(\pi f \frac{d}{c}(1+\cos\theta_{null})\right)}{\sin\left(\pi f \frac{d}{c}(1-\cos\theta_{null})\right)}, \tag{4.1}$$

where f indicates the signal frequency and θ_{null} is the angle between the null and the microphone plane. With the approximation that $\sin\theta \approx \theta$, this filter gain is frequency-independent as

$$\mathbf{w}[k] = \frac{1+\cos\theta_{null}}{1-\cos\theta_{null}}. \tag{4.2}$$

In 2017, Samtani proposed both ASIC [55] and FPGA [56] implementations of the above null-forming beamformer; however, in this thesis, we concentrate on the power-efficient ASIC design. The schematic of the adaptive beamformer is displaed in figure 4.1. In this design, the adaptive filtering method LMS is applied to estimate the signal gain [56]

$$as\mathbf{w}[k+1] = \mathbf{w}[k] + 2\mu\mathbf{y}[k]\mathbf{b}[k]. \tag{4.3}$$

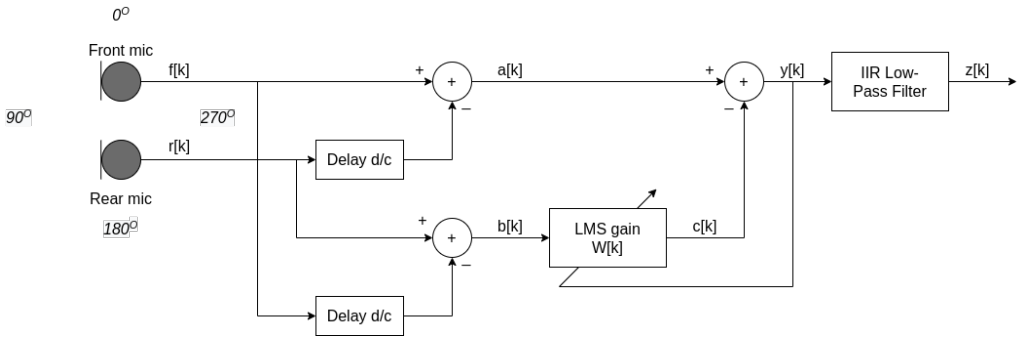

Figure 4.1. *Schematic of the adaptive null-forming beamformer*

His system employs a simple first-order IIR low-pass filter to compensate for the relation between the microphone output and the signal frequency [56], of which the impulse response is given by

$$H[z] = \frac{C_1 + C_2 z^{-1}}{1 - C_3 z^{-1}}, \tag{4.4}$$

where $C1$, $C2$ and $C3$ are some user-defined constants. Both fixed beamformer and LMS module are realized as Finite State Machine, while the IIR low-pass filter is implemented using MATLAB where $C1 = C2 = 0.2759$ and $C3 = 0.9758$ [55].

The system initially deploys 5 Radix-2 Booth's multipliers [57] and three addition operators. Later on, due to some optimizations in computing, the authors successfully reduce to only one multiplier that is active at all states. Eventually, this Radix-2 multiplier is supplanted by a Radix-4 one, which reduces the dynamic power dissipation significantly with a reasonable area score [55]. The power and area result of the implementations with those multipliers is depicted in table 4.1 [55]. There is a 93% reduction in total power and a 47% reduction in area score, which are extremely high for just a component substitution.

Architecture	Power (μW)			Area (μm^2)
	Static	Dynamic	Total	
5 radix-2 multipliers	1.39	839.70	841.10	102879.682
2 radix-2 multipliers	0.92	191.32	192.24	69895.902
1 radix-2 multiplier	0.72	106.62	107.34	57191.961
1 radix-4 multiplier	0.68	59.86	60.54	54216.081

Table 4.1. *Area and power result of the adaptive beamformer*

The architecture was first functionally tested on Xilinx® Artix 7 FPGA family before and then implemented in ASIC using 0.18 μm SCL library [58] and Cadence® EDA tools. The result of the two test cases is illustrated in table 4.2 [55]. The proposed beamformer shows noticeable SNR gains with different null directions and input signal characteristics.

Noise type	Null direction (°)	SNR gain (dB)
Monotone	180	20.81
	150	22.10
	105	21.21
Speech signal	180	15.08
	150	15.12
	105	14.92

Table 4.2. *SNR gain of the adaptive beamformer*

4.2 Feedback Cancellation with LMS algorithm

As discussed in section 3.4, shifting from current-time dependency in (Normalized) LMS to multiple input operation in APA rises the accuracy of the feedback purification but also the complexity of the algorithms. Hence, towards the low-power hearing-aid applications, LMS seems to be a more potential solution for AFC because of its ease of implementation and plausible performance. In 2003, Kim [59] proposed a novel implementation of the so-called Delay LMS (DLMS) based on the *pseudo nMOS* logic technology.

With regard to the electronics, Kim leveraged pseudo nMOS in the sub-threshold region, referred to as *sub-pseudo nMOS*. Here, sub-threshold or weak-inversion region means the supply voltage is less than the regular value, which are in turns $0.4V$ and $3.3V$ in this case. Sub-threshold digital logics such as sub-pseudo nMOS are extraordinarily suitable for applications which prioritize small power consumption over performance, including hearing aids [60]. They not only preserve the outstanding characteristics of the standard CMOS, which are exceptional robustness, wide noise margin and low power usage, but also show a better transconductance and hence improve the voltage transfer characteristics (VTC). In the weak-inversion region with V_{DD} = 0.5 V, the pseudo nMOS consumed a bit more power, but outperformed the standard CMOS in terms of delay and power-delay product (PDP). The actual comparison for this particular voltage is displayed in table 4.3.

Gate	CMOS			Pseudo nMOS		
	Power	Delay	PDP	Power	Delay	PDP
INVERSE	29 nW	68 ns	1.98 fJ	31 nW	45 nW	1.40 fJ
NAND	33 nW	124 ns	4.13 fJ	25 nW	77 nW	1.97 fJ
NOR	36 nW	133 ns	4.78 fJ	46 nW	44 nW	2.01 fJ
XOR	84 nW	430 ns	36.00 fJ	76 nW	248 nW	18.90 fJ

Table 4.3. *Statistics of the folded LMS and non-folded DLMS*

In the algorithm selection, the author analyzed the performance of *folded* LMS and *non-folded* DLMS with filter length of L. The term *non-folded* implies a parallel platform with numerous functional units (FU) and pipeline phases that can be used for a single-cycle execution, whereas *folded* architecture resembles a GPP that employs only one single FU for all operations [59]. The conventional LMS cannot take advantage of the pipelining, as the adaptation may only be applied when the error calculation is valid [61]. Meanwhile, introducing a constant delay to the LMS filter enables it to exploit the parallelism [62] and thus reduce the power dissipation of the system. Both algorithms considering real numbers are summarized in table 4.4 [63].

	Folded LMS	Non-folded DLMS
Number of FUs	1	L
Error calculation	$e[k] = d[k] - \mathbf{w}^T[k]\mathbf{y}[k]$	$e[k-L] = d[k-L] - \mathbf{w}^T[k-L]\mathbf{y}[k-L]$
Adaptation	$\mathbf{w}[k+1] = \mathbf{w}[k] + \mu\mathbf{y}[k]e[k]$	$\mathbf{w}[k+1] = \mathbf{w}[k] + \mu\mathbf{y}[k-L]e[k-L]$

Table 4.4. *Statistics of the folded LMS and non-folded DLMS*

The conventional LMS experienced a faster rate of convergence and smaller achieved MSE as expected, because the DLMS was designed to reduce the power consumption at the cost of poorer performance. The proposed architecture was able to meet the desired clock frequency of 22 kHz when the supply voltage was scaled down to just 0.4 V. In general, the sub-threshold implementation of the standard CMOS saved 87% of the total active energy. However, this came with an area trade-off where the total transistors used increased by 3.6 times. Moreover, the design achieved a further reduction of 27% in energy efficiency by replacing the CMOS with the pseudo nMOS architecture [59]. Table 4.5 summarizes the design configuration and aforementioned result.

Algorithm	Architecture	Clock frequency	Supply voltage	Energy per operation	Transistor count
LMS	Folded standard CMOS	748 kHz	0.65 V	21.71 nJ	31k
DLMS	Non-folded Sub-CMOS	22 kHz	0.45 V	2.80 nJ	111k
DLMS	Non-folded Sub-Pseudo nMOS	22 kHz	0.40 V	2.05 nJ	86k

Table 4.5. *Implementation result for three AFC schemes*

5. HEARING-AID ARCHITECTURES

The Digital Signal Processor (DSP) is the heart of digital hearing aids, where all audio processing techniques are executed to bring the most purified speech signals to the wearers. Recent technology developments reduces the size of the dies, which allows the designers to adapt more advanced audio processing techniques. In this chapter, we split the DSP architecture into three main types, specifically FFP, ASIP and ASIP with hardware accelerators.

5.1 Fixed-function Processor

In the discussed digital fixed-function processor (FFP) architecture, all functioning parts are soldered before manufacturing and cannot be changed afterwards [64]. Despite this drawback, the FFPs bring a noticeable performance and a considerable cutback in power consumption [65]. A digital low-power FFP implementation, presented in [66], made a compromise between accuracy and power efficiency by using stochastic approximation. The design consists of a FIR digital filter which was hardwired shifted, ripple-carry adders and multipliers for the gain and volume adjustments, multiplexers and internal storage such as D Flip-Flops to create delays. The authors decided to use the standard CMOS style mostly due to its robustness against the stochastic estimation [66].

The digital filter was coded in VHDL, and then verified with MP3 test files as well as Synopsys HSPICE® simulation output [66]. Table 5.1 describes the result when the voltage were scaled from 0.8 V to 0.4 V. Compared to the case when V_{DD} = 0.8 V, setting it to 0.4 V yielded a 82% reduction in power dissipation without any functional error [66].

Supply voltage (V)	Power consumption (μW)
0.8	120.00
0.7	86.22
0.6	57.36
0.5	37.08
0.4	21.93

Table 5.1. *Power result*

This DSP architecture considers the digital audio signals. Besides, there are other exceptional low-power implementations for analog [67, 68] or mixed-signal [69, 70] that are not cover in this thesis work.

5.2 Application-Specific Instruction-set Processor

As described in section 2.3, Application-Specific Instruction-set Processor (ASIP) refers to the architecture where the instruction set of a processor is customized for a particular purpose. Programmable ASIP provides a higher flexibility for different applications, which reduces the development and manufacturing costs. However, this feature also leads to a commonly greater power dissipation than the hardwired versions [64]. To handle this problem, a remarkable hearing-aid processing unit with ASIP standard was presented in 2011 that competed with a strict specification: silicon area of 22 mm^2 and a battery lifetime of at least 50 hours with 1.35 V voltage supply in [45]. The structure of the proposed system is displayed in figure 5.1.

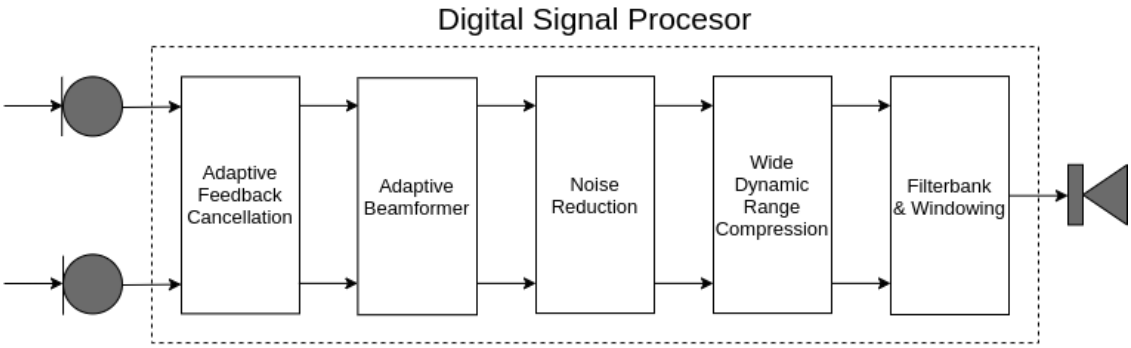

Figure 5.1. *Structure of the DSP as ASIP*

The proposed hearing-aid DSP consists of all essential building blocks that are FBC, NR, BMF, Filterbank (FB) and Wide Dynamic Range Compression (WDRC). Familiar audio processing techniques are adopted, particularly LMS for FBC, GSC for BMF and wrapped FIR filter for FB. The noise suppression [71, 72] and dynamic range compression [73] were done by the combination of FIR filters and frequency-domain computing methods. To further improve the processing capability, several optimizations were applied to both hardware and software sides such customized processor instructions, loop optimizations (merging, unrolling and cache involved), circular buffer, etc [74].

At the clock frequency of 11 MHz and supply voltage of 0.8 V, the system consumes only 0.964 mW. Based on this DSP, a full hearing-aid chip could be implemented, of which the area occupation and battery lifetime was 0.49 mm^2 and over 300 hours [74].

5.3 ASIP with hardware accelerators

Even though the standalone ASIP brings many valuable merits, it is still slow and power-consuming for complex arithmetic operations compared to hardwired versions [75]. The performance can be improved by relocating sophisticated and intensive computing tasks onto different specialized hardware accelerators [64]. On the other hand, the central ASIP is now responsible for light parallel computing, data management and controlling the whole system [76]. Based on this concept, a high-performance DSP architecture [77] was introduced in 2019. Figure 5.2 demonstrates the block diagram of the design.

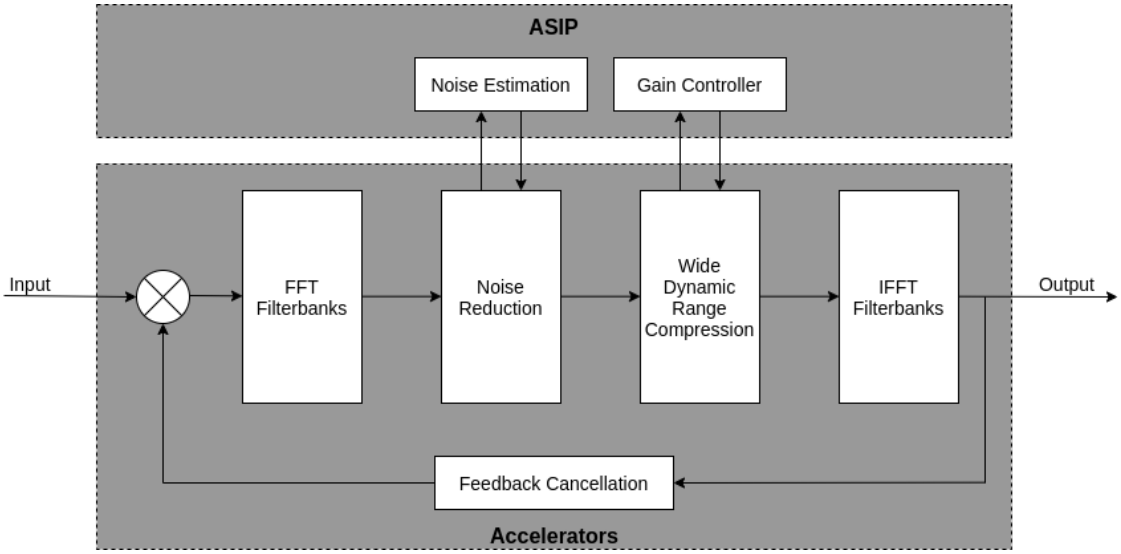

Figure 5.2. *Block diagram of the high-performance DSP*

The proposed configuration represents a rational combination of ASIP and ASIC. Heavy computations such as FB, FBC, WDRC and most of NR are executed on dedicated accelerators, which are ASICs in this case for power and performance constraint [77]. Meanwhile, the noise estimation and gain controller of WDRC are implemented on ASIP, because they depend on actual user preferences as well as the environment conditions [77]. The adopted algorithms for each processing block are:

- 128-point FFT/IFFT Filter Banks
- Multiband spectral subtraction (*mband*) [78] and Voice activity detection [79] for Noise Reduction
- Subband adaptive filter (SAF) [80] for Feedback Cancellation

The design mostly focuses on frequency-domain processing, which introduces the need for FFT/IFFT filter banks. Indeed, it is the first implemented System-on-Chip (SoC) for low-power hearing aids that leverages the 128-point architecture [77].

The accelerators are designed as ASICs, while the ASIP is taken from Xilinx Zynq® FPGA with some pipeline optimizations and customized instruction sets with respect to the noise

estimation method [77]. The authors made some comparisons in noise suppression level with other NR techniques in [81] using the NOIZEUS database [82]. They revealed that the *mband* algorithm showed a great balance between computation complexity and NR level; however, other NR methods which are specialized for particular noise patterns can also be applied due to the flexibility of ASIP [77].

	2011 [74]	2012 [65]	2014 [68]	2015 [83]	2019 [84]	2019 This work
Technology	Digital	Digital	Analog	Mixed-signal	Mixed-signal	Digital
Architecture	ASIP only	ASIP+acc.	FFP	ASIP+acc.	ASIP+acc.	ASIP+acc.
Algorithms	NR, WDRC, BMF, FBC	NR, WDRC, FBC	DRC, AGC	NR, WDRC, FBC	NR, WDRC, FBC	NR, WDRC, FBC
Power	964 μW	1.3 mW	1.6 mW	1.2 mW	1.1 mW	1.3 mW

Table 5.2. *Comparison of different hearing-aid systems*

The implementation of Kim's work is compared with other efficient related models, which is illustrated in table 5.2. The analog platform seems to be outperformed by the mixed-signal and especially digital versions. The result also illustrates the superior in power consumption of ASIP architecture with or without accelerators over the FFPs, where the analog FFP in [68] has the worst power dissipation with the least number of algorithms adapted. This can be explained by the hardware and power cost of analog ports and processing blocks. Additionally, mixed-signal designs require some front-end blocks for transferring and receiving analog signals, which might be costly than digital platforms in terms of hardware expense and power optimization. Consequently, digital ASIP with accelerators is the most promising implementation architecture for DSPs in modern low-power hearing aids.

6. CONCLUSION

This thesis analyzes some of the state-of-the-art signal processing techniques in digital hearing-aid applications, particularly Beamforming, Noise Reduction and Feedback Cancellation. Generally, the rudimentary approaches are still widely-used due to their simplicity and adequate performance. These methods can also be efficiently combined to preserve the quality and intelligibility of the speech as well as the acoustic schemes. Their advanced versions might be applied in the future when modern technology permits a larger amount of processing units to be integrated on small circuits as in low-power hearing aids. The researching also shows two power-efficient implementations of adaptive beamformer and feedback cancellation to visualize the application of the discussed Least-Mean Square algorithm.

In the sequel, three main types of latest digital signal processors (DSPs) are presented with detailed specifications. Comparisons between those DSP platforms, specifically FFP, ASIP and ASIP with accelerators, are made to identify the most outstanding architecture for low-power hearing-aid purposes. Digital and mixed-signal implementations are replacing analog versions because of their broader ranges of supported algorithms and power efficiency. The FFP is outperformed by ASIPs, which is unexpected because FFPs are well-known for their solid computing performance and low power consumption. In conclusion, digital ASIP with several hardware accelerators is best-suited for the low-power hearing-aid applications due to its low power dissipation and several intricate audio processing algorithms involved.

Our work can be further expanded in several ways. First, other major processing parts of the hearing aids such as Dynamic Range Compression or Amplification can be covered. Second, filter banks are worthwhile to mention since most of the findings relies on the FIR Filters. Finally, some full-chip implementations of low-power hearing aids can be included to show how the introduced DSPs are integrated in a complete system.

REFERENCES

[1] Yueh, B., Shapiro, N., MacLean, C. H. and Shekelle, P. G. Screening and Management of Adult Hearing Loss in Primary CareScientific Review. *JAMA* 289.15 (Apr. 2003), pp. 1976–1985. ISSN: 0098-7484. DOI: 10.1001/jama.289.15.1976.

[2] Lin, F. R., Metter, E. J., O'Brien, R. J., Resnick, S. M., Zonderman, A. B. and Ferrucci, L. Hearing Loss and Incident Dementia. *Archives of Neurology* 68.2 (Feb. 2011), pp. 214–220. ISSN: 0003-9942. DOI: 10.1001/archneurol.2010.362.

[3] Martin, R., Heute, U. and Antweiler, C. Binaural Signal Processing in Hearing Aids: Technologies and Algorithms. *Advances in Digital Speech Transmission*. 2008, pp. 401–429. DOI: 10.1002/9780470727188.ch14.

[4] Harvey, D. *Hearing aids*. New York, USA: Thieme. United States: Thieme Medical Publishers, Incorporated, 2012. ISBN: 9781604068108.

[5] Lee, S. J., Kim, S. and Yoo, H. A Low Power Digital Signal Processor with Adaptive Band Activation for Digital Hearing Aid Chip. *2007 IEEE International Symposium on Circuits and Systems*. 2007, pp. 2730–2733. DOI: 10.1109/ISCAS.2007.378526.

[6] Eargle, J. *The Microphone Book*. Audio Engineering Society Presents Series. Focal Press, 2004. ISBN: 9780240519616.

[7] Zölzer, U. *Digital Audio Signal Processing*. Wiley, 1997. ISBN: 9780471972266.

[8] Werner, N., Payá-Vayá, G. and Blume, H. Case Study: Using the Xtensa LX4 Configurable Processor for Hearing Aid Applications. Nov. 2013.

[9] Gerlach, L., Payá-Vayá, G., Liu, S., Weißbrich, M., Blume, H., Marquardt, D. and Doclo, S. Analyzing the trade-off between power consumption and beamforming algorithm performance using a hearing aid ASIP. *2017 International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS)*. 2017, pp. 88–96. DOI: 10.1109/SAMOS.2017.8344615.

[10] Hartig, J., Gerlach, L., Payá-Vayá, G. and Blume, H. Customizing a VLIW-SIMD application-specific instruction-set processor for hearing aid devices. *2014 IEEE Workshop on Signal Processing Systems (SiPS)*. 2014, pp. 1–6. DOI: 10.1109/SiPS.2014.6986072.

[11] Customized Parallel Computing group. *TTA-based Co-Design Environment (TCE)*. URL: http://openasip.org/.

[12] Smith, M. *Application-Specific Integrated Circuits*. Addison-wesley Vlsi Systems. Addison Wesley Professional, 1997. ISBN: 9780321602756.

[13] Barkalov, A., Titarenko, L. and Mazurkiewicz, M. *Foundations of Embedded Systems*. Studies in Systems, Decision and Control. Springer International Publishing, 2019. ISBN: 9783030119614.

[14] Saha, A. and Manna, N. *Digital Principles and Logic Design*. Engineering series. Jones & Bartlett Learning, LLC, 2009. ISBN: 9780763773731.

[15] Doclo, S. and Moonen, M. GSVD-based optimal filtering for single and multimicrophone speech enhancement. *IEEE Transactions on Signal Processing* 50.9 (2002), pp. 2230–2244. DOI: 10.1109/TSP.2002.801937.

[16] Benesty, J., Sondhi, M. and Huang, Y. *Springer Handbook of Speech Processing*. Springer Handbook of Speech Processing. Springer Berlin Heidelberg, 2007. ISBN: 9783540491279.

[17] Souden, M., Benesty, J. and Affes, S. On Optimal Frequency-Domain Multichannel Linear Filtering for Noise Reduction. *IEEE Transactions on Audio, Speech, and Language Processing* 18.2 (2010), pp. 260–276. DOI: 10.1109/TASL.2009.2025790.

[18] Frost, O. L. An algorithm for linearly constrained adaptive array processing. *Proceedings of the IEEE* 60.8 (1972), pp. 926–935. DOI: 10.1109/PROC.1972.8817.

[19] Hadad, E., Gannot, S. and Doclo, S. Binaural Linearly Constrained Minimum Variance Beamformer for Hearing Aid Applications. *IWAENC 2012; International Workshop on Acoustic Signal Enhancement*. 2012, pp. 1–4.

[20] Haykin, S. and Liu, K. *Handbook on Array Processing and Sensor Networks*. Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control. Wiley, 2010. ISBN: 9780470487051.

[21] Yu, L., Lin, N., Liu, W. and Langley, R. Bandwidth Performance of Linearly Constrained Minimum Variance Beamformers. *2007 International workshop on Antenna Technology: Small and Smart Antennas Metamaterials and Applications*. 2007, pp. 327–330. DOI: 10.1109/IWAT.2007.370141.

[22] Van Trees, H. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Detection, Estimation, and Modulation Theory. Wiley, 2004. ISBN: 9780471463832.

[23] Capon, J. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE* 57.8 (1969), pp. 1408–1418. DOI: 10.1109/PROC.1969.7278.

[24] Hadad, E., Doclo, S. and Gannot, S. The Binaural LCMV Beamformer and its Performance Analysis. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 24.3 (2016), pp. 543–558. DOI: 10.1109/TASLP.2016.2514496.

[25] Koutrouvelis, A. I., Hendriks, R. C., Heusdens, R. and Jensen, J. Relaxed Binaural LCMV Beamforming. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 25.1 (2017), pp. 137–152. DOI: 10.1109/TASLP.2016.2628642.

[26] Griffiths, L. and Jim, C. An alternative approach to linearly constrained adaptive beamforming. *IEEE Transactions on Antennas and Propagation* 30.1 (1982), pp. 27–34. DOI: 10.1109/TAP.1982.1142739.

[27] Benesty, J., Chen, J., Huang, Y. and Dmochowski, J. On Microphone-Array Beamforming From a MIMO Acoustic Signal Processing Perspective. *IEEE Transactions on Audio, Speech, and Language Processing* 15.3 (2007), pp. 1053–1065. DOI: 10.1109/TASL.2006.885251.

[28] Doclo, S., Spriet, A., Wouters, J. and Moonen, M. Speech Distortion Weighted Multichannel Wiener Filtering Techniques for Noise Reduction. *Speech Enhancement*. Springer Berlin Heidelberg, 2005, pp. 199–228. ISBN: 978-3-540-27489-6. DOI: 10.1007/3-540-27489-8_9.

[29] Spriet, A., Moonen, M. and Wouters, J. Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction. *Signal Processing* 84.12 (2004), pp. 2367–2387. ISSN: 0165-1684. DOI: https://doi.org/10.1016/j.sigpro.2004.07.028.

[30] Doclo, S., Spriet, A., Wouters, J. and Moonen, M. Frequency-Domain Criterion for the Speech Distortion Weighted Multichannel Wiener Filter for Robust Noise Reduction. *Speech Communication* 49.7–8 (July 2007), pp. 636–656. ISSN: 0167-6393. DOI: 10.1016/j.specom.2007.02.001.

[31] Breed, B. R. and Strauss, J. A short proof of the equivalence of LCMV and GSC beamforming. *IEEE Signal Processing Letters* 9.6 (2002), pp. 168–169. DOI: 10.1109/LSP.2002.800506.

[32] Van Veen, B. D. and Buckley, K. M. Beamforming: a versatile approach to spatial filtering. *IEEE ASSP Magazine* 5.2 (1988), pp. 4–24. DOI: 10.1109/53.665.

[33] Spriet, A., Moonen, M. and Wouters, J. Robustness analysis of GSVD based optimal filtering and generalized sidelobe canceller for hearing aid applications. *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575)*. 2001, pp. 31–34. DOI: 10.1109/ASPAA.2001.969535.

[34] Hoshuyama, O., Sugiyama, A. and Hirano, A. A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters. *IEEE Transactions on Signal Processing* 47.10 (1999), pp. 2677–2684. DOI: 10.1109/78.790650.

[35] Yinman Lee and Wen-Rong Wu. A robust adaptive generalized sidelobe canceller with decision feedback. *IEEE Transactions on Antennas and Propagation* 53.11 (2005), pp. 3822–3832. DOI: 10.1109/TAP.2005.858851.

[36] Vaseghi, S. V. *Advanced Digital Signal Processing and Noise Reduction*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006. ISBN: 047009494X.

[37] Diniz, P. *Adaptive Filtering: Algorithms and Practical Implementation*. Kluwer international series in engineering and computer science. Springer US, 2008. ISBN: 9780387686066.

[38] Haykin, S. *Adaptive Filter Theory*. Pearson, 2014. ISBN: 9780132671453.

[39] As'ad, H., Bouchard, M. and Kamkar-Parsi, H. A Robust Target Linearly Constrained Minimum Variance Beamformer With Spatial Cues Preservation for Binaural Hearing Aids. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 27.10 (2019), pp. 1549–1563. DOI: 10.1109/TASLP.2019.2924321.

[40] Spriet, A., Moonen, M. and Wouters, J. Stochastic gradient implementation of spatially preprocessed multi-channel Wiener filtering for noise reduction in hearing aids. *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 4. 2004, pp. iv–iv. DOI: 10.1109/ICASSP.2004.1326762.

[41] Cornelis, B., Moonen, M. and Wouters, J. Performance Analysis of Multichannel Wiener Filter-Based Noise Reduction in Hearing Aids Under Second Order Statistics Estimation Errors. *IEEE Transactions on Audio, Speech, and Language Processing* 19.5 (2011), pp. 1368–1381. DOI: 10.1109/TASL.2010.2090519.

[42] van Waterschoot, T. and Moonen, M. Fifty Years of Acoustic Feedback Control: State of the Art and Future Challenges. *Proceedings of the IEEE* 99.2 (2011), pp. 288–327. DOI: 10.1109/JPROC.2010.2090998.

[43] Guo, M., Jensen, S. H. and Jensen, J. An improved probe noise approach for acoustic feedback cancellation. *2012 IEEE 7th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. 2012, pp. 497–500. DOI: 10.1109/SAM.2012.6250548.

[44] Schepker, H., Tran, L. T. T., Nordholm, S. and Doclo, S. Improving adaptive feedback cancellation in hearing aids using an affine combination of filters. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2016, pp. 231–235. DOI: 10.1109/ICASSP.2016.7471671.

[45] J. Benesty and Y. Huang, eds. *Adaptive Signal Processing: Applications to Real-World Problems*. Springer Berlin Heidelberg, 2003. ISBN: 978-3-540-00051-8.

[46] Ozeki, K. *Theory of Affine Projection Algorithms for Adaptive Filtering*. Mathematics for Industry. Springer Japan, 2015. ISBN: 9784431557388.

[47] Hayes, M. H. *Statistical Digital Signal Processing and Modeling*. USA: John Wiley & Sons, Inc., 1996. ISBN: 0471594318.

[48] Dogancay, K. *Partial-Update Adaptive Signal Processing: Design Analysis and Implementation*. USA: Academic Press, Inc., 2008. ISBN: 0123741963.

[49] Sayed, A. *Fundamentals of Adaptive Filtering*. Wiley - IEEE. Wiley, 2003. ISBN: 9780471461265.

[50] Xiaoya Wu, Mingjiang Wang and Xiaoyong Zhang. Research on adaptive feedback cancellation in digital hearing aid. *2016 2nd IEEE International Conference*

on Computer and Communications (ICCC). 2016, pp. 1038–1042. DOI: 10.1109/CompComm.2016.7924862.

[51] Ozeki, K. and Umeda, T. An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties. *Electronics and Communications in Japan Part I-communications* 67 (1984), pp. 19–27.

[52] Sankaran, S. G. and Beex, A. A. L. Convergence behavior of affine projection algorithms. *IEEE Transactions on Signal Processing* 48.4 (2000), pp. 1086–1096. DOI: 10.1109/78.827542.

[53] Paleologu, C., Benesty, J. and Ciochina, S. Regularization of the Affine Projection Algorithm. *IEEE Transactions on Circuits and Systems II: Express Briefs* 58.6 (2011), pp. 366–370. DOI: 10.1109/TCSII.2011.2158162.

[54] Fa-Long Luo, Jun Yang, Pavlovic, C. and Nehorai, A. Adaptive null-forming scheme in digital hearing aids. *IEEE Transactions on Signal Processing* 50.7 (2002), pp. 1583–1590. DOI: 10.1109/TSP.2002.1011199.

[55] Samtani, K., Thomas, J., Deepu, S. P. and David, S. S. Area and power optimised ASIC implementation of adaptive beamformer for hearing aids. *2017 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. 2017, pp. 1–4. DOI: 10.1109/BIOCAS.2017.8325119.

[56] Samtani, K., Thomas, J., Varma, G. A., Sumam David S. and Deepu S. P. FPGA implementation of adaptive beamforming in hearing aids. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2017, pp. 2239–2242. DOI: 10.1109/EMBC.2017.8037300.

[57] Booth, A. D. A SIGNED BINARY MULTIPLICATION TECHNIQUE. *The Quarterly Journal of Mechanics and Applied Mathematics* 4.2 (Jan. 1951), pp. 236–240. ISSN: 0033-5614. DOI: 10.1093/qjmam/4.2.236.

[58] *Semi-conductor laboratory*. http://www.scl.gov.in.

[59] Kim, C. H. .-., Soeleman, H. and Roy, K. Ultra-low-power DLMS adaptive filter for hearing aid applications. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 11.6 (2003), pp. 1058–1067. DOI: 10.1109/TVLSI.2003.819573.

[60] Soeleman, H. and Roy, K. Ultra-Low Power Digital Subthreshold Logic Circuits. *Proceedings of the 1999 International Symposium on Low Power Electronics and Design*. ISLPED '99. 1999, pp. 94–96. ISBN: 158113133X. DOI: 10.1145/313817.313874.

[61] Harada, A., Nishikawa, K. and Kiya, H. Pipelined Architecture of the LMS Adaptive Digital Filter with the Minimum Output Latency(Special Section on Digital Signal Processing). *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* 81 (1998), pp. 1578–1585.

[62] Meyer, M. D. and Agrawal, D. P. A high sampling rate delayed LMS filter architecture. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 40.11 (1993), pp. 727–729. DOI: 10.1109/82.251841.

[63] Hyung-Il Kim and Roy, K. Ultra-low power DLMS adaptive filter for hearing aid applications. *ISLPED'01: Proceedings of the 2001 International Symposium on Low Power Electronics and Design (IEEE Cat. No.01TH8581)*. 2001, pp. 352–357. DOI: 10.1109/LPE.2001.945431.

[64] Gerlach, L., Payá-Vayá, G. and Blume, H. A Survey on Application Specific Processor Architectures for Digital Hearing Aids. *Journal of Signal Processing Systems* (2021). ISSN: 1939-8115. DOI: 10.1007/s11265-021-01648-0.

[65] Chang, K., Chen, Y., Kuo, Y. and Liu, C. A low power hearing aid computing platform using lightweight processing elements. *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2012, pp. 2785–2788. DOI: 10.1109/ISCAS.2012.6271888.

[66] Shaer, L., Nahlus, I., Merhi, J., Kayssi, A. and Chehab, A. Low-power digital signal processor design for a hearing aid. *2013 4th Annual International Conference on Energy Aware Computing Systems and Applications (ICEAC)*. 2013, pp. 40–44. DOI: 10.1109/ICEAC.2013.6737634.

[67] Serra-Graells, F., Gomez, L. and Huertas, J. L. A true-1-V 300 μW CMOS-subthreshold log-domain hearing-aid-on-chip. *IEEE Journal of Solid-State Circuits* 39.8 (2004), pp. 1271–1281. DOI: 10.1109/JSSC.2004.831469.

[68] Wang, X., Yang, H., Li, F., Yin, T., Huang, G. and Liu, F. A programmable analog hearing aid system-on-chip with frequency compensation. *Analog Integrated Circuits and Signal Processing* 79 (2014), pp. 227–236.

[69] Gata, D. G., Sjursen, W., Hochschild, J. R., Fattaruso, J. W., Fang, L., Iannelli, G. R., Jiang, Z., Branch, C. M., Holmes, J. A., Skorcz, M. L., Petilli, E. M., Chen, S., Wakeman, G., Preves, D. A. and Severin, W. A. A 1.1-V 270 μA mixed-signal hearing aid chip. *IEEE Journal of Solid-State Circuits* 37.12 (2002), pp. 1670–1678. DOI: 10.1109/JSSC.2002.804328.

[70] Kim, S., Cho, N., Song, S. and Yoo, H. A 0.9 V 96 μW Fully Operational Digital Hearing Aid Chip. *IEEE Journal of Solid-State Circuits* 42.11 (2007), pp. 2432–2440. DOI: 10.1109/JSSC.2007.907198.

[71] Hirsch, H. G. and Ehrlicher, C. Noise estimation techniques for robust speech recognition. *1995 International Conference on Acoustics, Speech, and Signal Processing*. Vol. 1. 1995, 153–156 vol.1. DOI: 10.1109/ICASSP.1995.479387.

[72] Scalart, P. and Filho, J. V. Speech enhancement based on a priori signal to noise estimation. *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*. Vol. 2. 1996, 629–632 vol. 2. DOI: 10.1109/ICASSP.1996.543199.

[73] Kates, J. M. Principles of Digital Dynamic-Range Compression. *Trends in Amplification* 9.2 (2005), pp. 45–76. DOI: 10.1177/108471380500900202.

[74] Qiao, P., Corporaal, H. and Lindwer, M. A 0.964mW digital hearing aid system. *2011 Design, Automation Test in Europe*. 2011, pp. 1–4. DOI: 10.1109/DATE.2011.5763297.

[75] Gerlach, L., Payá-Vayá, G. and Blume, H. KAVUAKA: A Low Power Application Specific Hearing Aid Processor. *2019 IFIP/IEEE 27th International Conference on Very Large Scale Integration (VLSI-SoC)*. 2019, pp. 99–104. DOI: 10.1109/VLSI-SoC.2019.8920354.

[76] Chen, C., Chen, L., Fan, J., Yu, Z., Yang, J., Hu, X., Hei, Y. and Zhang, F. A 1V, 1.1mW mixed-signal hearing aid SoC in 0.13μm CMOS process. *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2016, pp. 225–228. DOI: 10.1109/ISCAS.2016.7527211.

[77] Kim, S., Kim, M. and Kim, J. High-performance DSP platform for digital hearing aid SoC with flexible noise estimation. *IET Circuits, Devices and Systems* 13.5 (2019), pp. 717–722. ISSN: 1751-858X. DOI: 10.1049/iet-cds.2018.5374.

[78] Cheng-Wen, W., Cheng-Chun, T., Yi, F., Tian-Sheuan, C. and Shyh-Jye, J. Analysis and implementation of low-power perceptual multiband noise reduction for the hearing aids application. *IET Circuits, Devices & Systems* 8 (6 2014), 516–525(9). ISSN: 1751-858X. DOI: 10.1049/iet-cds.2013.0326.

[79] Bing-Fei Wu and Kun-Ching Wang. Robust endpoint detection algorithm based on the adaptive band-partitioning spectral entropy in adverse environments. *IEEE Transactions on Speech and Audio Processing* 13.5 (2005), pp. 762–775. DOI: 10.1109/TSA.2005.851909.

[80] Farhang-Boroujeny, B. and Wang, Z. Adaptive filtering in subbands: Design issues and experimental results for acoustic echo cancellation. *Signal Processing* 61.3 (1997), pp. 213–223. ISSN: 0165-1684. DOI: 10.1016/S0165-1684(97)00106-0.

[81] Loizou, P. *Speech Enhancement: Theory and Practice*. Signal processing and communications. Taylor & Francis, 2007. ISBN: 9780849350320.

[82] Yi Hu and Loizou, P. C. Subjective Comparison of Speech Enhancement Algorithms. *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*. Vol. 1. 2006, pp. I–I. DOI: 10.1109/ICASSP.2006.1659980.

[83] Chen, L.-M., Yu, Z.-H., Chen, C.-Y., Hu, X.-Y., Fan, J., Yang, J. and Hei, Y. A 1-V, 1.2-mA fully integrated SoC for digital hearing aids. *Microelectronics Journal* 46.1 (2015), pp. 12–19. ISSN: 0026-2692. DOI: 10.1016/j.mejo.2014.09.013.

[84] Chen, C. and Chen, L. A 79-dB SNR 1.1-mW Fully Integrated Hearing Aid SoC. *Circuits, Systems, and Signal Processing* 38 (July 2019). DOI: 10.1007/s00034-018-1002-6.