# Cost effectiveness filter design for low-latency audio analogue to digital converter (Σ-ΔADC)

**Xiangyu Zhu**

**A thesis submitted in partial fulfilment of the requirements of**

**Birmingham City University for the degree of Doctor of Philosophy**

**June 2021**

**Digital Media Technology Lab,**

**The Faculty of Computing, Engineering and the Built Environment,**

**Birmingham City University**

**Abstract**

The current technical challenge posed in professional audio industry is to have a small size portable device, which can support real-time interactive applications. $\Sigma$-$\Delta$ Modulation based audio system becomes the mainstream due to the higher resolution and fewer auxiliary circuits. However, extremely high sampling frequency brings severe challenges to its decimation or interpolation filter design and performance. The current optimal filter design parameter calculation methods have room for improvement, such as complex calculations and results need further rounding. Therefore, Author proposed a new optimal decimation or interpolation rate selection approach converts the optimization problem to factorization and permutation problem which improves the efficiency significantly and provides directly usable integer solutions. Furthermore, due to the lack of theoretical relationship between latency of filter and multi-stage design parameters a latency estimation equation is derived by author. The analysis of this equation shows that optimal computational cost design and optimal latency design have contradictory requirements. Hence, the optimization of filter design parameter is added to optimize the other costs as much as possible. Thus, author proposed a new numeric optimization based method to design the cost efficient low-latency multi-stage multi-rate filter. This approach further reduces the costs of the filter on the basis of the previous optimal design. At last, a MATLAB GUI based filter design and evaluation framework has been established which can help user to search the optimal design parameters and design optimal filters with different filter types.

**Acknowledgements**

I'd like to thank everyone in the Digital Media Technology (DMT) Lab at Birmingham City University and Prof. We Hu's team in Wuhan University of Science and Technology. In particular my supervisor Prof. Cham, Dr. Yonghao Wang and Prof. We Hu who supported and encouraged me alot during my PhD research. Also, a big thank you to all my dear friends for their help and support not only in research but also mentally and daily life. Finally, to my family, thank you. None of these would have happened without you.

## Acronyms

| | |
|---|---|
| PDE | Partial Differential Equations |
| IC | Integrated Circuit |
| VLSI | Very Large Scale Integration |
| DSP | Digital Signal Processing (Processor) |
| LTI | Linear Time-Invariant |
| ADC | Analogue to Digital Converter |
| DAC | Digital to Analogue Converter |
| FIR | Finite Impulse Response |
| IIR | Infinite Impulse Response |
| CIC | Cascaded Integrator-Comb |
| PCM | Pulse-code Modulation |
| SDM | Sigma-Delta Modulation |
| FFT | Fast Fourier Transform |
| DTFT | Discrete-time Fourier Transform |
| DFT | Discrete Fourier Transform |
| SRC | Sampling Rate Conversion (Changing) |
| SNR | Signal-Noise Ratio |
| SA | Simulated Annealing |
| GUI | Graphic User Interface |

## Mathematical Notation

| | |
|---|---|
| $\mathbb{N}_1$ | Set of Natural Numbers |
| $\sum_{i=1}^{K} D_i$ | Accumulation from $D_1$ to $D_K$ |
| $\prod_{i=1}^{K} D_i$ | Multiplication from $D_1$ to $D_K$ |
| $\binom{n}{k}$ or $C_k^n$ | Binomial coefficient or all combinations |
| $\Delta f$ | Transition bandwidth |
| $\delta_1$ | Passband ripple |
| $\delta_2$ | Stopband attenuation |
| $f_p$ | Passband edge |
| $f_s$ | Stopband edge |
| $F_s$ | Sampling frequency |
| $D$ | Oversampling rate |
| $K$ | Number of stage |
| $\nabla$ | Partial derivative of vector |

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Motivation

At present, the development of science and technology has a great impact on human entertainment activities. The digital technology especially has a great influence on the professional audio industry. The interesting technical challenge posed in professional audio industry is to have a small size portable device, which can support real-time interactive (with low latency) applications.

Nowadays, more and more entertainment functions are embedded in the portable devices, such as using smartphones and tablet devices to play, control and interact with audio data. Due to the limitation of the physical size and power supply, the hardware efficiency has become more and more important to the system design, in other words, the design needs to achieve the maximum effect with minimum resources. This puts a challenge on both analogue and digital design parts of the audio system. For the digital part, the hardware efficiency normally means the architecture with fewer digital logical gates to perform arithmetic operations.

For larger audio performance shows in live, the professional audio processing system chain is used in real-time for live mixing and monitoring, audio processing, live broadcasting, and live recording. It is well known that the distinguishable latency for human ear is about 30ms (Haas 1972). According to Lester and Boley (2007), the audio latency in the scale of a few milliseconds can be perceived by human ear, especially in the professional studio or live concert. In the article Lester and Boley (2007), the perceptions of musicians who play different musical instruments are tested by using different

latencies for live sound monitoring devices. For some musicians, the very low latency in-ear monitoring system is preferred in which the latency is less than 2ms, for example: singers and saxophonists.

The current professional audio processing systems can be separated into analog based system and digital based system. The analogue based system has a strong ability of audio signal restoration and an extremely low group delay. However, the analog based system normally has a large volume and power cost due to the application of a large number of analog components. Kester (2015) stated that with the development of technology, the $\Sigma$-$\Delta$ Modulation, which uses high sampling frequency to trade for high sample resolution is developed and the capabilities of hardware are improved. This means the digital system could have enough resolution to achieve high level audio processing. At the same time, the digital system has the advantages of low power cost and portability over to analogue system. Therefore, the current trend is to use digital system to process high resolution audio signal.

Although the volume and power cost of digital system has been greatly improved compared to analog system, it still cannot fully meet the increasing stringent requirements for portability. Also, the current digital devices are battery powered. Therefore, efficient digital systems have become the trend of design. Hence, the computational cost which cause the most power cost become an important indicator to measure a digital system design.

According to Wang, Stables, and Reiss (2010), Reiss (2008) and above introduction presented, it can be known that the challenge basically raises two technical aspects in terms of digital domain:

1. Hardware efficient digital filter design for audio signal processing.
2. Real-time DSP algorithm and architecture with low latency.

Therefore, it is interesting to concentrate on the technical challenges of design hardware efficient low latency $\Sigma$-$\Delta$ based ADC.

## 1.2 Aim and Objectives

The aim of this research is to carry out a comprehensive evaluation as well as propose a framework for the digital filter design with both low latency and

low cost for Σ-Δ modulation based ADC/DAC.

In order to achieve this goal, several objectives are set up.

1 Investigate into the optimal decimation or interpolation rate calculation methods and improve the current approach.

2 Explore the relationship between latency and filter design parameters; and derive the latency estimation equation.

3 Explore the optimal filter design parameters finding method, and propose the optimal filter design approach based on optimal design parameters.

4 Develop cost efficient low-latency filter design optimization method and evaluation framework using numeric optimization algorithm.

## 1.3 Thesis Structure

Based on the aim and objectives, the sketched research structure can be constructed as:

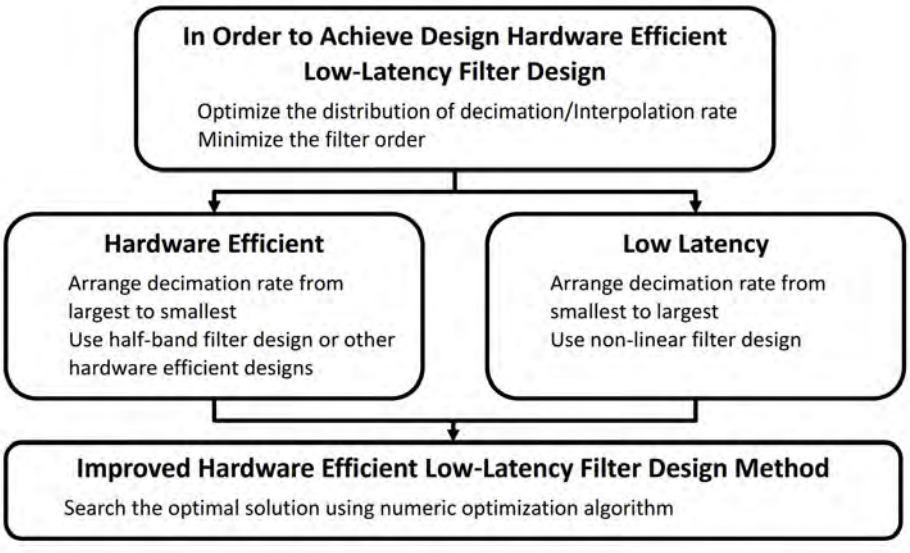



Figure 1.1: Structure of Research

As Figure 1.1 shows, in order to achieve the aim of research, the costs and latency of decimation or interpolation filters need to be minimized. According

to Crochiere and Rabiner (1975), the distribution of decimation or interpolation rate affects cost and latency of the multi-stage multi-rate filter design significantly. Therefore, it is important to find out the optimal decimation or interpolation rate distribution for the design. Also, the improved of filter design will reduce the costs and latency as well.

However, cost efficient and low latency designs have contradictory design requirements. Hence, using suitable types of filters and filter design parameter optimization become the key to achieving the aim. Nevertheless, there are many design parameters need to be optimized, and each parameter's change affects the filter performance. Therefore, these parameters have to be adjusted together. Base on these characteristics, the numeric optimization algorithm is used to search the optimal solution.

Following the above research structure, the thesis are mainly constituted by six parts.

**Chapter 2 Literature Review** presents the background upon which this thesis will be developed. It begins with a discussion of the digital audio system. Different popular modulation techniques and filter requirements are introduced along with different systems. Finally, optimal decimation or interpolation rate calculation methods and optimal filter design techniques have been discussed.

**Chapter 3 Methodology** discusses the methodology of research work based on the 1.2 Objectives.

**Chapter 4 Cost Efficient Decimation Rate Selection and Filter Design** investigates into the optimal decimation or interpolation rate calculation methods and the optimal multi-stage multi-rate filter design techniques for different aspects. A highly efficient decimation or interpolation rate selection method is proposed. Lastly, based on this selection method, the look-up to 3D table optimal filter design method has been introduced.

**Chapter 5 Research and Derivation for Theoretical Latency of Multi-stage Multi-rate Filter** investigates into the relationship between filter's latency and filter design parameters. The derived latency estimation function for the optimal narrow band linear phase FIR filter design fills the gap. From analysis of this estimation function, the relationship between latency and filter design parameters is discovered. Also the relationship between latency and different filter costs has been discussed as well.

**Chapter 6 Low Latency Cost Efficient Filter Design Using Optimization Method** integrates all previous findings into a realistic model based low latency cost efficient filter design framework. Meanwhile, the widening transition bandwidth and overshoot problem, the low cost filter design using half-band structure and passband ripple allocation problem have been discussed.

**Chapter 7 Conclusion and Future Work** concludes the thesis. Research problem and findings are discussed and the suggestions for future work were considered.

## 1.4 Contributions

The main contribution of this thesis is the proposed filter design evaluation and development framework for Σ-Δ Modulation based high level audio ADC. In achieving this, a number of other contributions are made as follows:

- The first contribution is we proposed a new method for fast identifying the optimal decimation or interpolation factors for area cost and computational cost efficient multi-stage multi-rate filter design. This new method improved the efficiency of optimal decimation/interpolation rate selection significantly.

  Author did this based on the limitations of current decimation or interpolation rate selection approaches. From 1975 to 2007, Crochiere and Rabiner and Coffey have proposed different approaches. However, they all have the complicated calculation or search process and cannot provide integer solutions. Although Huang (2003,2009) proposed exhaustive and genetic algorithm based search methods to obtain the integer solutions using set theory, this method is not ideal in efficiency.

- Second contribution is to solve the optimal latency filter design problem. To construct the relationship between filter's latency and design parameters.

  The second contribution we made is because currently, there is no theoretical relationship between filter's latency and multi-stage design parameters. Therefore, my supervisor Leo Wang and I derived the latency

estimation equation inspired by Crochiere and Rabiner 1975 (1975)'s assumption and filter order estimation equations. From this equation, author discovered that the key factors affect the latency performance are transition bandwidth, passband ripple and stopband attenuation. These findings guide us to design the optimal low cost and low latency filters which leads to my 3rd contribution.

- At last, based on above contributions, author proposed a numeric optimization based cost efficient low-latency filter design method to improve the performance of filter cost and latency at the same time.

  The proposed method uses the annealing algorithm to optimize the key factors affect the filter performances, so that the designed filter can be further improved by about 3%-4% on the basis of previous optimal design. To the best of our knowledge, there is no theoretical framework to optimize multi-stage multi-rate filter using the properties we discovered in previous research. Although in commontial package such as MATLAB, they do use the optimization method which haven't been published. However, author proposed method shows better results than the existing software as well.

## 1.5 Publications

- **Most of research in Chapter 4** was published as: Zhu, Xiangyu et al. (2016) "*Practical Considerations on Optimising Multistage Decimation and Interpolation Processes*". In: *Digital Signal Processing (DSP), 2016 IEEE International Conference on.* IEEE, pp. 370-374.

  The author of the thesis wrote and did the factorisation algorithm, database creation, optimal solution set selection algorithm, as well as the implementation, analysis, and verification of the program. Yonghao Wang has discovered the regularity of optimal solution sets, and proposed the simplified search and balanced design algorithms. Other authors had and editing and supervising role. The program and data can be browsed from Zhu and Wang (2016).

- **The preliminary work of Chapter 6** was published as: Ma, R. et al. (2018) "*Optimum Design of Multistage Half-band FIR Filter for Audio Conversion Using a Simulated Annealing Algorithm*". In: *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 74-78.

  The author of the thesis proposed the design methods and filter design and measure algorithm. Yonghao Wang proposed the idea of using wider transition to reduce the filter cost. Rongxuan Ma proposed using simulated annealing algorithm to find out the optimal solution. Other authors had an editing and supervising role.

# Chapter 2

# Literature Review

## 2.1 Digital Audio System

According to Trick and T. (1975) and Leslie and Singh (1990), most of the real-time digital audio processing systems have the following common audio processing chain structures.

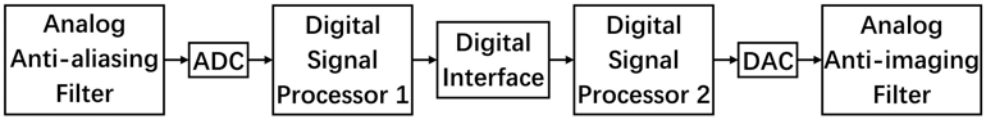



Figure 2.1: Audio Processing Chain

Lapsley et al. (1997) mentioned that there are four main blocks in the audio processing chain: analog to digital converter (ADC) module, digital signal processor (DSP) module, digital interface module and digital to analog converter (DAC) module. For audio signal processing, an anti-aliasing filter can be added into the chain if the system does require one. And with the same argument, the anti-imaging filter can be added into the chain as well.

Wang (2011) mentioned the commonly used ADC/DAC technologies in the audio processing chain normally have latency about 1ms, and this 1ms delay is significant for a low-latency audio system in some situations. Wang and Reiss (2012) have pointed out the main delay of ADC/DAC is the group delay which is generated by the structure of $\Sigma$-$\Delta$ Modulation (SDM) based ADC/DAC.

### 2.1.1 Traditional PCM Based Digital Audio System

According to Mitra and Kaiser (1993) and Proakis and Manolakis (1996), in the Pulse-Code Modulation(PCM) based digital audio system, the amplitude of analogue signal is sampled regularly at uniform intervals, and each sample is quantized to the nearest value within a range of digital steps like the following example shows.

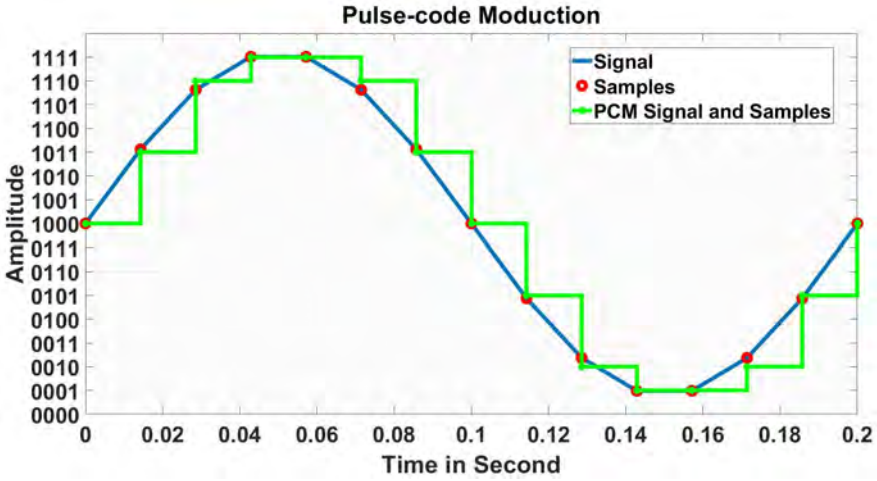



Figure 2.2: Pulse-code Modulation

As the Figure 2.2 demonstrated, the blue curve is input signal, red dots are samples and the green step signal is the PCM signal. Y axis is the amplitude of the signal, where the amplitude is represented by the binary expression of the voltage. Base on this example and theory of PCM, the lack of using PCM in high-level audio signal processing area can be found.

- **Resolution (Bit width)**
  Both Lyons (2011) Leung et al. (1988) mentioned that the voltage difference between two neighbouring voltage level is very small, because the power supply can not be very large. As an example, for a 16-bit ADC/DAC with ±5V power supply, the range between two digital steps is only $10V/2^{16}$ which is around 0.00015V Therefore, it is hard to separate the voltage into too many levels due to the stability and accuracy problem. According to the datasheet of Texas Instruments PCM56, normally, regulated power supplies with 1% or less ripple are recommended for use with the DAC. Hence, the highest resolution for PCM

based system is around 16-bit which is far behind the Σ-Δ Modulation based system.

- **Quantization error (Rounding error)**
  Also Vegte (2001) and Stein (2001) have mentioned that the samples are sampled with the fixed space in time domain, and the amplitude of samples are rounded to the nearby voltage level as the Figure 2.2 demonstrates. Even though, the difference between the sample's amplitude and nearby voltage level is tinny.

- **Analogue anti-aliasing and anti-imaging filter and other support analog circuit are needed**

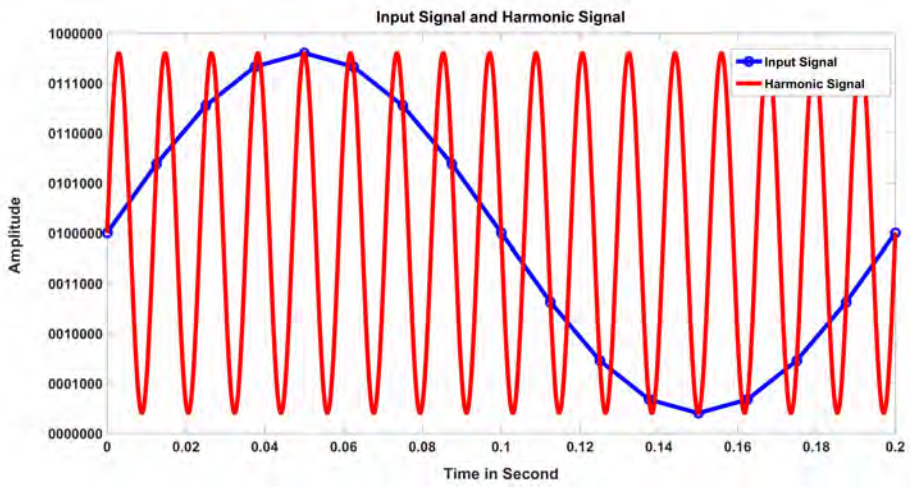



Figure 2.3: Input Signal and Harmonic Signal

According to Oppenheim, Schafer, and Buck (1999), Madisetti and Williams (1998) and Mitra (2011), as Figure 2.3 shows, the input signal is normally sampled with a certain sampling rate. Mintzer and Liu (1977) mentioned these samples can be reconstructed with different combinations of sine waves which means the non-existent high frequencies (Red harmonic signal and its harmonic signals) are added to the signal. Hence, the anti-aliasing and anti-imaging filters are needed to filter out these extra high frequencies.

Based on the Nyquist Sampling Theory, in order to abstract all information from the signal, sampling frequency of the system has to be

more than two times of signal's frequency. However, the sampling frequency can not be too large due to the efficiency problem and aliasing or imaging problem. Therefore, the system's sampling frequency is usually chosen according to the limitation of Nyquist Sampling Frequency to avoid these problems.

Meanwhile, Schreier, Temes, and Norsworthy (1997) and Friedman et al. (1989) have pointed out that currently, even though lots of digital audio systems use PCM format signal as the standard, but due to the stability of voltage cannot be maintained when the resolution is higher than 16-bit, the sampling process has been replaced by Σ-Δ Modulation technique.

### 2.1.2 Σ-Δ Modulation Based Digital Audio System

Candy (1986) and Stojnic and Babic (2011) stated the concept of Σ-Δ Modulation has been proposed for decades, but because of the limitation of Integrated Circuit (IC) in the past, Σ-Δ Modulation technique had been stuck in a concept for long time until the rapid development of Very Large Scale Integration (VLSI) in recent years. Now, according to Ansari and Liu (1985) Hejn, Kale, and Kurek (1993), Σ-Δ Modulation technique has been widely used in digital audio processing area because of its high resolution. It also reduces the design complexity of external analog anti-aliasing filter and anti-imaging filter via the internal filters. But Williams and Taylor (2006) also state that the Σ-Δ Modulation technique has the shortcoming of the latency of the internal filters.

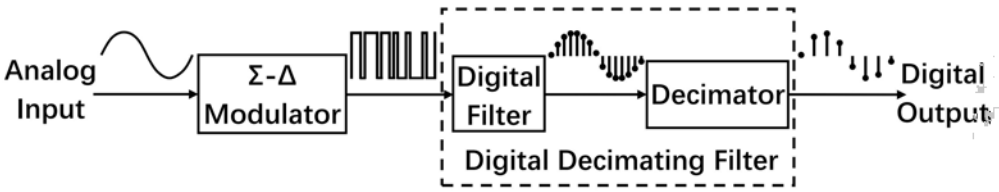



Figure 2.4: Block Diagram for Σ-Δ ADC

Figure 2.4 shows a typical SDM based ADC, Aziz, Sorensen, and Spiegel (1996)it can be regarded as two parts: one is the SDM modulator and the other one is digital filter (decimation filter in this case). According to Gray (1987) and Boser and Wooley (1988), for Σ-Δ Modulator Block, there are integrator, comparator and 1bit digital to analog converter in a negative

feedback loop. Furthermore, Candy (1985) and Chae and Han (2009) added more integrator or feedback loop of the Σ-Δ modulator to improve the resolution and accuracy of the system. However, comparing with the costs and latency caused by digital decimator or interpolator parts, these modifications are insignificant.

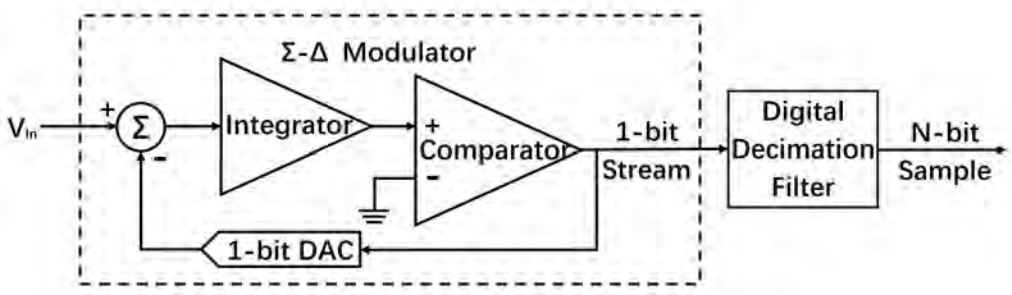



Figure 2.5: First-Order Σ-Δ ADC

Lipshitz and Vanderkooy (2001) stated that following the basic structure of the Σ-Δ modulator, like Figure 2.5 shows, the signal will be imported to the integrator after adding the input signal and the negated DAC output together; the integrator will export a ramp signal, the slope of this signal is proportional to the amplitude of the input signal of the integrator; and the output signal of integrator will be compared with the reference signal in the comparator and export 1-bit output. The binary output signal of comparator will be sent to the digital decimation filter based on the oversampling of ADC which is the rate of oversampling. Each bit of the comparator's output signal represents the direction of the output signal, after cycling for several times, the N-bit sample representation is exported from the digital decimation filter, and it represents the quantized values of the output signal. In fact, the input signal is matched with the average output of the DAC by the feedback loop, and the digital decimation filter will counterpoise the bit stream and export the N-bit samples with accepted sampling frequency.

Wherein the Σ-Δ Modulator block does not cause the latency, Candy and Benjamin (1981) mentioned the significant latency is caused by the internal filter which is designed as high performance, high resolution filter. Wang and Reiss (2012) have proved that the low-latency filter for Σ-Δ ADC can be realized mathematically, but there is no detailed literature showing if these filters can be realized in hardware concisely.

According to Kester (2000), the Σ-Δ ADC has the following characteristics:

- Low Cost, High Resolution (to 24-bits)
- Excellent Differential Linearity
- Low Power, but Limited Bandwidth (Voiceband, Audio)
- Key Concepts are Simple, but Math is Complex
  - Oversampling
  - Quantization Noise Shaping
  - Digital Filtering
  - Decimation
- Ideal for Sensor Signal Conditioning
  - High Resolution
  - Self, System, and Auto Calibration Modes
- Wide Applications in Voiceband and Audio Signal Processing

Hence, the $\Sigma$-$\Delta$ Modulation technique is suitable for high-level audio signal processing area. According to the view of Kester (2000), at present, there are several kinds of ADC/DAC architecture, for example, the Pulse-code Modulation based ADC/DAC, Successive Approximation Register (SAR) based ADC/DAC, $\Sigma$-$\Delta$ Modulation based ADC/DAC and so on. And each of these systems has different internal structures and performances. The $\Sigma$-$\Delta$ based ADC/DAC becomes de facto standard converters used in audio system nowadays due to its architectural advantages in supporting small and portable size of audio devices.

## 2.2 Filters

### 2.2.1 LTI Discrete-Time System

Oppenheim, Schafer, and Buck (1999) and Mitra (2011) have mentioned the linear system is the most widely used discrete-time system in digital signal processing. For high level audio signal processing area, the Linear Time-Invariant(LTI) System is the most common used system format. There are two important properties for LTI system, linearity and the time invariance.

- **Time invariance**
  According to Tan and Jiang (2018), the time invariance system can be described as the follows: the output signal will be identical delay for $N$ seconds while the input signal import to the system $N$ seconds from now. This means there is a time-invariance system:

$$y(n) = H\{x(n)\} \tag{2.1}$$

  The system must meet the requirement of:

$$y(n-N) = H\{x(n-N)\} \tag{2.2}$$

  Following the above description of Time invariance, the system can simply be viewed as following figure 2.6 shows. No matter how long the signal delayed, for a certain input signal, the system will export the same output with added delay. In another word, the output signal is only related to the input signal.

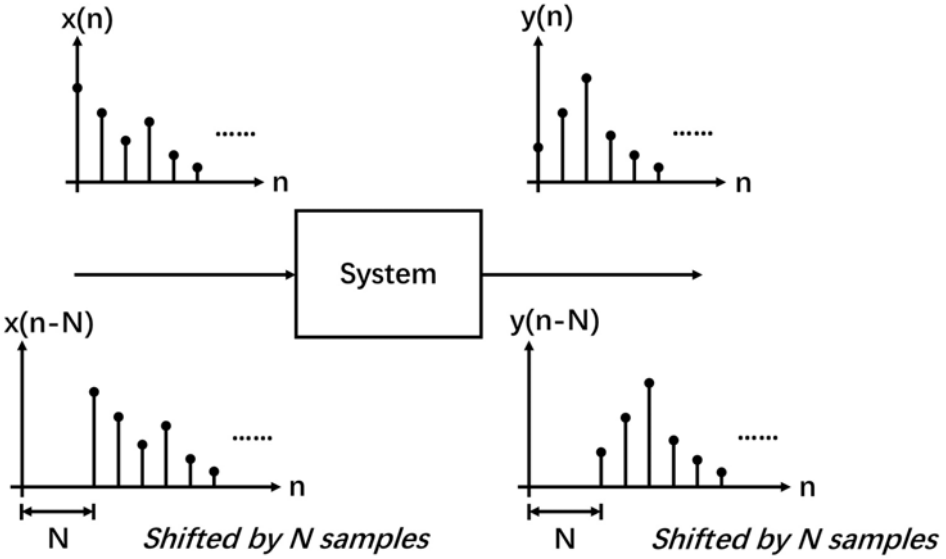



Figure 2.6: Sketched Diagram of Time Invariance

This property is the basic feature of Σ-Δ Modulation based decimation/interpolation filter. Otherwise it will be very hard to reconstructed the filtered digital signal to desired analog signal due to the correctness of the input signal cannot be guaranteed.

- **Linearity**
  Linearity describes the relationship between system's input and output is a linear map. For example, A system $H$ which has input signal $x_1(n)$ produces the output signal $y_1(n)$, and the input signal $x_2(n)$ produces the other output signal $y_2(n)$. If the system meets the condition of input signal $a_1x_1(n) + a_2x_2(n)$ producing the output signal $a_1y_1(n) + a_2y_2(n)$ it can be called as a linear system. According to Tan and Jiang (2018), this system can be sketched as:

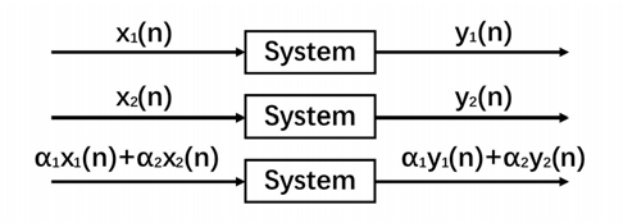



Figure 2.7: Sketched Diagram of Linear System

The system can be written as:

$$a_1y_1(n) + a_2y_2(n) = H\{a_1x_1(n) + a_2x_2(n)\} \tag{2.3}$$

Then, this can be extended to an arbitrary number of terms, and so for real numbers $a_1, a_2, \ldots, a_k$

$$\sum_k a_ky_k(n) = H\{\sum_k a_kx_k(n)\} \tag{2.4}$$

Linearity is important for decimator or interpolator in Σ-Δ modulation based ADC/DAC due to the accuracy of the reconstructed signal. However, Wang and Reiss (2012) pointed out that using non-linear filter can reduce the system's latency significantly. In exchange, higher frequency signals will have larger groupdelay. This may cause discomfort in high level audio area.

As Mitra (2011) mentioned, there are two fundamental types of digital filters, the finite impulse response (FIR), and the infinite impulse response (IIR). Although it is not accurate enough, when referring to FIR filters, they usually refer to linear FIR filters. With the same argument, IIR filters normally

refer to non-linear IIR filters.

### 2.2.2 Infinite Impulse Response (IIR) Filter

As shown by the name of infinite impulse response filter, the impulse response of this type of filter extends for an infinite period of time due to the recursive structure or feedback loop. An example of IIR filter has been given in Figure 2.8. Kester (2000) has mentioned that although IIR filter can be implemented with fewer computations than FIR filters, IIR filters do not match the performance achievable with FIR filters, and there is no computational advantage when the output of an IIR filter is decimated because each output value must always be calculated. And there is one of the most important problems for using IIR filter in the audio signal processing area is IIR filter do not have linear phase.

Kester (2000) and Ifeachor and Smithson (1995) have also pointed out that the Infinite Impulse Response Filter is a highly efficient filter type. Meanwhile, IIR filter can achieve better magnitude-frequency characteristics with less hardware resources. But there are prominent characteristics for IIR filter which are non-linear phase and the performance not able to compete against FIR filter. Therefore, although IIR has the theoretical advantages comparing with FIR filter, IIR has disadvantages in applications.

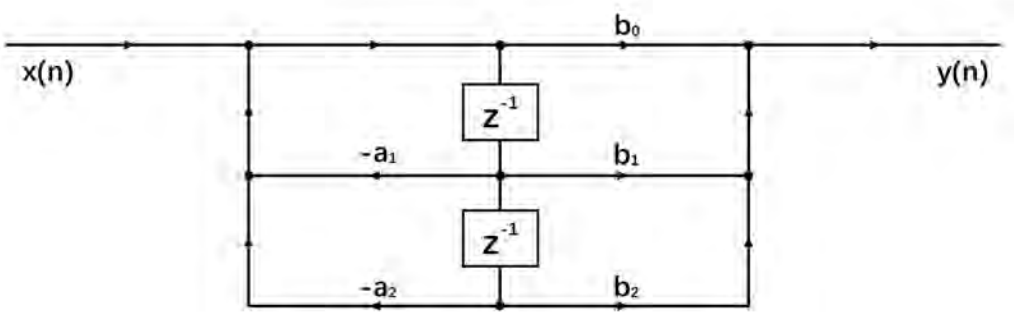



Figure 2.8: Second-Order IIR Filter Simplified Notation

IIR filter's impulse response is infinite, and its transfer function can be written as:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{i=0}^{M} b_i z^{-1}}{1 - \sum_{j=1}^{N} a_j z^{-1}} \tag{2.5}$$

Where:
$M$ is feedforward filter order
$b_i$ are feedforward filter coefficients
$N$ is feedback filter order
$a_j$ are feedback filter coefficients

According to Kester (2000), IIR filter has following characteristics:

- Uses Feedback (Recursion)
- Impulse Response has an Infinite Duration
- Potentially Unstable
- Non-Linear Phase
- More Efficient than FIR Filters
- No Computational Advantage when Decimating Output
- Usually Designed to Duplicate Analog Filter Response
- Usually Implemented as Cascaded Second-Order Sections (Biquads)

Even though, the IIR filter has advantage of cost efficient and low-latency comparing with FIR filter, Johnson (1984) point out that the accuracy of calculation and storage is required due to the feedback loop, and the stability check is needed. McClellan, Parks, and Rabiner (1973) mentioned that the IIR filter need analogue support circuit which may increase the volume of system. Also the vary latency may cause the mismatch in high level audio area. Sozanski (2013) stated that although using linear-phase IIR filters as interpolator can reduce the filter order and operations significantly, linear-phase IIR filters are only suitable for applications where a linear-phase is required and long delay time is acceptable. Hence, for the $\Sigma$-$\Delta$ Modulation technique, IIR filter is not the best choice.

### 2.2.3 Finite Impulse Response (FIR) Filter

Mitra 2011 (2011) and Smith (1997; 2003) have mentioned that the Finite Impulse Response Filter is a kind of filter which has finite length of impulse response curve. Comparing with IIR filter, the most important advantage of FIR filter is its linear phase. But not all FIR filters have this characteristic.

Only symmetrical FIR filter has this characteristic.

The most elementary form of an FIR filter is given in Figure 2.9.

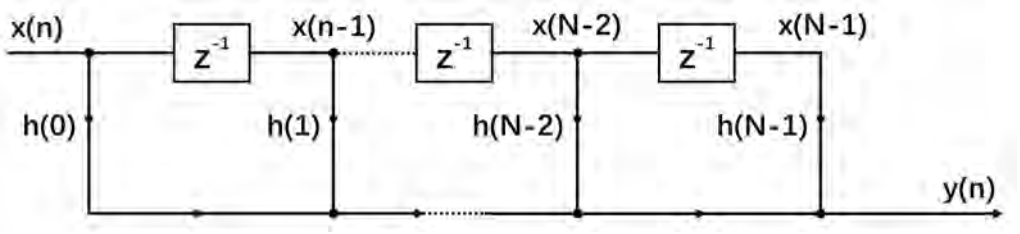



Figure 2.9: $N_{th}$ Order FIR Filter Simplified Notation

Based on the above Figure 2.9, the FIR filter's transfer function can be written as:

$$H(z) = \sum_{n=0}^{N-1} h(n) z^{-n} \tag{2.6}$$

According to Kester (2000), the characteristics of FIR filters can be summarized as:

- Impulse Response has a Finite Duration (N Cycles)
- Linear Phase, Constant Group Delay (N must be odd)
- No Analog Equivalent
- Unconditionally Stable
- Can be adaptive
- Computational Advantages when Decimating Output
- Easy to Understand and Design
  - Windowed-Sinc Method
  - Fourier Series Expansion with Windowing
  - Frequency Sampling by Using Inverse FFT-Arbitrary Frequency Response
  - Parks-McClellan Program with Remez Exchange Algorithm

### 2.2.4 Discussion

Kester (2000) summarized the following Table 2.1:

Table 2.1: Comparison Between FIR and IIR Filters

| **IIR Filters** | **FIR Filter** |
|---|---|
| More Efficient | Less Efficient |
| May Be Unstable | Always Stable |
| Non-Linear Phase Response | Linear Phase Response |
| No Efficiency Gained by Decimation | Decimation Increases Efficiency |

As mentioned above, Σ-Δ Modulation based system is widely used in audio area. Therefore high performance digital decimation or interpolation filters are needed. From section 2.2.2, it can be known that IIR filters are more efficient than FIR filters because the less memory usage and fewer multiplier are needed. However, IIR filters need extra analog support circuit and may exhibit instability problems, but this is much less likely to occur if higher order filters are designed by cascading second-order systems. On the other hand, FIR filters require more taps and multiplier for a given design. But linear phase characteristics and the steepness of transition band are the advantage of FIR filters. If the number of multiplication is not prohibitive, and the linear phase is a requirement, the FIR filter should be chosen which means for Σ-Δ Modulation based high-level audio ADC/DAC, FIR filters are better than IIR filters.

## 2.3 Filter Cost

With the rapid development of science and technology, high quality audio or video media are no longer a symbol of large professional equipments; rather, more and more portable devices have the capability to run or play high quality entertainment resources. Therefore, filter cost become one of the most important consideration for filter design.

### 2.3.1 Different Filter Cost

Filter cost is not only one aspect or a parameter, there are several different types of cost. For example, Computational Cost, Area Cost and so on. Each

of these cost describes one or several performances of filters. This project is mainly focused on the "Computational", "Area" and "Latency" cost which are the most important evaluation factors for the filter.

### Computational Cost

Oppenheim et al. (1976) pointed out that there are several different methods to evaluate the filter performance. One of the most common used method to evaluate efficiency of filter is the so called "Computational Cost" which can evaluate the computation complexity of the system. Normally, the computational cost can be expressed as A/IS (Addition per Input Sample) and M/IS (Multiplication per Input Sample). A/IS and M/IS are the addition or multiplication operation needed for each input sample during the filtering process.

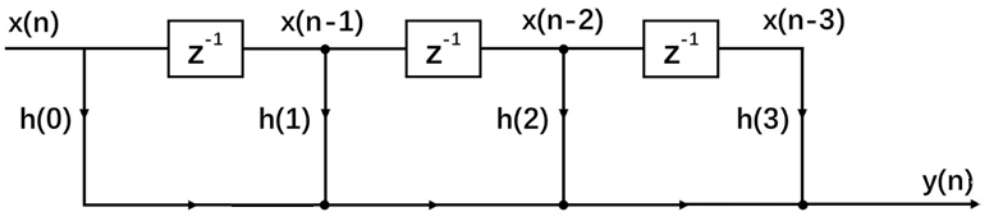



Figure 2.10: $3_{rd}$ Order FIR Filter Simplified Notation

As above Figure 2.10 shows, for this $3_{rd}$ order FIR filter, in order to obtain the current output signal $y(n)$, the current input signal $x(n)$ needs to multiply with filter coefficient $h(0)$, and then add the accumulates of previous signals and filter coefficients in sequence which means, the current input signal $x(n)$ needs to do one multiplication and three additions to obtain the current output signal $y(n)$.

### Area Cost

As Crochiere and Rabiner (1975; 1976 and 1981) mentioned, the difference between Computational Cost and Area Cost is that the "Computational Cost" represents the computation complexity and the "Area Cost" represents the hardware complexity. The "Area Cost" can be expressed as how many storage capacities are needed for the system.

As in the same example given in Figure 2.10, the filter coefficients $h(0)$ to $h(3)$, previous input signal $x(n-1)$ to $x(n-3)$ and also the accumulates of

filter coefficients and previous input signal $x(n) \times h(0)$ to $x(n-3) \times h(3)$ need to be stored for future operations. Memory usage and occupation can be used to present the "Area Cost".

**Power Cost**

As Brandt and Wooley (1994) mentioned, the high-level audio signal processing system pay more and more attention to portability. Therefore, most of these devices are battery driven. Hence, the battery life is an extremely important design indicator which means the system must achieve high performance with relatively low power cost.

According to Agarwal, Pavankumar, and Yokesh (2008) adders and multipliers are the most important arithmetic units in a general microprocessor and the major source of power dissipation. Yuan and Liu (2014) have pointed out that with the advancement of Very Large Scale Integrated Circuit (VLSI) design technology, high performance processing chips have become an indispensable part of communications, electronics and space technology. Therefore, modules containing a large number of decimal multiplications are also frequently applied to various chips and circuits such as digital filters and digital signal processors. For general multiplications, its power cost and area cost are mainly derived from the number of internal addition operations. Therefore, if the number of internal addition operations can be reduced while maintaining the accuracy of the multiplication operation, the power cost and area cost of the multiplier can be reduced effectively.

Therefore, the computational cost of the filter design is the key factor of power cost of filter.

### 2.3.2 Latency

The auditory perception of latency has important effects on many live audio applications. In many cases if the latency between two consequent expected events is beyond a certain threshold, it causes negative effects. Lester and Boley (2007), Farner et al. (2009), and Chafe, Caceres, and Gurevich (2010) have summarized the typical latency thresholds for different applications, for musical ensembles, the performance can be naturally synchronised when the latency is between 8ms to 25ms.

As above mentioned, Wang and Reiss (2012) have pointed out the main delay of ADC/DAC is the group delay which is caused by the filters. And

the system's latency is well known as the delay between input signal and output signal. With the same example given by Figure 2.10, the filter needs four samples to start the process, therefore, the filter starts to export the output signal after importing the four input samples. This delay is so called the group delay which is the main delay of ADC/DAC for Σ-Δ Modulation based system due to the extreme rigorous requirements.

For a digital filter which has N sample impulse response h(n), where n is the input serial number. If we perform Discrete-time Fourier Transform (DTFT) on this filter, the following transfer function can be obtained:

$$H(\omega) = M(\omega)e^{j\phi(\omega)} \tag{2.7}$$

Where $M(\omega)$ is the magnitude response of filter, $\Phi(\omega)$ is phase response of filter and $\omega$ is the angular frequency. Derivative $H(\omega)$ to $\omega$, it can be obtained:

$$\frac{d[H(\omega)]}{d\omega} = M(\omega)\frac{d[e^{j\phi(\omega)}]}{d\omega} + e^{j\phi(\omega)}\frac{d[M(\omega)]}{d\omega} \tag{2.8}$$

After derivative expansion of the first term on the right side of the equal sign and simplifying, the equation becomes:

$$\frac{jd[H(\omega)]/d\omega}{M(\omega)e^{j\phi(\omega)}} = -\frac{d[\phi(\omega)]}{d\omega} + j\frac{d[M(\omega)]/d\omega}{M(\omega)} \tag{2.9}$$

Because the $jd[H(\omega)]/d\omega$ is the DTFT of $n \cdot h(n)$, $M(\omega) \cdot e^{j\phi(\omega)}$ is $H(\omega)$ which is DTFT of $h(n)$ and $-d[\Phi(\omega)]/d\omega$ is the group delay of the filter. If we use DFT instead of DTFT, the above equation 2.9 can be rewritten as:

$$GD = real[\frac{DFT[n \cdot h(n)]}{DFT[h(n)]}] \tag{2.10}$$

where the GD is the filter's group delay. And from this equation 2.10 it can be found that the group delay calculation method is the process of deconvolution of the phase or it can be said that the group delay is the differentiation of phase versus frequency.

According to Section 2.2.2 and 2.2.3, and also the above Equation 2.10, it can be known that the linear FIR filter has constant group delay and IIR filter has the group delay that varies with frequency due to the different phase

characteristic. Although the non-linear phase filter has lower group delay comparing with linear phase FIR filters, the varies group delay makes the non-linear filter a great disadvantage for time-sensitive signals in industry control signal, live music signal and so on.

### 2.3.3 Discussion

Zambreno, Nguyen, and Choudhary (2004) mentioned that cost efficient system with high throughput has higher group delay. With the same argument, the low-latency system must has larger area cost or smaller throughput. Zambreno, Nguyen, and Choudhary (2004) proposed that relationship between different filter design aspects can be sketched like following figure shows.

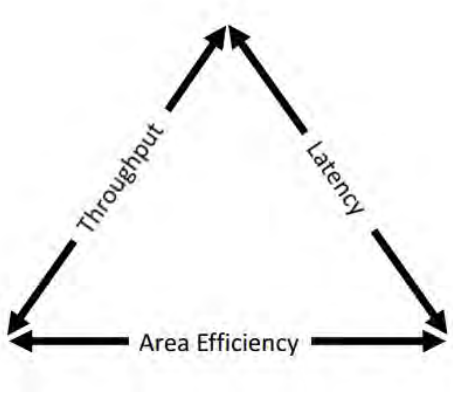



Figure 2.11: Relationship Between Different Filter Design Aspects

In above figure 2.11, throughput can be regarded as computational cost and area efficiency can be treat as area cost.

As mentioned above, both computational cost and area cost are highly correlated with filter order which means if the filter order can be estimated the computational cost and area cost can be estimated as well. Simultaneously, filter design parameters can be optimized to obtain the minimum filter order. Hence, the problem becomes how to estimate filter order accurately. Also, if the storage optimization and resource sharing can be involved, the area cost may further reduced.

However, as far as author knows, there is no systematic article on the latency estimation of high level audio filter. Therefore, the filter latency estimation

method related with filter order is derived by author and author's supervisor Yonghao Wang. This estimation method fill the gap of correlation between area cost, computational cost and latency.

### 2.3.4 Filter Order and Estimation Functions

Base on the discussion in 2.3.3, filter order is the belt that links the area cost, computational cost, and latency of the filter. Therefore, it is very important to estimate filter order accurately. Unfortunately, filters using in Σ-Δ modulation based ADC/DAC have not made new breakthroughs in these decades. Hence, even though there are several filter order estimation functions can be found for filters using in Σ-Δ modulation based ADC/DAC, the most common used estimation function is still Crochiere and Rabiner (1975) one.

**Crochiere and Rabiner 1975**

$$N_i \cong \frac{D_\infty(\delta_1, \delta_2)}{\Delta F_i} - f(\delta_1, \delta_2)\Delta F + 1 \tag{2.11}$$

Where:

$$\begin{aligned} D_\infty(\delta_1, \delta_2) =& [5.309 \times 10^{-3}(log_{10}\delta_1)^2 + 7.114 \times 10^{-2}(log_{10}\delta_1) \\ & - 0.4761]log_{10}\delta_2 - [2.66 \times 10^{-3}(log_{10}\delta_1)^2 \\ & + 0.5941(log_{10}\delta_1) + 0.4278] \end{aligned} \tag{2.12}$$

$$\Delta F = \begin{cases} \frac{(f_{ri} - f_s) - f_p}{L_i f_{r(i-1)}} & \text{For Decimation} \\ \frac{(f_{ri} - f_s) - f_p}{M_i f_{ri}} & \text{For Interpolation} \end{cases} \tag{2.13}$$

$$f(\delta_1, \delta_2) = 0.51244log_{10}(\frac{\delta_1}{\delta_2}) + 11.01217 \tag{2.14}$$

$N$ is the estimated filter order.
$\Delta F$ is the transition band width normalized to the sampling frequency.
$\delta_1$ is the tolerance in the magnitude response in the passband divide $K(\delta_p/K)$.
$\delta_2$ is the tolerance in the magnitude response in the stopband.

For high-level audio Σ-Δ ADC/DAC, most of the stages will have relatively narrow-band filters. Hence, $\Delta F$ in Equation 2.11 will be relatively small and insignificant. Therefore, the further approximation of filter order estimation

equation can be written as:

$$N \cong \frac{D_\infty(\delta_1, \delta_2)}{\Delta F} \tag{2.15}$$

**Herrmann, Rabiner, and Chan 1973**

$$N \cong -2\frac{log_{10}(10\delta_1\delta_2)}{3\Delta F} - 1 \tag{2.16}$$

Where:
$N$ is the estimated filter order.
$\Delta F$ is the transition band width normalized to the sampling frequency.
$\delta_1$ is the tolerance in the magnitude response in the passband.
$\delta_2$ is the tolerance in the magnitude response in the stopband.

**Kaiser 1974**

$$N \cong -20\frac{log_{10}(\sqrt{\delta_1\delta_2}) - 13}{14.6\Delta F} \tag{2.17}$$

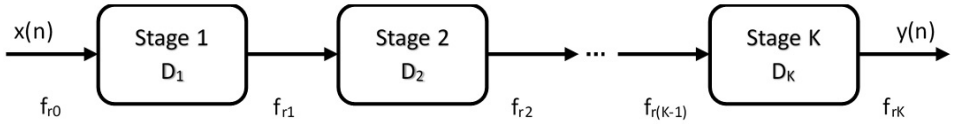
Where:
$N$ is the estimated filter order.
$\Delta F$ is the transition band width normalized to the sampling frequency.
$\delta_1$ is the tolerance in the magnitude response in the passband.
$\delta_2$ is the tolerance in the magnitude response in the stopband.

Based on these filter order estimation functions, the filter order can be estimated accurately. According to the relationship between the filter order and different cost, the required resources can be estimated as well. Hence, it is possible to find a way to design the optimal filters with above filter order estimation functions.

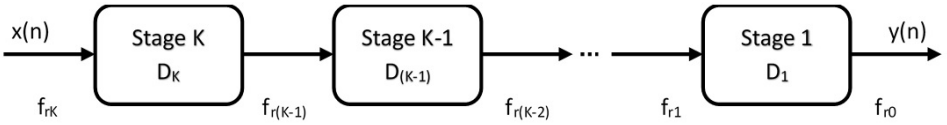
## 2.4 Optimal Decimation/Interpolation Rate Calculation Methods

Lesso and Magrath (2005) and Craven (2004) pointed out that the filter order can be very large if the transition bandwidth is narrow. As the above section 2.1.2 expresses, the Σ-Δ ADC requires very narrow transition bandwidth due to the oversampling technique. From the tests and experiences, the filter order could be more than 1000. Therefore, it will be very hard to realize the high-level audio ADC/DAC with single stage filter design.

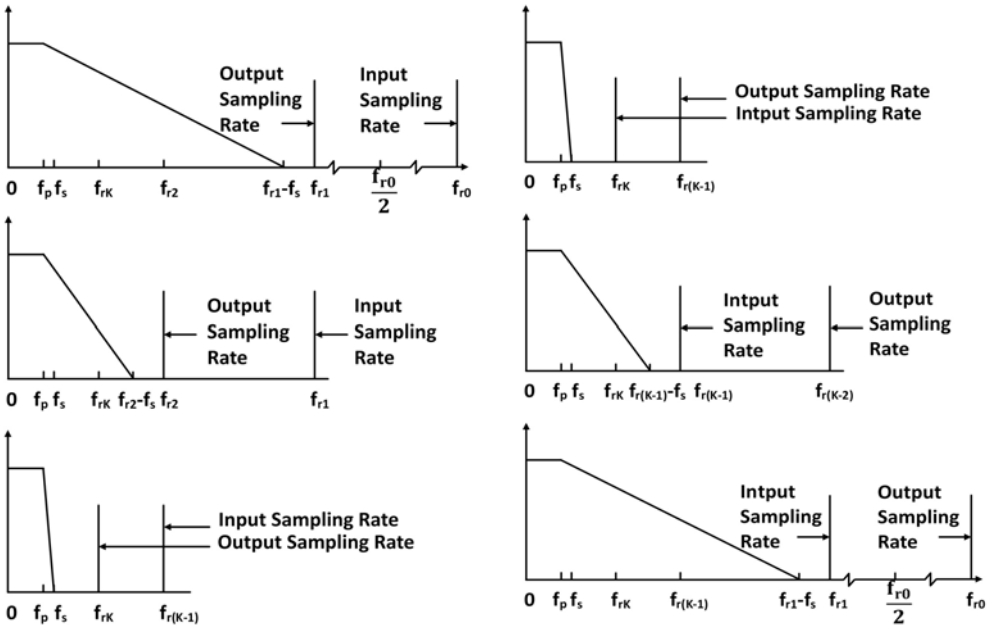
However, the multi-stage filter design technique makes the Σ-Δ ADC/DAC achievable with significant reduction of computational or area cost. Bellanger, Daguet, and Lepagnol (1974) and Nelson, Pfeifer, and Wood (1972) implemented decimation using several decimation stages which reduced the filter order significantly; however Crochiere and Rabiner (1975) mentioned that they restricted their results by only using factors of two at each stage. Shively (1975) considered a more general approach with integer decimation for a two stage design. Also, Rabiner and Crochiere (1975) extend and generalize the work of Shively (1975) using the more decimation stage. At that time, the filter order can be reduced to few hundreds which represents the filter cost becomes acceptable and the Σ-Δ modulation based ADC/DAC can be realized. Hereafter, how to design the optimal multi-rate multi-stage filter becomes the major problem. At that time, there is no affirmatory support theory to design the optimal decimation or interpolation filter mathematically. Therefore, Crochiere and Rabiner have proposed a theory to design the optimal multi-rate multi-stage filter in 1975, 1976 and 1981 (Crochiere and Rabiner 1975; 1976 and 1981). But these theories are based on the theoretical model which can be treated like the following sketched diagram.




(a) Illustration of a $K$-stage decimator




(b) Illustration of a $K$-stage interpolator




(c) Frequency response (Decimation) (d) Frequency response (Interpolation)

Figure 2.12: Sketched Block Diagram of Crochiere and Rabiner (1975)

### 2.4.1 Optimal Computational Cost(Crochiere and Rabiner 1975)

According to Section 2.3, the filter costs are mainly caused by the filter order $N$. The filter order $N$ can be estimated by using Equations 2.11, 2.16 and 2.17. Therefore, Crochiere and Rabiner (1975) convert the optimal decimation or interpolation rate selection problem into a minimum filter order finding problem.

Also, 2.11, 2.16 and 2.17 demonstrate that the key factors affect the filter order most are passband ripple, stopband attenuation and transition bandwidth. Due to the absolute values of passband ripple and stopband attenuation are fixed relatively, the transition bandwidth becomes the target to be optimized.

Hence, the sketched diagram of theoretical model for optimal multi-rate multi-stage filter design was proposed by Crochiere and Rabiner in 1975. In this model, the passband and stopband ripples are ignored as Figure 2.12 shows.

According to this model, the total computational cost $R_T$ can be treated as the summation of each filter stage's cost $R_i$.

$$R_T = \sum_{i=1}^{K} R_i \tag{2.18}$$

And the current filter's computational cost $R_i$ can be expressed as the product of current stage's filter order $N_i$ times the sampling rate of the filter.

$$R_i = \frac{N_i L_i f_{r(i-1)}}{L_i M_i} = \frac{N_i f_{r(i-1)}}{M_i} \tag{2.19}$$

According to the above model, the transition bandwidth $\Delta F$ can be summarized as:

$$\Delta F = \frac{(f_{ri} - f_s) - f_p}{L_i f_{r(i-1)}} \tag{2.20}$$

Substitute Equations 2.15 and 2.20 into 2.19, it can be obtained that:

$$R_i \cong D_\infty(\frac{\delta_p}{K}, \delta_s) \frac{D_i f_{ri}^2}{f_{ri} - f_s - f_p} \tag{2.21}$$

Hence, the total computational cost $R_T$ can be summarized as:

$$R_T \cong D_\infty(\frac{\delta_p}{K}, \delta_s) f_{r0} \sum_{i=1}^{K} \frac{D_i}{(\prod_{j=1}^{i} D_j)(1 - \frac{f_s+f_p}{f_{r0}} \prod_{j=1}^{i} D_j)} \tag{2.22}$$

Assume the last stage's sampling frequency $f_{rK}$ is equal to two times of the stopband edge $f_s$, it can be obtained that $\Delta f = \frac{f_s - f_p}{f_s}$. If we abstract the last stage from the summation part of Equation 2.22, the estimation function for the total computational cost $R_T$ can be simplified as:

$$R_T \cong D_\infty(\frac{\delta_p}{K}, \delta_s) f_{r0} S \tag{2.23}$$

Where $D_\infty(\frac{\delta_p}{K}, \delta_s)$ can be calculated by equation 2.12 and:

$$S = \frac{2}{(\Delta f \prod_{j=1}^{K} D_j)} + \sum_{i=1}^{K-1} \frac{D_i}{(\prod_{j=1}^{i} D_j)(1 - (\frac{2-\Delta f}{2D}) \prod_{j=1}^{i} D_j)} \tag{2.24}$$

### 2.4.2 Optimal Memory Usage(Crochiere and Rabiner 1976)

With the same argument of estimation function of computational cost $R_T$, Crochiere and Rabiner find that the area cost can be expressed as the sum of the lengths of the filter in each stage. Therefore, Crochiere and Rabiner has proposed the estimation function for filter's area cost (memory usage) in 1976. And the approximate objective function can be written as:

$$N_T \cong G \sum_{i=1}^{K} N_i \tag{2.25}$$

Where $N_T$ is the total number of necessary storage requirement, $N_i$ is the length of the FIR filter for $i$-stage, $K$ is the total number of stages and G is a proportionality constant.

$$N_i \approx \frac{D_\infty(\frac{\delta_p}{K}, \delta_s) D_i L_i}{1 - \frac{2-\Delta f}{2D} \prod_{j=1}^{i} D_j} \tag{2.26}$$

Where $D$ is the total upsampling or downsampling rate, $\Delta f = (f_s - f_p)/f_s$, $f_p$ is the passband edge, and $f_s$ is stopband edge which can be assumed as

half of the final sampling frequency.

$$N_T \approx GD_\infty(\frac{\delta_p}{K}, \delta_s) \sum_{i=1}^{K} \frac{D_i}{1 - \frac{2-\Delta f}{2D} \prod_{j=1}^{i} D_j} \tag{2.27}$$

Hence, the estimation function can be simplified to:

$$N_T \approx GD_\infty(\frac{\delta_p}{K}, \delta_s)T \tag{2.28}$$

Where $D_\infty(\delta_p/K, \delta_s)$ depending only upon the stopband and passband ripple levels is only a slowly varying function of $K$. The key multivariable factor $T$ can be summarized as:

$$T = \frac{2}{\Delta f} \frac{D}{\prod_{j=1}^{K-1} D_j} + \sum_{i=1}^{K-1} \frac{D_i}{1 - \frac{2-\Delta f}{2D} \prod_{j=1}^{i} D_j} \tag{2.29}$$

Therefore, the total area cost $N_T$ is minimized by minimizing $T$ while the number of filter stage $K$ is fixed.

### 2.4.3 Improved Optimal Computational Cost(Coffey 2003)

As the above part mentioned, Crochiere and Rabiner have proposed the optimal decimation or interpolation rate calculation method for optimal computational cost filter design in 1975 and 1981. According to Crochiere and Rabiner (1975; 1981), the proposed optimal decimation or interpolation rate finding problem is an optimization process. Therefore, the solution can not be found directly. Hence, Coffey (2003) has proposed another method to find the optimal decimation or interpolation rate for each stage based on Crochiere and Rabiner (1975; 1981). This approach reduces the complexity of the problem by mathematically using Partial Differential Equations (PDE).

According to Coffey (2003), for the given number of stages, the minimized computational cost $R_T$ can be found by minimizing the multivariate function $S$ as mentioned in Equation 2.24.

Coffey (2003) has demonstrated how to calculate the decimation or interpolation rate of each stage for optimal computational cost $R_T$ when the filter stage is less than 5. If we assume $f_2 = 2 - \Delta f$, for a 2-stage filter design

$(K = 2)$, S depends only upon $D_1$, and the optimal solution, found by putting $dS/dD_1 = 0$, is given by

$$D_1 = \frac{2D}{f_2} \frac{1}{(1 + \sqrt{D\Delta f/f_2})} \tag{2.30}$$

$$D_2 = D/D_1 \tag{2.31}$$

When filter stage $K = 3$, the factor $S$ in Equation 2.24 can be written as:

$$S = \frac{2}{\Delta f D_1 D_2} + \frac{1}{1 - f_2 D_1/2D} + \frac{1}{D1(1 - f_2 D1D2/2D)} \tag{2.32}$$

In order for $S$ to be minimized, it is necessary that $\partial S/\partial D_1 = \partial S/\partial D_2 = 0$. Hence, two equations can be deduced.

$$\begin{aligned} D_1^2 \frac{\partial S}{\partial D_1} = &- \frac{2}{\Delta f D_2} + \frac{f_2}{2D} \frac{D_1^2}{(1 - f_2 D_1/2D)^2} \\ &- \frac{1}{(1 - f_2 D_1 D_2/2D)} + \frac{f_2}{2D} \frac{D_1 D_2}{(1 - f_2 D_1 D_2/2D)^2} = 0 \end{aligned} \tag{2.33}$$

$$D_2^2 \frac{\partial S}{\partial D_2} = -\frac{2}{\Delta f D_1} + \frac{f_2}{2D} \frac{D_2^2}{(1 - f_2 D_1 D_2/2D)^2} = 0 \tag{2.34}$$

Therefore, according to Equation 2.34, it can be deduced that

$$D_2(D_1) = \frac{2D}{f_2} \frac{1}{(D_1 + \sqrt{D\Delta f/f_2}\sqrt{D_1})} \tag{2.35}$$

By combining Equations 2.33 and 2.34, the equation can be found for $D_2$ in terms of $D_1$. If we combine the equation $(D_2/D_1)\partial S/\partial D_2 = 0$ and $\partial S/\partial D_1 = 0$, it can be deduced

$$D_2(D_1) = \frac{\alpha D_1^2 - \alpha^2 D_1^2 + 2\alpha D_1 - 1}{\alpha^2 D_1^3} \tag{2.36}$$

Where $\alpha = f_2/2D$.

Hence, the first decimation rate or the last interpolation rate $D_1$ can be calculated by using the combination of Equations 2.33 and 2.35 or Equations

2.34 and 2.36. Therefore $D_2$ can be found if we substitute $D_1$ into the equations. Finally, $D_3$ can be solved by using the relationship between the total decimation or interpolation rate and each stage's decimation or interpolation rate which is $D = D_1 D_2 D_3$.

For four stages, we can still start from Equation 2.24. It can be found that:

$$D_3(D_1 D_2) = \frac{2D}{f_2} \frac{1}{\sqrt{D_1 D_2}(\sqrt{D_1 D_2} + \sqrt{D\Delta f / f_2})} \tag{2.37}$$

$$\begin{aligned} D_1^2 \frac{\partial S}{\partial D_1} =& -\frac{2}{\Delta f D_2 D_3} + \frac{f_2}{2D} \frac{D_1^2}{[1 - f_2 D_1/2D]^2} - \frac{1}{[1 - f_2 D_1 D_2/2D]} \\ &+ \frac{f_2}{2D} \frac{D_1 D_2}{[1 - f_2 D_1 D_2/2D]^2} - \frac{1}{D_2} \frac{1}{[1 - f_2 D_1 D_2 D_3/2D]} \\ &+ \frac{f_2}{2D} \frac{D_1 D_3}{[1 - f_2 D_1 D_2 D_3/2D]^2} = 0 \end{aligned} \tag{2.38}$$

and

$$\begin{aligned} D_2^2 \frac{\partial S}{\partial D_2} =& -\frac{2}{\Delta f D_1 D_3} + \frac{f_2}{2D} \frac{D_2^2}{[1 - f_2 D_1 D_2/2D]^2} \\ &- \frac{1}{D_1} \frac{1}{[1 - f_2 D_1 D_2 D_3/2D]} \\ &+ \frac{f_2}{2D} \frac{D_2 D_3}{[1 - f_2 D_1 D_2 D_3/2D]^2} = 0 \end{aligned} \tag{2.39}$$

Since $D_3$ can be known from Equation 2.37, the problem becomes a two dimensional root finding problem. If we combine Equation 2.38 and 2.39, it can be found that:

$$D_2(D_1) = \frac{\alpha D_1^2 - \alpha^2 D_1^2 + 2\alpha D_1 - 1}{\alpha^2 D_1^3} \tag{2.40}$$

Therefore, the two dimensional root finding problem can be simplified to a one dimensional root finding problem. By combining the Equation $(D_3/D_2)\partial S/\partial D_3 = 0$ with $\partial S/\partial D_2 = 0$, it can be found:

$$D_3(D_1, D_2) = \frac{\alpha D_1 D_2^2 - \alpha^2 D_1^2 D_2^2 + 2\alpha D_1 D_2 - 1}{\alpha^2 D_1^2 D_2^3} \tag{2.41}$$

According to Crochiere and Rabiner (1975; 1981), the filter's computational

cost $R_T$ can be decreased significantly by using multi-stage design. However, Crochiere and Rabiner (1975) also mentioned that the influence of key factor $S$ would be reduced significantly as well while the filter had four or more stages. Hence, only 2-4 stage's solution methods are proposed by Coffey (2003).

### 2.4.4 Improved Optimal Memory Usage(Coffey 2007)

In 2007, Coffey published another paper based on Crochiere and Rabiner (1976) and Coffey (2003) which proposed the optimal decimation or interpolation rate calculation method for area cost (storage requirement).

With the same argument of relationship between $S$ and computational cost $R_T$, only 2 and 3 stages' optimal decimation or interpolation rate calculation method for area cost has been deduced by Coffey (2007).

For stage $K = 2$, $T$ depends only upon $D_1$, therefore, the optimal solution can be found by letting $dT/dD_1 = 0$ which is:

$$D_1(D) = \frac{2D}{f_2} \frac{1}{(1 + \sqrt{2D\Delta f}/f_2)} \tag{2.42}$$

$$D_2 = D/D_1 \tag{2.43}$$

Where $f_2 \equiv 2 - \Delta f$

When $K = 3$, $T$ can be simplified to

$$T = \frac{2D}{\Delta f D_1 D_2} + \frac{D_1}{1 - \alpha D_1} + \frac{D_2}{1 - \alpha D_1 D_2} \tag{2.44}$$

In order to minimize the key factor $T$, it is necessary to make $\partial T/\partial D_1 = \partial T/\partial D_2 = 0$. Therefore, it can be obtained:

$$D_2(D_1) = \frac{2D}{f_2} \frac{1}{(D_1 + \sqrt{2D_1 D\Delta f}/f_2)} \tag{2.45}$$

By letting the result $1-[(2-\Delta f)/2D]D_1D_2 = \sqrt{\Delta f D_1/2D}D_2$ from $\partial T/\partial D_2 = 0$ into the equation $\partial T/\partial D_1 = 0$, latter equation can be written as:

$$\frac{\partial T}{\partial D_1} = -\frac{2D}{\Delta f D_1^2 D_2} + \frac{1}{(1 - \frac{2-\Delta f}{2D} D_1)^2} + (\frac{2 - \Delta f}{2D}) \sqrt{\frac{2D}{\Delta f} \frac{1}{D_1}} D_2 = 0 \tag{2.46}$$

## 2.5 Summary

In this chapter, author has reviewed the background of this project. With the pursuit of high-level audio and advancement of technology, high-level audio has become one of the indispensable components of modern electronic products. However, the requirements for high resolution and low-latency make the design of high-level audio system extremely demanding. Due to the principle of PCM, the accuracy of sampling cannot be guaranteed at such a high resolution, and there is also the need for additional anti-aliasing and anti-imaging filters. This has lead to the $\Sigma$-$\Delta$ modulation based digital audio system gradually becoming the mainstream. The main feature of the $\Sigma$-$\Delta$ modulation based digital audio system is to use high sampling frequency in exchange for high resolution. At the same time, aliasing and imaging noise can be filtered out in the process of up or down sampling, thereby reducing the external support circuits. However, the $\Sigma$-$\Delta$ modulation based digital audio system has very stringent design requirements for filters in decimator and interpolator, which causes most of the system latency is generated by these filters. Therefore, how to reduce the latency as much as possible while meeting the design requirements has become the main difficulty in the decimation or interpolation filter design.

Based on the above literature review, author further reviewed the two major characteristics of the filter: linearity and time-invariance. Then based on these two characteristics, representative filter design types: IIR and FIR filters are briefly analysed. Among them, IIR filter has the advantage of cost efficient and low-latency, but the changing latency and stability concern make it difficult to perform well in the high-level audio area. As a linear filter, FIR filter has advantage of constant latency and high stability. However, the latency is quite high compared to the IIR filter.

Hereafter, author reviews and discusses several kinds of filter cost, and points out that the filter order is the key factor affects different costs. Therefore, the problem becomes how to reduce and estimate filter order. Since the optimal FIR filter design has not made substantial progress in recent decades, the most authoritative and accurate estimation method at this stage is the still the estimation equations proposed by Crochiere and Rabiner in 1975. While at the same time, they also proposed the method of using multi-stage filter design to reduce the filter order. Kale et al. (1995) stated this method will be able to significantly reduce a single stage filter with several thousand orders to several hundred orders with only a small increase in filter latency. This means the $\Sigma$-$\Delta$ modulation based digital audio system can finally be

truly realized. However, in the filter design, the calculation of decimation or interpolation rate of each stage is very complicated and inefficient. Hereafter, Coffey improved the decimation or interpolation rate calculation method for optimal computational and area cost filter design in 2003 and 2007, thereby convert decimation or interpolation rate calculation from the optimization into a one-dimensional root finding. However, the result of this calculation method is still non-integer, and further rounding and selection is needed to find the optimal integer solutions. Hence, author proposed a new optimal decimation or interpolation rate selection approach which has advantage of computationally convenient and also can provide directly usable optimal integer solutions.

Meanwhile, author found that although many researchers have made a lot of research and contributions to the hardware design of low-latency filter (Peled and Liu, 1974), no researchers have proposed the latency estimation method for optimal FIR filter. Similarly, author also found that the current research on the optimization of filters is mainly focused on a certain performance of the filter rather than global optimization. However, as above mentioned, it can be known that the $\Sigma$-$\Delta$ modulation based digital audio system has comprehensive requirements for filters, not only for latency but also different costs and performances. Therefore, how to optimize the overall filter performance and construct a filter evaluation and design framework with only few amount of relevant literature is a great challenge.

# Chapter 3

# Methodology

As mentioned above, the current background is more and more mobile devices equipped with high quality audio, At the same time, in the professional area, more and more stringent technical indicators have been proposed. Like the latency requirement for singer and musicians' in-ear monitoring.

The Σ-Δ modulation based audio system is the solution to solve the high-resolution requirement problem. However, the concept of the Σ-Δ modulation based system is using very high sampling frequency to represent the signal. Therefore, the Σ-Δ modulation based system need to use oversampling compared to traditional PCM based system which means the system needs decimation and interpolation process. And this process will worsen the latency performance and add extra calculations or operations.

The aim of research is to carry out a comprehensive evaluation as well as to propose a framework for the digital filter design with both low latency and low cost for Σ-Δ Modulation based ADC/DAC. Following above mentioned requirements and limitations, the research question is How to design a cost efficient low-latency multi-stage multi-rate filter? If this question is asked in detail, it will lead us to the following 3 questions:

1. Can the parameter selection be improved?
2. How to trade off contradictory design requirements?
3. Is there any existing design or trade off method can meet both low latency and hardware efficient design requirements?

In order to answer the first question, optimal filter design and parameter selection or calculation methods need to be researched in depth which is link to objective 1 and part of objective 2 listed in Section 1.2.

Although there are many filter design methods and parameters, most of them are limited by design requirements. Therefore, only few parameters need to be researched. From previous research, it can be obtained that filter order is the most important factor which affects filter costs directly. Thus, the problem becomes how to minimize the filter order. However, filter order is a property of filter itself. therefore, the reduction of filter order can only be achieved by adjusting other design parameters or using special filter types.

Even though, the filter order can be reduced significantly using special filter types, the use of these filters has many limitations or may worsen other filter performances. For example, halfband filter can reduce the filter's computational and area cost significantly with increasement of latency. Also, halfband filter has limitation that the passbaand edge and stopband edge cannot be set directly. And the biggest limitation on the use of halfband filter is the oversampling rate can only be set to two. Hence, using different filter types to design cost efficient filter can only be an auxiliary design method. Therefore, this research is focus on how to optimize the filter design parameters to reduce the filter order.

Nelson, Pfeifer, and Wood (1972) and Bellanger, Daguet, and Lepagnol (1974) implemented decimation using several decimation stages which reduced the filter order significantly, Shively (1975) considered a more general approach with integer decimation for a two stage design. Also, Crochiere and Rabiner (1975) extend and generalize the above work and using more decimation stages which reduce the filter order from thousands to hundreds.

Although these methods can reduce the filter order greatly, the calculation process is quite complicated and the results cannot be used directly due to the non-integer solutions. Therefore, how to calculate out direct usable integer results becomes the new problem. Coffey (2003,Coffey 2007) proposed a approach which can reduce the complexity of calculation significantly, however, the results are still non-integer. Hereafter, Huang (2003) and Huang and Hung (2009) considered an Exhaustive and Genetic algorithm based approach using set theory which can provide direct usable integer solutions. However, the search process of this approach is not much more efficient than the previous calculation methods.

Hence, a new optimal decimation or interpolation rate selection approach is proposed by author using set theory and knowledge based search algorithm. The new method has relatively convenient calculation process and usable integer solutions. This approach converts the optimization problem becomes the factorization and permutation problem. Also based on this approach a knowledge based lookup tables searching method has proposed by author as well. At this point, the first question is answered by author.

During the research, author found that currently, there is no theoretical relationship between latency and multi-stage design parameters. As mentioned above, latency is a very important property of filter which affect the feelings of user. Therefore, in order to built up the link between latency and filter design parameters for further analysis and optimization, the latency estimation equation is derived based on the assumption and filter order estimation equations of Crochiere and Rabiner (1975).

However, during the analysis of derived latency estimation equation, author found optimal latency and optimal computational or area cost filter designs have contradictory requirements. For example, fewer filter stages leads to more computational or area cost but lower latency, distribution of optimal decimation rate for computational or area cost is from large to small, however, overall latency tends to be smaller when distributing larger decimation rate towards later stage. This finding leads to the second question: How to trade off contradictory design requirements?

Since the above mentioned characteristics can be proved by mathematics, hence, there is no design method that can calculate out a distribution of decimation or interpolation rate has both optimal latency and computational or area cost. So the author proposes two directions to solve this problem. One is to abandon the optimal design and search a relatively balanced semi-optimal solution which is published in Zhu et al. (2016). The other one is to optimize other filter design parameters based on optimal computational or cost or latency design. Since the first direction is already been published, here will focus on the second direction.

Based on the analysis of estimation equation for different filter costs, author found the passband ripple and transition bandwidth still have room to be optimized due to the imperfect allocation rules. So far, author has completed the first two objectives while answering the first two questions.

From the previous research, it can be known that the current optimal multi-

stage multi-rate filter design always allocate passband ripple equally. To author's best knowledge, there is no filter has same costs for each stage. Simultaneously, it can be deduced from estimation equations that the both passband ripple and transition bandwidth can improve performance of latency, area and computational cost at the same time. Hence, the problem becomes how to allocate given passband ripple and wider transition bandwidth.

In order to solve the above question, numeric optimization method is needed to search the optimal solutions. At this moment, author found larger stopband attenuation may lead to wider transition bandwidth. Therefore, the third factor is introduced. Based on the principle of filter design, each of these three factors link to filter coefficients which means these factors need to be adjusted together otherwise, any one of them is adjusted the whole filter will be changed.

Hence, the Simulated Annealing algorithm has been introduced to search the optimal solutions. Base on the searched optimal design parameters, cost efficient low-latency multi-stage multi-rate filter can be designed. As the result, filter designed by author proposed approach has around 3%-4% improvement for different costs. Till now, all three questions are answered by author and only the last objective is left.

For the last objective, MATLAB GUI based cost efficient low-latency filter design optimization method and evaluation framework is proposed by author. this framework is based on the above research outcomes which can help user to search the optimal design parameters, also this framework support filter design with different filter types for example, halfband filter. In the future, more filter types could be added as well.

# Chapter 4

# Cost Efficient Decimation Rate Selection and Filter Design

## 4.1 Limitations of Existing Approaches

As Chapter 2 mentioned, most of the optimal decimation or interpolation rate selection methods are based on Crochiere and Rabiner (1975; 1976), and the nature of these methods are optimization problem. Therefore, the results cannot be found directly. Hence, Coffey (2003; 2007) drastically reduce the complexity of the problem mathematically by using PDE. But there are two limitations on the applications of these methods.

1. Although Coffey converts the problem into a one dimensional equation. Commonly, the order of the equation of 3-stage filter is eight. And the order of the 4-stage filter is ten. Thus, numeric methods are still needed to solve the equations, and for each design, the roots of the equation still need to be put back to cost function to find out the optimal solution set.

2. The optimal solutions are often groups of non-integer real numbers that cannot be implemented in practical system. A manual adjustment of the results is needed.

Nevertheless, Coffey's approaches provide a clear route to enable us to investigate the solutions to further study some interesting and useful properties. Huang (2003; 2009) can yield the integer solution directly. Huang (2003) represents this problem in integer domain by using set theory, and Huang and Hung (2009) showed that the problem can be solved by using either an

exhaustive search or a genetic algorithm.

However, based on the experiments, some important properties of real valued and integer valued solutions are found by author which can help to simplify the search algorithm. Also when looking into the optimal solutions sets, author infer that the optimal results might be a limited number of sets for any types of design within certain constrains and design specifications. Therefore, a new knowledge based search algorithm mainly based on the properties of the distribution of the solution sets is proposed by author. Based on our inference, this algorithm is also used to generate a set of 3D optimal solution database for practical useful design specifications.

## 4.2 Description of the Background

Based on the Equation 2.22 to 2.24 and Equation 2.28 to 2.29 in Chapter 2, the computational cost and area cost of filter design can be calculated using design parameters. Among them, $D_\infty$, $f_{r0}$ and $G$ can be regarded as constants under the given design requirements of stage number, passband ripple, stopband attenuation and initial sampling frequency of input signal. Therefore, in these cases where other design parameters can be regarded as constants, $S$ and $T$ become the key factors affect the computational cost and area cost.

Since the design method of the optimal linear FIR filter used in $\Sigma$-$\Delta$ modulation based audio system has not changed in recent decades, therefore, the selection of optimal filter design parameters are still follows the rules which proposed by Crochiere and Rabiner (1975). However, the calculation method has been improved several times by Coffey (2003, 2007) and Huang (2003, 2009).

Therefore, in order to find out the distribution of optimal filter decimation or interpolation rate and improve the current optimal filter design method, the influence of $S$ and $T$ on optimal computational and area cost designs are analysed and tested in the following sections.

### 4.2.1 Computational Cost

As mentioned above, the key factor affecting the computational cost $R_T$ is $S$ in Equation 2.24. For the certain filter design requirements, the oversampling rate $D$, filter stage $K$ and transition bandwidth $\Delta f$ are given, therefore, the decimation or interpolation rate for each stage $D_i$ is the only parameter that can affect the $S$. Thereby, As the following Figure 4.1 demonstrated, Crochiere and Rabiner 1975 (1975) calculated out the distribution of $S$ with different transition bandwidth and filter stages relative to the oversampling rate.

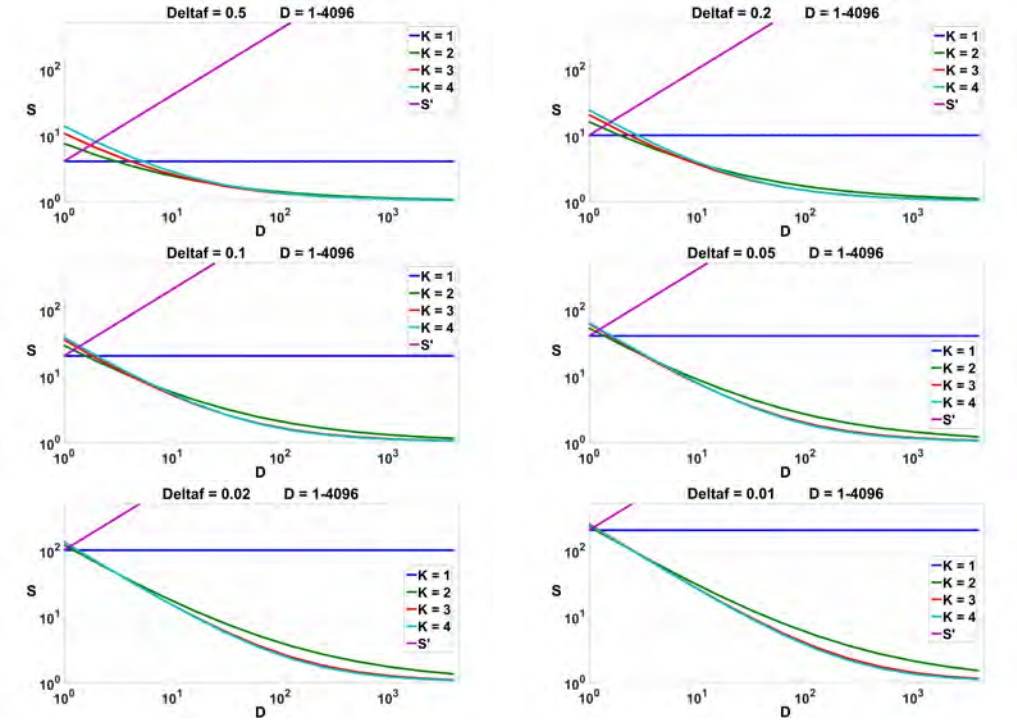



Figure 4.1: Distribution of Minimized '$S$' (Crochiere and Rabiner,1975)

Where $S' = f_{r0}/(f_s - f_p)$.

From the above Figure 4.1, it can be obtained that the value of key factor $S$ is heavily influenced by the system's transition bandwidth $\Delta f$. And the multi-stage filter design method can reduce the computational cost significantly. Also the difference between 3-stage and 4-stage design is slight in comparison with the difference between 3-stage and 2-stage or single stage design.

In order to test and verify the above derivation and analysis, the following experiment has been made by author. The goal of this experiment is to plot the distribution of decimation or interpolation rate for each stage in the reasonable region. Therefore, the total oversampling rate $D$ has been limited below 4100, and the transition bandwidth $\Delta f$ has been limited below 0.5 due to the Nyquest sampling theory. Also the other typical filter design parameters are given in Table 4.1.

Table 4.1: Test Conditions

| | |
|---|---|
| Filter Stage ($K$) | 1-4 |
| Sampling Frequency ($Fs$) | 48kHz |
| Passband Ripple ($A_p$) | 0.0001dB |
| Stopband Attenuation ($A_{st}$) | 120dB |

Thus, the distribution of optimal decimation or interpolation rate for each stage can be obtained and plotted once the optimal solutions have been solved with Coffey (2003).

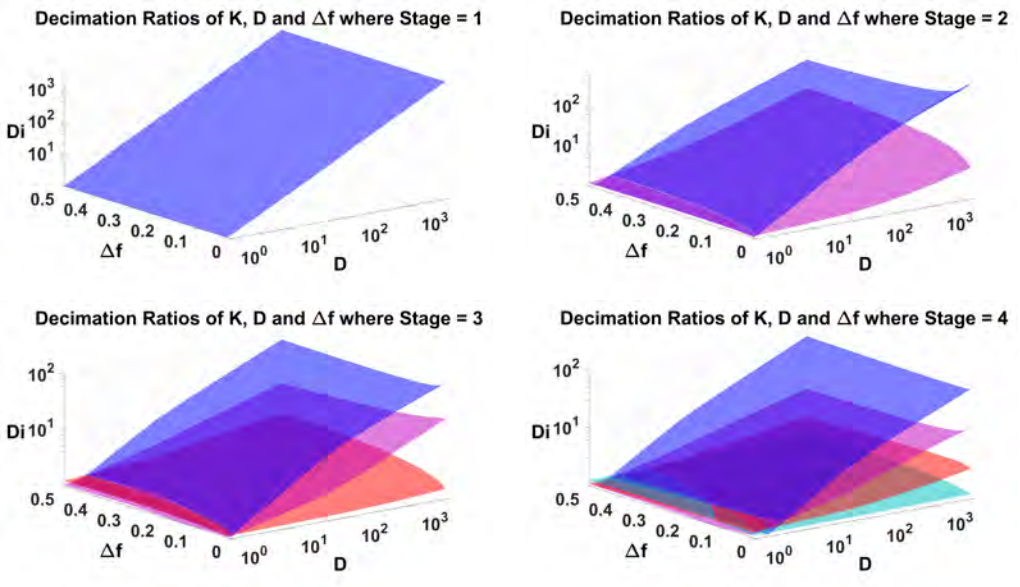



Figure 4.2: Decimation Rate Distribution of Optimal Computational Cost Filter Design

The above Figure 4.2 are the test results. Each surface represents oversampling rate distribution of one stage where the blue surface is the first stage, the pink one is the second stage, the orange one is the third stage and the cyan surface is the fourth stage. Also $D$ is the overall oversampling rate, $Di$ is the oversampling rate for each stage, and $\Delta f$ is the transition bandwidth.

From observation, it can be known that in the two to four stage optimal computational cost filter design, no matter how many stages the filter has and how narrow the transition bandwidth is, the distribution of decimation rate always followed the rule which is the front stage has larger decimation rate.

Therefore, base on the above findings and discussions, the following characteristics can be summarized:

- The filter's computational cost $R_T$ is mainly affected by the transition bandwidth and number of stages.
- The distribution of filter's optimal decimation rate is arranged from large value to small value (small value to large value for interpolation).
- The passband ripple and stopband attenuation affect the computational cost $R_T$ as well. But due to the absolute value of $D_\infty$ is not in the same magnitude comparing with the filter's initial sampling frequency $f_{r0}$ and key factor $S$, and the changing of passband ripple and stopband attenuation affect the $D_\infty$ with limited impact. Therefore, Crochiere and Rabiner (1975) and Coffey (2003) didn't pay too much attention to these factors.
- Although the optimal decimation or interpolation rate for each stages can be calculated by these mentioned design methods, the implementation problem is still exist because of the decimal results.

### 4.2.2 Area Cost

According to Crochiere and Rabiner (1976), the key factor affecting the area cost $N_T$ is $T$ in Equation 2.29. With the same argument mentioned above, the characteristic curve can be drawn for the certain filter design requirements. Hence, if the filter stage is limited in the range of 1-3, the system's oversampling rate is below 4096, and the $\Delta f$ is selected between 0.01 and 0.5. The following curves can be obtained using Equation 2.29.

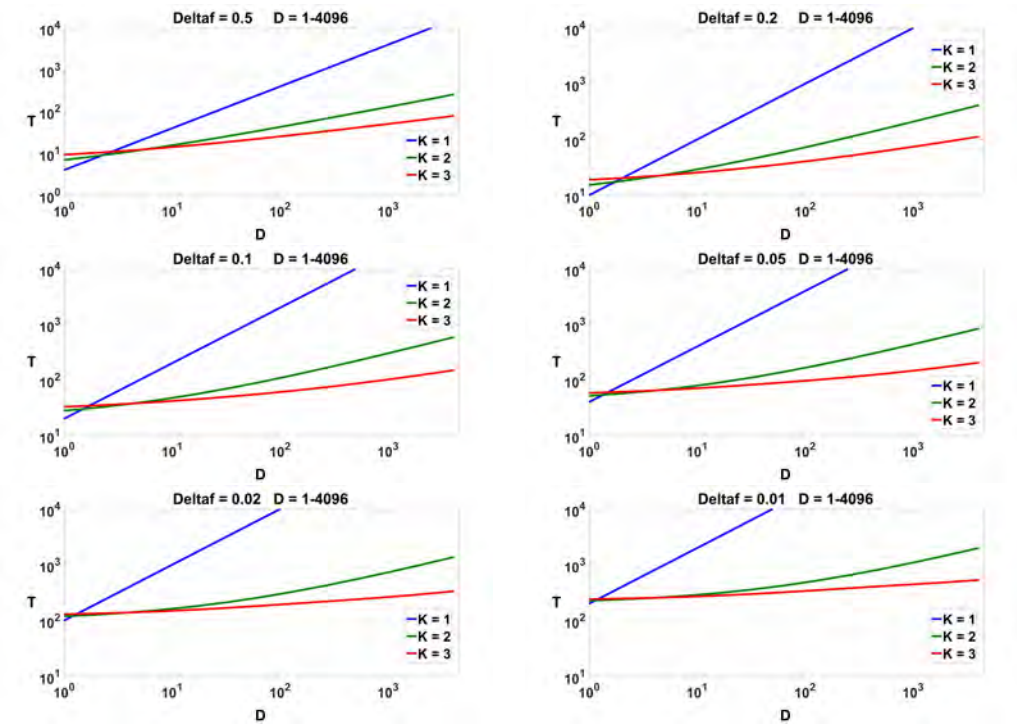



Figure 4.3: Distribution of Minimized '$T$' (Crochiere and Rabiner,1976)

Comparing with the above $S$ curves of computational cost, the characteristics of $T$ curves are similar to $S$ curves which increase heavily with the decrease of filter transition bandwidth and the value of key factor $T$ can be reduced significantly by using multi-stage design method. The only difference between $S$ curves and $T$ curves is the tendency of these curves. In Figure 4.1, although the $S$ curves are strongly affected by the transition bandwidth $\Delta f$, in Figure 4.3, the $T$ curves are more affected by the transition bandwidth $\Delta f$ than the $S$ curves.

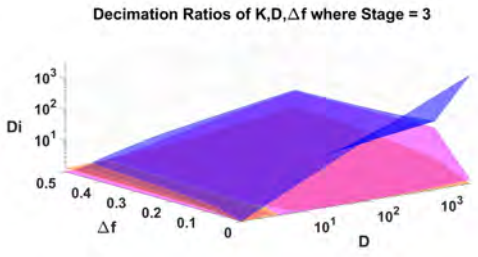



Figure 4.4: Decimation Rate Distribution of Optimal Area Cost Filter Design

Following the same test method and conditions mentioned in section 4.2.1, the decimation rate distribution of the typical 3-stage optimal area cost filter design can be drawn as Figure 4.4 shows. Comparing with the decimation rate distribution of optimal computational cost filter design, just like mentioned above, the decimation rate of optimal area cost filter design is strongly affected by the transition bandwidth $\Delta f$. Also the decimation rate distribution of optimal area cost filter design follows the same rule with optimal computational cost filter design which is the front stage has larger decimation rate. Therefore, it can be obtained that the distribution of optimal computational and area cost filter design have the same characteristics, the only different is the degree of influence by filter design parameters.

Thus, base on the above research and findings, whether it is the optimal computational or area cost filter design, the following characteristics are followed:

- Fewer filter stages leads to more computational or area cost.
- When filter stage is greater than three, the reduction of cost is insignificant.
- Decimation rate distribution of both optimal computational or area cost follows the same rule which is front stage has larger decimation rate.
- The key factors affect computational or area cost are transition bandwidth and overall oversampling rate.

Although above characteristics have been found by author, how to design the optimal filter efficiently is still an unsolved problem. However, base on the previous research and findings, if the key factors like number of filter stage, oversampling rate for each stage and the transition bandwidth of each stage can be determined, the optimal filter design problem can be easily solved. Hence, a new optimal decimation or interpolation rate selection approach is proposed by author.

## 4.3 Optimal Decimation/Interpolation Rate Selection Approach

In order to calculate and implement the optimal multi-rate multi-stage filter design efficiently, a new decimation/interpolation rate selection approach has

been proposed by author. The basic idea of this approach is to find out the optimal combination of decimation/interpolation rates for each stage in a limited set instead of finding out the solution mathematically. Precision and efficiency are the advantages of this new approach, because the operation complexity has been reduced significantly and the follow-up rounding and adjustments can be omitted.

As Candy (1986) and Yang, Sculley, and Abraham (2007) stated for $\Sigma$-$\Delta$ Modulation based ADC/DAC design, the oversampling rate, passband edge, passband ripple, stopband edge and stopband attenuation are normally given in the filter design requirements. Therefore, $\Delta f$ can be calculated and the only two unknown parameters are: filter stage N and decimation/interpolation rate for each stage $D_i$. Hence, the problem can be converted to two problems. the first one is to find the optimal decimation/interpolation rate, and the other one is to find the optimal filter stage.

From the above discussion and experience, it can be found that the distribution of optimal decimation/interpolation rate for optimal computational cost design is from large scopes to small scopes. Therefore, the optimization problem becomes to the permutation problem. Thus, the approach can be designed as:

1. **Find optimal solution for N-stage filter design**
   (a) Oversampling rate $D$ factorization.
   (b) Find possible solution sets $D_i$.
   (c) Sort and rearrange factors and sets.
   (d) Search for the optimal set within the possible $D_i$ sets.
2. **Find out the optimal filter design from above results**
   (a) Calculate the computational cost $R_T$ for each solutions.
   (b) Find out the optimal result.

According to the $1^{st}$ step of this approach, the flow chart can be drawn as the Figure 4.5 demonstrates.

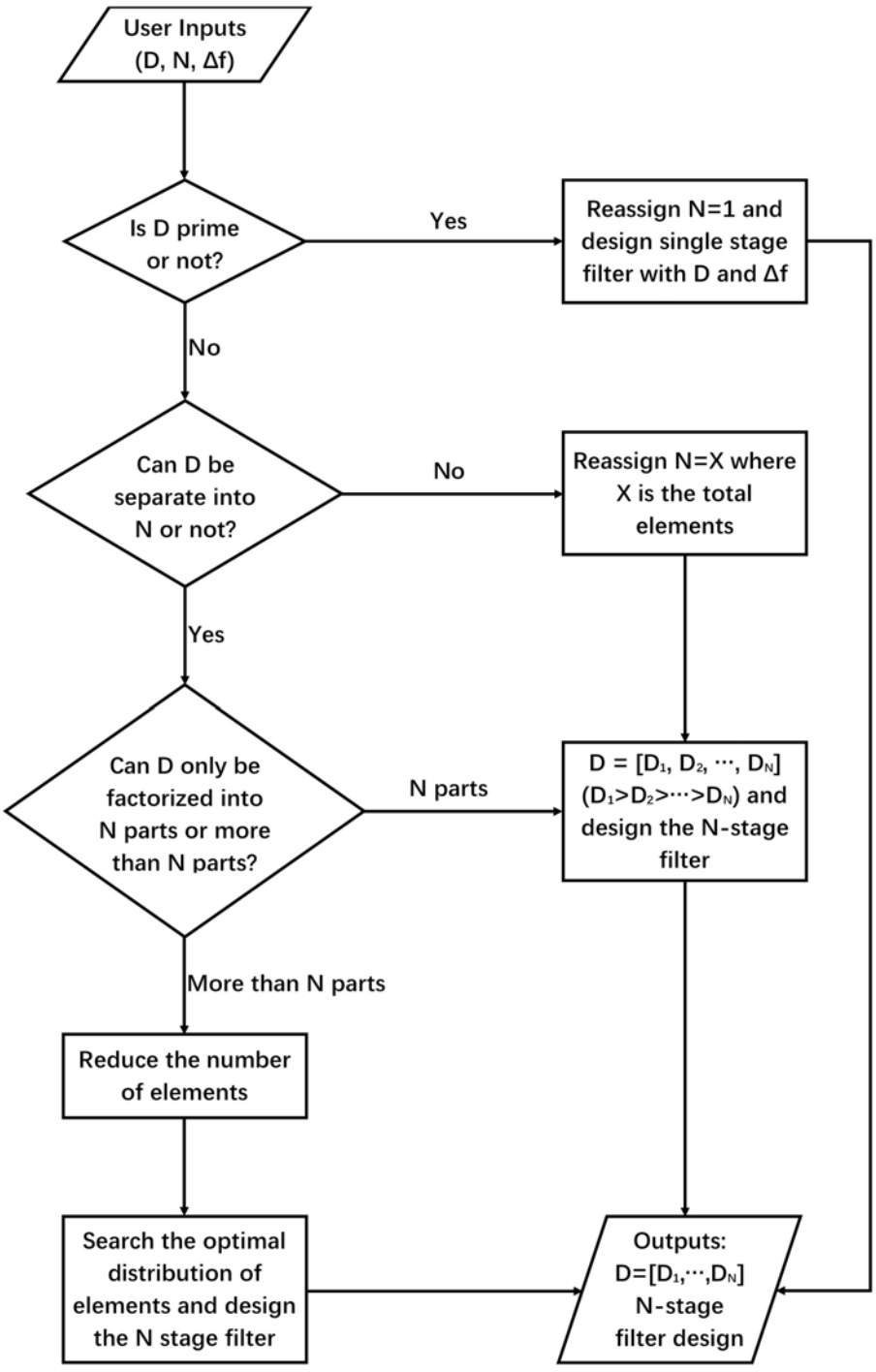



Figure 4.5: Flow Chart of Finding the Optimal Solution for N-Stage Filter Design

From the "User Inputs" the total oversampling rate $D$ and preset filter stage $N$ can be obtained. The next step is to check the relationship between $D$ and $N$. Therefore, 3 judgement blocks has been set. If $D$ is a prime number, the filter stage $N$ can only be set as "1" due to the prime number can not be factorized. If $D$ is a non-prime number and cannot be factorized to $N$ factors, the filter stage $N$ can be set to "X" which is the total number of factors. And if $D$ can only be factorized to $N$ elements, according to the above discussion, the optimal computational cost design can be obtained by sorting the elements from large to small. However, in the most situations, the number of factors is larger than the preset filter stage $N$. Therefore, the adjustment and rearrangement for factors are needed in the "Reduce the number of elements" block. Actually, the nature of this block is the classic "$n$ choose $k$" mathematical problem.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad for \quad 0 \leqslant k \leqslant n \tag{4.1}$$

From the Equation 4.1, it can be obtained that $k$ is fixed when the filter stage has been defined and the total number of combinations is determined by the number of total factors $n$. Hence, in order to increase the efficiency of this approach, the minimum $n$ is required.

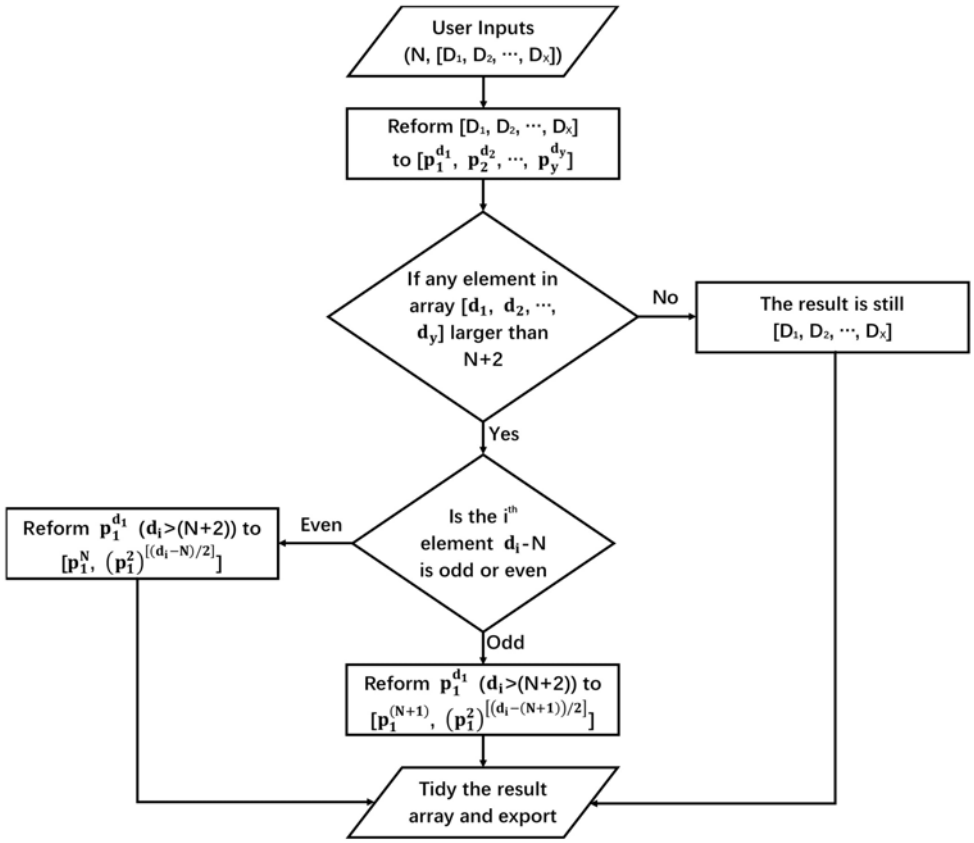



Figure 4.6: Flow Chart of Factor Reduction for N-Stage Filter Design

Therefore, the above flow chart Figure 4.6 has been designed to minimize the total elements $n$.

For 1-5000 total oversampling rate $D$, there are 4330 non-prime number and 670 prime numbers. And the characteristics of these non-prime numbers can be summarized as:

Table 4.2: Characteristic for 1-5000 non-prime numbers

| Factors | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sets | 1365 | 1273 | 832 | 452 | 224 | 104 | 47 | 22 | 7 | 3 | 1 |

For example, for a 3-stage filter design with 4096 ($2^{12}$) oversampling rate, $C_3^{12}$ is 220. However, if the above factor reduction has been done, the factors will become [2, 2, 2, 4, 4, 4, 8], and the $C_3^7$ will be reduced to 35 which is

around 16% reduction.

Based on the above discussions and examples, it can be known that 2-4 stage filter designs are relatively efficient. According to Table 4.2, it can be known that there are hundreds of sets which have more than 6 elements and most of them can be simplified due to the component element of these sets are repeated small prime number. Therefore, the efficiency of this approach can be increased significantly with the factor reduction approach.

For example, for $D < 5000$, 3 stage decomposition there are only 1692 (33.8%) number can be factorised in to multiple unique sets of 3 which will need to be put back into cost functions to evaluate the best one.

However, $C_k^n$ can only represent the complexity of this approach. Based on the result of above factor reduction process, the possible solution sets can be obtained by full permutation, rearrangement, calculation and matrix operation. The details of this approach are represented in the following Table 4.3.

Following this approach, for the typical design value $\Delta f = 0.18$, with $D < 5000$ of 2, 3, and 4 stage design, this approach provides around 85.4% average reduction when compared with exhaustive search in terms of the number of cost function tests, and around 65% computing time reduction. Computing time was averaged over 100 iterations using a standard Intel Core i7 based PC.

Table 4.3: Search the Optimal Distribution Solution

| | |
|---|---|
| Inputs ($D$, $N$, $\Delta f$ and $D_i$ sets) | $D$, $N$, $\Delta f$, $[D_1,D_2,\ldots,D_X]$ |
| | $\downarrow$ |
| Full permutation $D_i$ sets | $[D_1,D_2,\ldots,D_X]$<br>$\vdots$<br>$[D_X,\ldots,D_2,D_1]$ |
| | $\downarrow$ |
| Separate the matrix into two matrices<br>Matrix 1: the first N column<br>Matrix 2: the rest $(X-N)$ column<br>Repeat these two matrices $N^{(X-N)}$ times | $[D_1,\ldots,D_N]$ $[D_{N+1},\ldots,D_X]$<br>$\vdots$ $\vdots$<br>$[D_X,\ldots,D_{N+1}]$ $[D_N,\ldots,D_1]$ |
| | $\downarrow$ |
| Rearrange, permute and combine the elements in Matrix 2 and substitute into Matrix 1 | $[D_1 \times D_{N+1}\times\ldots\times D_X,\ldots,D_N]$<br>$\vdots$<br>$[D_X,\ldots,D_{N+1} \times D_N\times\ldots\times D_1]$ |
| | $\downarrow$ |
| Delete the duplicate row sorting and calculate $S$, $T$ or $P$ for different applications | $[D_1 \times D_{N+1}\times\ldots\times D_X,\ldots,D_N]\rightarrow S(1)$<br>$\vdots$<br>$[D_X,\ldots,D_{N+1} \times D_N\times\ldots\times D_1]\rightarrow S(Z)$ |
| | $\downarrow$ |
| Find out the optimal distribution and export the result | $[D_a,D_b,\ldots,D_x]\rightarrow S(minimum)$ |

Compared with mainstream approach, author proposed method can provide the directly usable results efficiently. Therefore, what does the integer decimation rate distribution of optimal computational or area cost look like?

Based on the above mentioned design conditions, the reasonable interval of oversampling rate $D$ is set at 1-5000, and the transition bandwidth $\Delta f$ is limited below 0.5. In this case, $D$ is highly composite value as $2^n$, therefore, the value of $D$ can be regarded as a array $[2^1\ 2^2\ \ldots\ 2^{11}\ 2^{12}]$, also the $\Delta f$ can be regarded as [0.01 0.02 0.03 ... 0.49 0.5] (the step of $\Delta f$ can be adjusted according to the requirement of accuracy). Thereby, a matrix or coordinate system can be constituted through these two arrays where the horizontal axis is $D$ and the vertical axis is $\Delta f$. At this time, parameters for optimal decimation rate calculation are prepared. Then, we only need to calculate the optimal decimation rate for each stages according to the above mentioned approach and coordinate system.

Hereafter, by sorting the results, the decimation rate of each stage can be constituted to different matrices, after adding these matrices as the Z axis to the coordinate system, the following figure 4.7 and 4.8 can be obtained

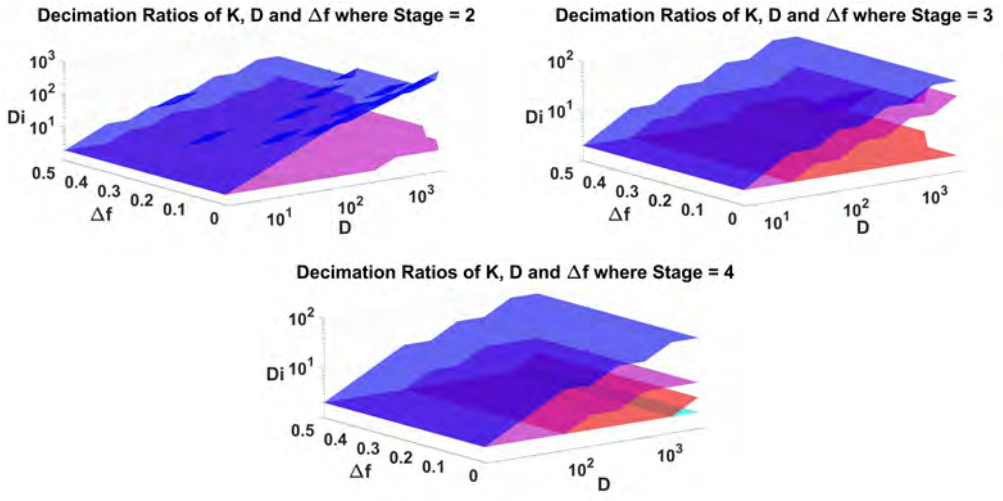



Figure 4.7: Integer Solution Sets Distribution for $D = 2^n$ (Computational)

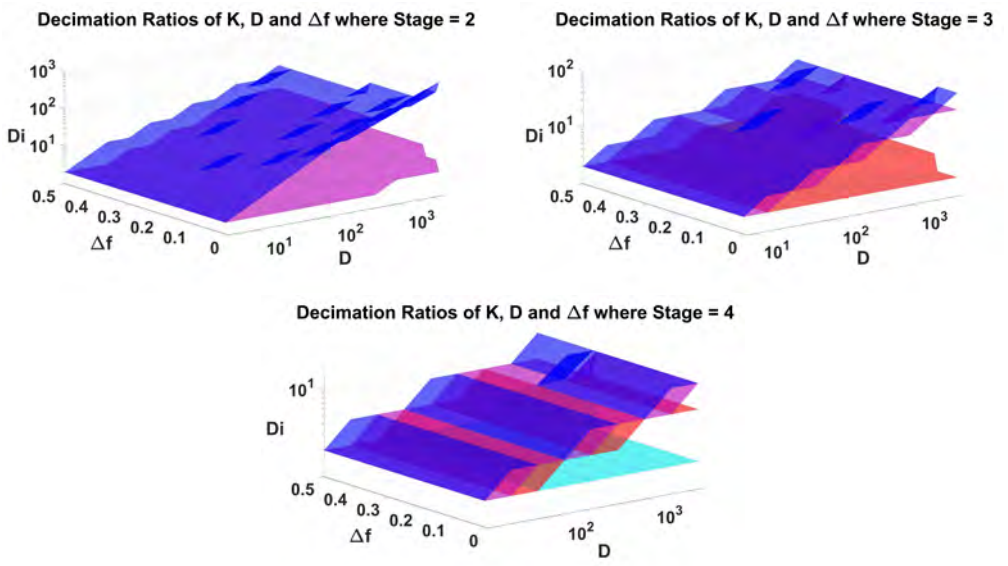



Figure 4.8: Integer Solution Sets Distribution for $D = 2^n$ (Area)

Base on previous research findings, as above Figure 4.7 and 4.4 demonstrated the optimal integer decimation or interpolation rate distribution only change a few times in the entire interval of reasonable transition bandwidth $\Delta f$. Also the above 3D matrix or coordinate system has advantage of easy to access. Therefore, an idea came up by the author which is make the optimal designs in reasonable design interval into a 3D table, so that the results can be found faster.

## 4.4 Knowledge Based Search and 3D Database

According to the above sections, it can be found that the decimation or interpolation rate distribution for optimal computational cost or area cost has certain disciplines.

1. Decimation or interpolation rate set $D_i$ is always in descending order for multistage decimation and in ascending order for multistage interpolation.

2. $\Delta f$ is related to the width of transition band. The variation of $\Delta f$ changes the order of the filter but not the sampling rate changing factor of each stage.

In addition, the variation of $\Delta f$ does not cause much change in the optimal integer values. For the same example of $D < 5000$ and 3-stage case, within the 1692 cases that have possible multiple solutions, in 994 (59%) cases for computational cost optimization and 1167 (69%) cases for memory cost optimization, the optimal solution sets change only once or twice over the (0 to 0.5) $\Delta f$ region.

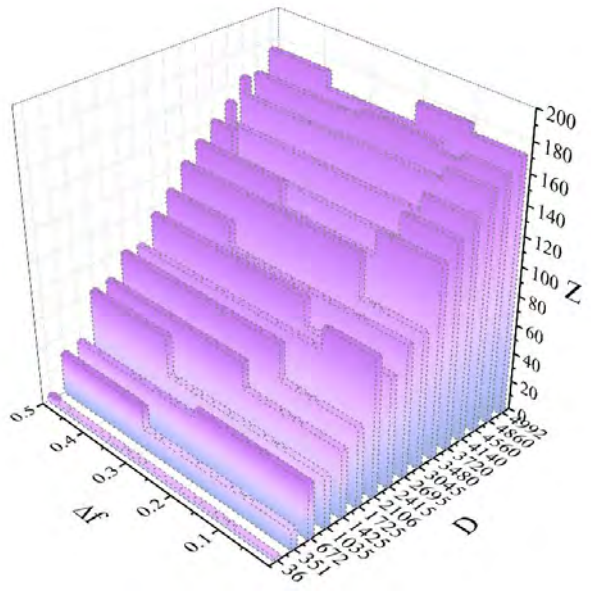



Figure 4.9: Changing of Optimal Value against Δf for Composite Number D

The above Figure 4.9 shows the changing of the optimal integer valued solution sets against $\Delta f$ for some highly composite number $D$. The $Z$ axis value is calculated by the formula 4.2 below. The height of $Z$ represent an unique solution set.

$$Z = \frac{D}{40} + \sqrt{\frac{1}{K}\sum_{i=1}^{K}(D_i - (\frac{1}{K}\sum_{i=1}^{K}D_i))^2} \tag{4.2}$$

Hence, based on the above disciplines and findings, the lookup tables optimal solution set finding method has been proposed for 2, 3, 4 stage filter design. In order to store these sets, tables have been created. Figure 4.9 shows the optimal integer solution sets can be same values for considerable range of design specification $\Delta f$. Both the optimal area cost and the computational

cost databases share the similar properties. The experiments show that for 2-4 stage designs, with $1 < D < 5000$ and $0 < \Delta f < 0.5$, the total number of optimal computational cost sets is 15783, and the total number of optimal memory usage sets is 15785. Therefore, we can create databases to store these optimal solutions with the critical values of $\Delta f$ that cause the changes of optimal solution sets. The structure of database can be depicted in Figure 4.10.

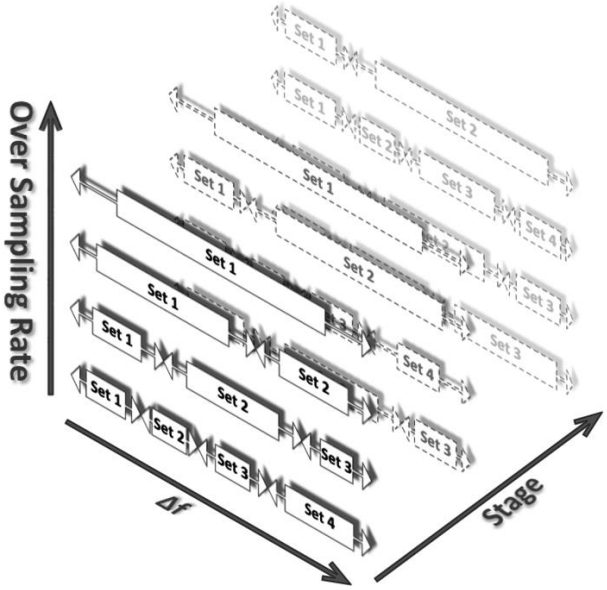



Figure 4.10: Sketched 3-dimensional storage matrix for Database

Where stage is the number of filter stage, $\Delta f$ is the transition bandwidth, and oversampling rate is the decimation or interpolation rate for each stage.

## 4.5 Summary

In this chapter, author reviewed the mainstream optimal decimation or interpolation rate calculation methods first, and point out two problems that can be further improved. One is that the calculation is complicated, and the other one is the calculated results cannot be used directly due to the non-integer solutions. Then, analysing and discussing current optimal decimation or interpolation rate calculation or selection approach, author finds out the key parameters affect the optimal filter design and the direction of

solving non-integer solution problems. Thereby, author proposed a new optimal decimation or interpolation rate selection approach based on Parks and McClellan (1972) proposed classical individual optimal FIR filter design. As a major contribution of this thesis, this approach can reduce the calculation complexity significantly while improving the search efficiency greatly. Finally, a knowledge based lookup tables search method has been proposed which can further improve the search efficiency.

# Chapter 5

# Research and Derivation for Theoretical Latency of Multi-stage Multi-rate Filter

## 5.1 Background

For digital filters used in high resolution audio conversions such as the decimation or interpolation filters in professional $\Sigma$-$\Delta$ ADC/DAC and Sample Rate Conversion (SRC), the performance of the filters needs to satisfy specific resolution and Signal to Noise Ratio (SNR) requirements, for example, greater than 24 bits bit resolution and greater than 120dB SNR. These requirements normally result in high order filters that have significant overall latency.

The associated filters are usually designed to be multistage. Each stage utilises half-band or N-band linear phase filter whereby it is possible to further reduce the computational and implementation cost. However, these structures often worsen the overall latency and deteriorate the responsiveness of the system. For high performance anti-aliasing FIR filter, the transition band can be less than 0.01 (normalised frequency). A single filter realisation will result in very high number of orders, usually over thousands which is very difficult to realize. Although multi-stage filter design method reduce the filter order significantly and makes design achievable. However, for exchange this method will worsen the latency performance.

To author's best knowledge, there is no literature provides equation of latency estimation. Since the aim of this thesis is to carry out a comprehensive eval-

uation as well as propose a framework for the digital filter design with both low latency and low cost for Σ-Δ modulation based ADC/DAC, therefore, it is necessary to construct a quantifiable latency evaluation and estimation method. Otherwise, latency and other filter costs will not be able to correlate which means the latency performance of design cannot be evaluated together with other filter performances in the proposed framework. Hence, it is very important to derive the latency estimation equations. Inspired by the estimation equation of computational cost proposed by Crochiere and Rabiner (1975) the following derivation are made by author.

## 5.2 Decimation Rate Distribution for Low Latency Filter Design

The latency of linear phase FIR filter with symmetric coefficients is equal to half of filter order (Tan and Jiang, 2018). Also, several filter order estimation methods are available from Equation 2.11 to 2.17. Therefore, it is possible to derive the equations of latency estimation

### 5.2.1 Equation Derivation

Assume Figure 2.12 shows the sketched block diagram of multi-stage multirate filter. As mentioned above, for linear FIR filter, the groupdelay can be estimated as half of the filter order. Therefore, the filter's groupdelay $GD$ in above Figure 2.12 can be calculated as:

$$GD_T = \sum_{i=1}^{K} GD_i \tag{5.1}$$

$GD_i$ in samples can be considered as half filter order length $\frac{1}{2}N_i$, $GD_i$ in seconds can be calculated using $GD_i$ (samples) and current sampling frequency which listed below.

$$GD_i = \begin{cases} \frac{1}{2}\frac{N_i}{f_{r(i-1)}} & \text{Decimation} \\ \frac{1}{2}\frac{N_i}{f_{ri}} & \text{Interpolation} \end{cases} \tag{5.2}$$

Substituting Equation 2.13 and 2.15 into Equation 5.2

$$GD_i = \begin{cases} \frac{1}{2f_{r(i-1)}} \frac{D_\infty(\delta_1,\delta_2) L_i f_{r(i-1)}}{f_{ri}-f_s-f_p} & \text{Decimation} \\ \frac{1}{2f_{ri}} \frac{D_\infty(\delta_1,\delta_2) M_i f_{ri}}{f_{ri}-f_s-f_p} & \text{Interpolation} \end{cases} \tag{5.3}$$

After simplification, the Equation 5.3 will become:

$$GD_i = \begin{cases} \frac{D_\infty(\delta_1,\delta_2)}{2} \frac{L_i}{f_{ri}-f_s-f_p} & \text{Decimation} \\ \frac{D_\infty(\delta_1,\delta_2)}{2} \frac{M_i}{f_{ri}-f_s-f_p} & \text{Interpolation} \end{cases} \tag{5.4}$$

$L_i$ and $M_i$ will be 1 in this case. Hence, Equation 5.4 can be written as:

$$GD_i = \frac{D_\infty(\delta_1, \delta_2)}{2(f_{ri} - f_s - f_p)} \tag{5.5}$$

Substituting Equation 5.5 into Equation 5.1 we have:

$$GD_T = \sum_{i=1}^{K} \frac{D_\infty(\delta_1, \delta_2)}{2(f_{ri} - f_s - f_p)} \tag{5.6}$$

Extract $\frac{D_\infty(\delta_1,\delta_2)}{2}$ from Equation 5.6

$$GD_T = \frac{D_\infty(\delta_1, \delta_2)}{2} \sum_{i=1}^{K} \frac{1}{(f_{ri} - f_s - f_p)} \tag{5.7}$$

If we extract the $K$-stage (last stage for decimation, first stage for interpolation), it can be obtained:

$$GD_T = \frac{D_\infty(\delta_1, \delta_2)}{2} (\frac{1}{f_s - f_p} + \sum_{i=1}^{K-1} \frac{1}{f_{ri} - f_s - f_p}) \tag{5.8}$$

From Figure 2.12c and Figure 2.12d, we can assume that stopband edge of last stage is half of sampling frequency which can be written as Equation 5.9, normalised transition bandwidth $\Delta f$ is regarded as Equation 5.10 shows. Also, the sampling frequency of current stage can be calculated by decimation rate and sampling frequency of previous stage as Equation 5.11 demonstrated.

$$f_s = \frac{f_{r0}}{2D} \tag{5.9}$$

$$\Delta f = \frac{f_s - f_p}{f_s} \tag{5.10}$$

$$f_{ri} = \frac{f_{r(i-1)}}{D_i} \tag{5.11}$$

By deform the Equation 5.10 it can be obtained:

$$f_p = f_s - f_s \Delta f \tag{5.12}$$

After substituting Equation 5.9 into 5.12, the passband edge $f_p$ can be regarded as:

$$f_p = \frac{f_{r0}}{2D}(1 - \Delta f) \tag{5.13}$$

Due to the Equation 5.8 and Equation 5.9 to 5.13 are universal equations for both decimation and interpolation, The derivation for interpolation equations is not necessary. There is only one thing need to be noticed, that is the filter stage order is reverse order for interpolation filter.

When we bring Equation 5.9 and 5.13 into Equation 5.8, the overall groupdelay of filter can be obtained:

$$\begin{aligned} GD_T = & \frac{D_\infty(\delta_1, \delta_2)}{2}\Big(\frac{1}{\frac{f_{r0}}{2D} - \frac{f_{r0}}{2D}(1 - \Delta f)} \\ & + \sum_{i=1}^{K-1} \frac{1}{\frac{f_{r0}}{\prod_{j=1}^{i} D_j} - \frac{f_{r0}}{2D} - \frac{f_{r0}}{2D}(1 - \Delta f)}\Big) \end{aligned} \tag{5.14}$$

After simplification, the following Equation 5.15 can be obtained:

$$GD_T = \frac{D_\infty(\frac{\delta_p}{K}, \delta_s)}{2}\Big(\frac{1}{f_s - f_p} + \sum_{i=1}^{K-1} \frac{1}{2D\frac{f_s}{\prod_{j=1}^{i} D_j} - f_s - f_p}\Big) \tag{5.15}$$

Which can be written as:

$$GD_T = \frac{1}{2} D_\infty(\frac{\delta_p}{K}, \delta_s) P \tag{5.16}$$

Where:

$$P = (f_s - f_p)^{-1} + \sum_{i=1}^{K-1} (2D\frac{f_s}{\prod_{j=1}^{i} D_j} - f_s - f_p)^{-1} \tag{5.17}$$

If the convenience of programming is considered, the variables in the estimation function need to be unified, so that a set of variables can be used to estimate different filter costs such as filter's computational cost, area cost and latency.

Hence, the equation can also be written as Equation 5.18 based on the Equation 5.14:

$$GD_T = \frac{D_\infty(\delta_1, \delta_2)}{2}\left(\frac{2D}{f_{r0}(1-(1-\Delta f))} + \sum_{i=1}^{K-1} \frac{1}{\frac{f_{r0}}{\prod_{j=1}^{i} D_j} - \frac{f_{r0}}{2D}(1+(1-\Delta f))}\right) \tag{5.18}$$

By simplifying the Equation 5.18, the Final groupdelay estimation equation can be obtained as:

$$GD_T = \frac{D_\infty(\delta_1, \delta_2)D}{f_{r0}}\left(\frac{1}{\Delta f} + \sum_{i=1}^{K-1} \frac{1}{\frac{2D}{\prod_{j=1}^{i} D_j} - 2 + \Delta f}\right) \tag{5.19}$$

Where:
$\delta_1$ is the tolerance in the magnitude response in the passband.
$\delta_2$ is the tolerance in the magnitude response in the stopband.
$D$ is overall decimation or interpolation rate.
$f_{r0}$ is the initial input sampling frequency.
$\Delta f$ is the transition bandwidth normalized stopband edge.
$K$ is the number of filter stage.
$D_j$ is decimation or interpolation rate of current stage.

Along with the quantized latency estimation equation has been derived, further analysis of equations can help us find the key factors affect the latency performance and provide guidance for optimal latency filter design.

From Equation 5.19, for a certain design, passband ripple $\delta_1$, stopband attenuation $\delta_2$, overall oversampling rate $D$, filter stage $K$ and initial input sampling frequency $f_{r0}$ are given. Therefore, these design parameters can be traded as constant. Thus, the only two variables affect the total groupdelay are transition bandwidth $\Delta f$ and decimation or interpolation rate distribution $D_j$. Next, author will analyse the influence of decimation or interpolation rate distribution on the latency performance in detail.

### 5.2.2 Analysis the Properties of Equation

Although the total order of the filter can be reduced significantly using multistage filter design method, the filter latency will increased because the later stages have lower input sampling frequencies. thereby, the latency caused by longer sampling period at later stage.

This can be proved below. The Equation 5.16 and 5.17 can be rewritten into the summation of two parts A and B.

$$GD_T = A + B \tag{5.20}$$

Where A is defined as:

$$A = \frac{1}{2} D_\infty(\frac{\delta_p}{K}, \delta_s)(f_s - f_p)^{-1} \tag{5.21}$$

And B is defined as:

$$B = \frac{1}{2} D_\infty(\frac{\delta p}{K}, \delta s) \sum_{i=1}^{K-1} (2D \frac{f_s}{\prod_{j=1}^{i} D_j} - f_s - f_p)^{-1} \tag{5.22}$$

When the filter is single stage, part $A$ is the only one left in $GD_T$

$$\begin{aligned} GD_T = A &= \frac{1}{2} D_\infty(\frac{\delta_p}{K}, \delta_s)(f_s - f_p)^{-1} \\ &= \frac{1}{2} D_\infty(\frac{\delta_p}{K}, \delta_s) \frac{1}{\Delta F f_{r0}} \end{aligned} \tag{5.23}$$

From above Equation 5.23, it can be obtained that the latency of single stage filter design does not depend on the overall decimation rate $D$ and when the latency is equivalent to the half of number of order when filter order $N \gg 1$.

According to 'TABLE I' in Crochiere and Rabiner (1975), it can be found that for the certain $\delta p$ and $\delta s$ the $D_\infty(\delta_p/K, \delta_s)$ increases when filter stage $K$ increasing. And for the remaining part B, the maximum value of $\prod_{j=1}^{i} D_j$ is the overall decimation rate $D$, therefore, it can be proved that the running sum and part B are always greater than 0. The overall $GD_T$ is increasing. Hence, when $K > 1$, the latency increases when the filter stage $K$ increases.

If we regard the decimation rate $D$ as a composite number which can be factorised into the number $K$ of $D_i$. The latency of the filter in Equation 5.5 can be further written as:

$$GD_i = \frac{C}{f_{ri} - E} \tag{5.24}$$

$$C = \frac{1}{2} D_\infty(\frac{\delta_p}{K}, \delta_s) \tag{5.25}$$

$$E = f_s + f_p \tag{5.26}$$

Where both C and E is greater than 0 and $f_{ri}$ always greater than D. The latency is inversely proportional to the input sampling frequency.

Therefore, for the given number of stage $K$, it would be have maximum $f_{ri}$ to have minimum $GD_i$. The maximum value of $f_{ri}$ is $f_{r0}$. The $D_i$ should be '1' for that stage. Hence, the minimum latency should be the form of [1, 1, ..., $f_{r0}/f_{rK}$], which is equivalent to single stage filter design.

For the multistage structure where $K > 1$ and $D > 1$, let's take the partial derivative of Equation 5.15 with respect to $f_{ri}$:

$$\frac{\partial GD_i}{\partial f_{ri}} = \frac{-C}{(f_{ri} - E)^2} \tag{5.27}$$

For $f_{ri} > (f_s + f_p)$, the partial derivative is regarded as gradient: $\nabla GD_i$. If we regard the multistage decimation process as adding number of frequencies together to reach the highest input frequency $f_{r0}$ from $f_{rK}$. This basically shows when a small stage of frequency increasing *df* is needed to have lower increment of $GD_i$, the *df* should be put toward lower end of $f_{r0}$. For example, in theory when $K > 1$ and $D > 1$, the form of decimation to have smaller total latency will be [2, 2, ..., remains]. Hence, the properties of equation can be summarized as:

1. The overall latency increases as the number of stages increase.
2. For given $K$-stage, the overall latency tends to be smaller when distributing larger decimation factor $D$ towards the later stage.

## 5.3 Discussion of Low Latency Filter Design

In order to verify the above findings and analyse the latency characteristic of the filter better, a verification test has been made. In this experiment, several filters are designed in the commonly used design requirements for high-level audio, and the designed filter performances are summarized and analysed to verify whether it is consistent with the above conclusions.

Also, during this verification test, author found that the optimal filter design method can be further improved due to the non-optimal allocation of pass-band ripple and underutilized of transition bandwidth.

### 5.3.1 Verification Test

As mentioned above, this verification test will design several filters which follows the commonly used high-level audio filter design requirements. According to Hawksford (1994), Mitra (2011), Datasheet of PCM4220 (2009), PCM1840 (2019) and Zhu et al. (2016), the test conditions have been designed as following Table 5.1 listed.

Table 5.1: Test Conditions

| | |
|---|---|
| Filter Stage ($K$) | 1-3 |
| Oversampling Rate ($D$) | 64 |
| Sampling Frequency ($Fs$) | 48kHz |
| Passband Edge ($f_p$) | 20kHz |
| Stopband Edge ($f_s$) | 24kHz |
| Passband Ripple ($A_p$) | 0.0001dB |
| Stopband Attenuation ($A_{st}$) | 120dB |

Because the single stage filter design does not involve the parameter calculation of each stages, the design is relatively simple, however, due to the stringent design requirements the design takes nearly twenty minutes. And the filter costs are listed in below Table 5.2.

Table 5.2: Costs for Single Stage Filter Design

| Single Stage | |
|---|---|
| $D_i$ Sets | [64] |
| $NO.M$ | 5248 |
| $NO.A$ | 5247 |
| $MpIS$ | 82 |
| $ApIS$ | 81.9844 |
| $GD_{Sample}$ | 2623.5 |

Where:
$D_i$ Sets is distribution of decimation rate.
$NO.M$ is total number of required multipliers.
$NO.A$ is total number of required adders.
$MpIS$ is total number of multiplications required for each input sample.
$ApIS$ is total number of additions required for each input sample.
$GD_{Sample}$ is groupdelay in samples.

According to the listed costs, it can be found that this filter has around 5000 order which can not be implemented on the hardware.

Hereafter, the two stage filter design is a little bit more complicated than single stage filter design. Because all possible distribution of decimation rate need to be considered. Fortunately, all possible distribution of decimation rate can be calculated out using the optimal decimation or interpolation rate selection approach which is proposed in Section 4.3. Then, the next step is to calculate out design parameters for each stage, the overall passband ripple needs to be reassigned to each stage, stopband edge of each stage need to be calculated according to the equation $f_{ri} - f_s$ listed in Figure 2.12, also, the input and output sampling frequency of each stage are necessary as well. While all of above mentioned parameters are calculated, filters can be designed and the results are shown in Table 5.3.

Table 5.3: Decimation Rate Distribution and Costs for 2-Stage Filter

| 2-Stage | | | | | |
|---|---|---|---|---|---|
| $D_i$ Sets | [2,32] | [4,16] | [8,8] | [16,4] | [32,2] |
| $NO.M$ | 2710 | 1394 | 747 | 492 | 589 |
| $NO.A$ | 2708 | 1392 | 745 | 490 | 587 |
| $MpIS$ | 47.6719 | 28.5781 | 18.6719 | 14.5781 | 15.7031 |
| $ApIS$ | 47.1563 | 28.3125 | 18.5313 | 14.5 | 15.6563 |
| $GD_{Sample}$ | 2703 | 2742 | 2759.5 | 2825 | 2959.5 |

From the above results, it can be found that costs except latency are reduced significantly comparing with single stage filter design. Also, the costs difference between different decimation rate distribution are huge. Simultaneously, these results also confirm the previous inferences mentioned in Chapter 4 and Section 5.2.2. With the same arguments, the costs of 3-stage filter design can be calculated.

Table 5.4: Decimation Rate Distribution and Costs for 3-Stage Filter

| 3-Stage | | | | |
|---|---|---|---|---|
| $D_i$ Sets | [2,2,16] | [2,4,8] | [2,8,4] | [2,16,2] |
| $NO.M$ | 1412 | 737 | 433 | 398 |
| $NO.A$ | 1409 | 734 | 430 | 395 |
| $MpIS$ | 29.9688 | 20.2344 | 15.5625 | 14.8594 |
| $ApIS$ | 29.2031 | 19.5938 | 14.9844 | 14.3125 |
| $GD_{Sample}$ | 2793 | 2811 | 2854 | 3000 |

| 3-Stage | | | | | |
|---|---|---|---|---|---|
| $D_i$ Sets | [4,2,8] | [4,8,2] | [8,2,4] | [8,4,2] | [16,2,2] |
| $NO.M$ | 737 | 310 | 429 | 293 | 349 |
| $NO.A$ | 734 | 307 | 426 | 290 | 346 |
| $MpIS$ | 19.7344 | 13.2969 | 14.5625 | 12.5156 | 12.8281 |
| $ApIS$ | 19.3438 | 13 | 14.3594 | 12.3438 | 12.7188 |
| $GD_{Sample}$ | 2814 | 3008 | 2868 | 3024 | 3050 |

From the above tables, it can be found that the optimal delay filter design is the single stage filter design which is 2623.5 samples (around 0.854ms). And the largest delay is 3050 samples (around 0.9928ms). Therefore, it can be obtained that in this test, the different between the optimal latency design and worst latency design is 16.26% which is 426.5 samples (around 0.139ms). Although this is a big difference, this gap can be tolerated in most cases.

But if we look into the other costs, it can be found that the single stage filter design and some of 2 or 3 stage filter designs are hard to be realised due to the exaggerated area cost and computational cost. The different between optimal area cost and worst area cost designs are around 92.42% (4955) multipliers and around 94.47% (4957) adders. The difference between optimal computational cost and worst computational cost designs are around 84.75% (69.4844) multiplications and around 84.94% (69.6406) additions for each input sample. However, some 2 or 3 stage filter designs have balanced performance and are easy to realise. Therefore, from the above tests it can be obtained that the consideration priority of low latency cost efficient filter design is area cost and computational cost. The latency problem can only be considered based on the realisable cost efficient filter designs.

In this section, the verification test has been made by author. Base on the above test results and discussion, the correctness of the several important derivations and analyses mentioned above can be inferred which are:

- Fewer filter stages leads to more computational or area cost.
- Distribution of optimal decimation rate is arranged from large to small.
- The overall latency increases as the number of stages increase.
- For given $K$-stage, the overall latency tends to be smaller when distributing larger decimation factor $D$ towards the later stage.

### 5.3.2 Other Considerations for Groupdelay Optimization

From the previous derivation and discussion, it can be known that the decimation rate distribution of optimal computational cost is from large to small and the optimal latency is from small to large. Therefore, it is impossible to find out a distribution to have both optimal cost and latency. Hence, the problem becomes to find out another way to realize the cost efficient filter design with minimized latency.

From the Equation 5.19, it can be obtained that the key factors affect the groupdelay is $\Delta f$ when the sampling frequency and decimation rate distribution were set. In other words, the key factors affect the groupdelay of the multi-stage multi-rate filter design is the transition bandwidth.

Also, author found another design parameter may affect the latency and other performances of multi-stage filter which is passband ripple. Although it is a given design requirement, it has to be reassigned to each stage because the design requirement is the limitation for overall passband ripple and current optimal filter design approach always reassign the requirement to each stage evenly. Hence, author found here are two problems: One is the current optimal design approach didn't fully utilize design requirement. The other one is costs of each stage are different, therefore, reassign passband ripple evenly is unreasonable.

However, it is still unknown how much the above 2 factors $D_\infty(\delta_1, \delta_2)$ and $\Delta f$ will affect the groupdelay of the designed filter. Therefore, a commonly used multi-stage multi-rate filter design may help us to find out the future optimization direction. The design specifications are set to the following values:

Table 5.5: Filter Design Specifications of Example

| | Cascaded Filter | Stage 1 | Stage 2 | Stage 3 |
|---|---|---|---|---|
| $F_s$(in) | 3072kHz | 3072kHz | 384kHz | 96kHz |
| $F_s$(out) | 48kHz | 384kHz | 96kHz | 48kHz |
| $D$ | [8 4 2] | 8 | 4 | 2 |
| $F_p$ | 20kHz | 20kHz | 20kHz | 20kHz |
| $F_s$ | 24kHz | 360kHz | 72kHz | 24kHz |
| $A_p$ | 0.0001dB | 0.0001dB/3 | 0.0001dB/3 | 0.0001dB/3 |
| $A_{st}$ | 120dB | 120dB | 120dB | 120dB |

Where: $F_s$(in) is input sampling frequency, $F_s$(out) is output sampling frequency, $D$ is decimation rate, $F_p$ is passband edge, $F_s$ is stopband edge, $A_p$ is passband ripple, and $A_{st}$ is stopband attenuation.

As the above Table 5.5 exposed, the multi-stage multi-rate filter design based on Crochiere and Rabiner (1975) has the fixed passband ripple ($\delta_1$) and stopband attenuation ($\delta_2$) for each stage, and the stopband edge (related to $\Delta f$) has the fixed calculation method. And designed filter's magnitude response curve and related parameters are shown in following figure and table.

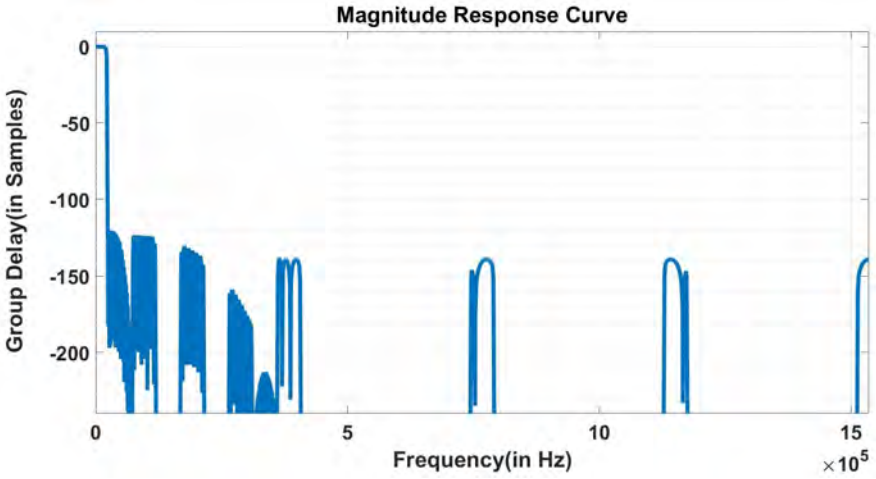



Figure 5.1: Magnitude Response Curve of Designed Filter

Table 5.6: Related Parameters of Designed Filter

| | Cascaded Filter | Stage 1 | Stage 2 | Stage 3 |
|---|---|---|---|---|
| $NO.A$ | 290 | 64 | 52 | 174 |
| $NO.M$ | 293 | 65 | 53 | 175 |
| $ApIS$ | 12.3438 | 8 | 13 | 87 |
| $MpIS$ | 12.5156 | 8.125 | 13.25 | 87.5 |
| $GD_{Sample}$ | 3024 | 32 | 26 | 87 |
| $GD_{ms}$* | 0.984375 | 0.01042 | 0.0677083 | 0.90625 |

Where $GD_{ms}$ is groupdelay in millisecond, * means approximately equal to.

As the Table 5.6 shown, the designed filter's costs and groupdelay are mainly caused by the last stage. Therefore, the cascaded filter's costs and groupdelay can be reduced effectively if the costs and groupdelay of last stage can

be reduced. According to Chapter 4 and above discussion, the costs and groupdelay can be changed by adjusting the filter's passband ripple, transition bandwidth and stopband attenuation.

Firstly, looking into the passband ripple of the filter, the magnitude response curves can be obtained as:

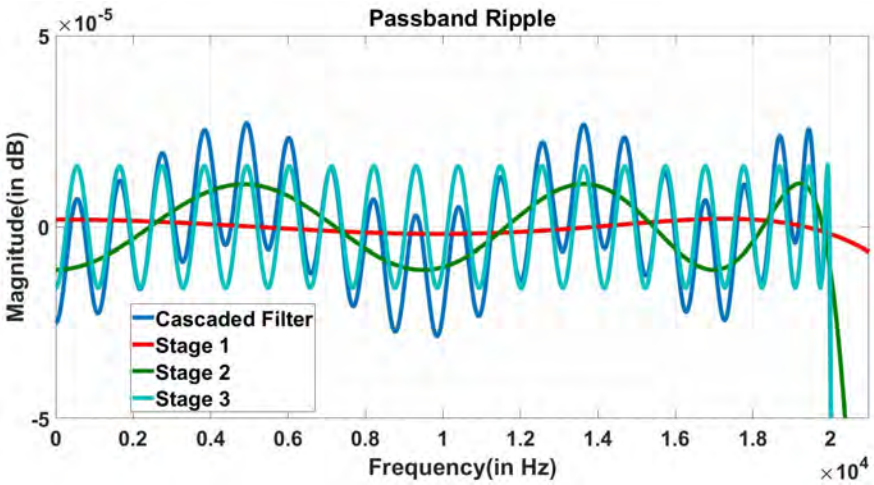



Figure 5.2: Passband Ripple of Designed Filter and Each Stage

The passband ripple of designed cascaded filter is around ±0.00003dB. However the design requirement is 0.0001dB which is around ±0.00005dB. Hence, the designed cascaded filter didn't make full use of the design requirement. In other words, the passband ripple can be optimized with specific strategies.

Secondly, after zooming into the transition bandwidth and stopband attenuation, the following magnitude response curves can be seen:

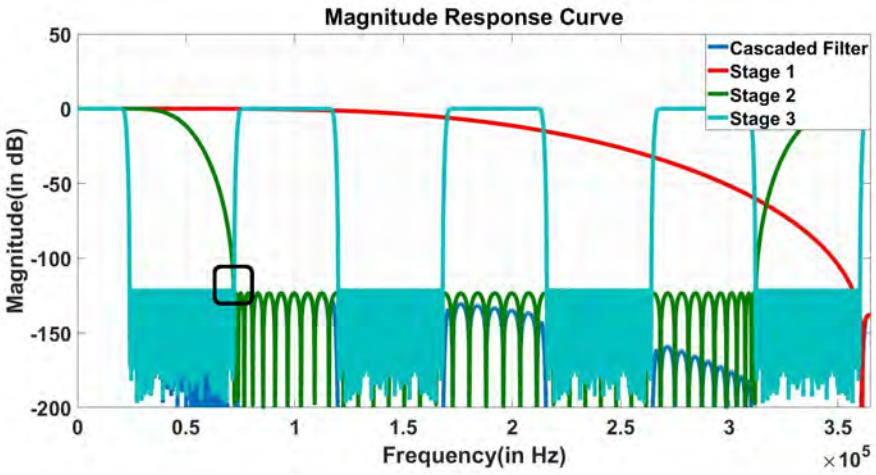



Figure 5.3: Stopband Edge and Attenuation of Designed Filters

Zooming into the black frame in Figure 5.3, the space can be found to increase the transition bandwidth which is shown in the purple frame of Figure 5.3.

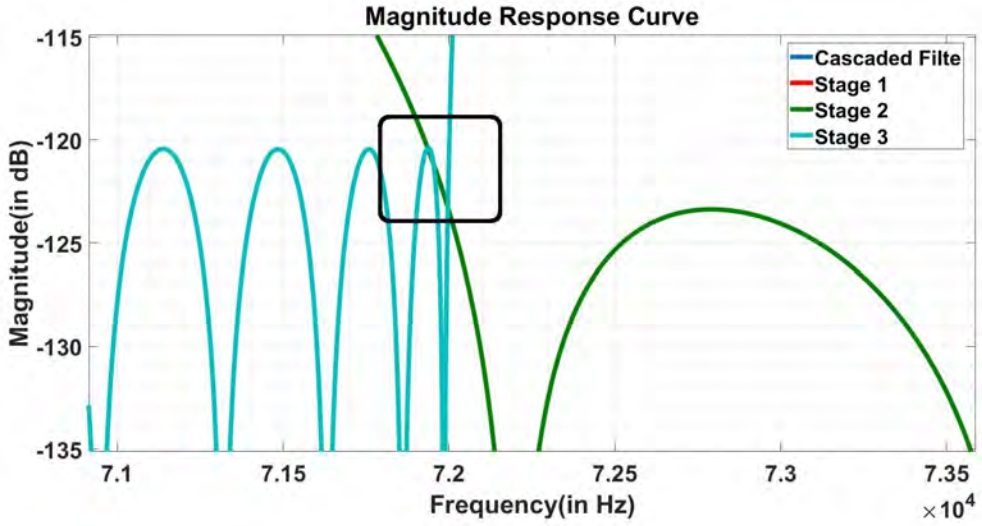



Figure 5.4: Details of Filter's Stopband Edge and Attenuation

As the above Figure 5.3 and 5.3 exposed, the designed filter's transition bandwidth still has potential to increase. Which means the length of filter coefficients or costs can be further reduced.

Hence, From this 5.3.2 section, it can be obtained that even the design is based on the theoretical optimal filter design method, the designed filter still

has the potential to be optimized. The optimization direction could be:

1. Maximize the usage of the passband ripple design requirement.

   This one is ensure the designed filter's passband ripple fully utilizes the design requirement.

2. Rearrange the scale of each stage's passband ripple.

   Rearrange the passband ripple for each stage to make sure the higher cost filter stage has larger passband ripple. This is because according to Equation 2.22, 2.27 and 5.19, the filter has larger passband ripple can achieve the lower costs or latency.

3. Widen the transition bandwidth as much as possible within the design requirements.

   According to Equation 2.22, 2.27 and 5.19, the filter has wider transition bandwidth will cause the lower costs or latency. Furthermore, if the stopband attenuation can be adjusted synchronously, the transition bandwidth can be further increased.

Although, the above-mentioned theories are easy to understand, the specific operation is not that easy. As the Figure 5.2 shown the magnitude response curve of the filter is not a straight line, so how to control the superimposed curves to meet the design requirements and find out the optimal design is a great challenge.

## 5.4 Summary

In this chapter, author reviewed the background first and pointed out there is no theoretical relationship between latency and multi-stage design parameters of filter. Therefore, in order to achieve the aim of this thesis, a quantifiable latency evaluation and estimation approach is needed. Thereby, author derived the latency estimation equation inspired by assumption and filter order estimation equations proposed by Crochiere and Rabiner (1975). As a major contribution of this thesis, these estimation equations built the link between latency and other filter costs which make the aim of this thesis achievable.

Hereafter, in the analysis of latency estimation equation, several important features are found. The first one is the overall latency increases as the

number of stages increase. The second one is for given $K$-stage, the overall latency tends to smaller when distributing larger decimation factor $D$ towards the later stage. These two characteristics determine that the optimal latency design and the optimal computational or area cost have contradictory requirements. Finally, author found that the key factors affect the overall latency performance are transition bandwidth $\Delta f$ and distribution of decimation or interpolation rate $D_i$

Then, in the section 5.3, a verification test has been made by author which corroborates the previous inference. Also, base on the third feature of latency estimation equation mentioned above, author point out that the current optimal filter design approach still has room to improve. One optimization direction is passband ripple, the other one is transition bandwidth.

# Chapter 6

# Low Latency Cost Efficient Filter Design Using Optimization Method

## 6.1 Limitation of Current Filter Design Method

From the above discussion, it is known that the current filter design method has the following limitations:

1. In theory, the distribution of decimation rate for optimal computational cost and optimal latency filter design has contradictory requirements. Therefore it is hard to design a filter with optimal computational cost and optimal latency. Therefore, the current filter design methods normally set the low latency consideration has lower priority than the area cost and computational cost when the latency of filter design meets the design requirement.

2. Current filter design methods are based on the ideal model of filter which has flat passband and stopband. But in reality, both passband and stopband have ripples, and these ripples may help us to improve the filter design method and make filter's transition band wider.

3. All of the optimal filter design methods are based on the optimal FIR design for example: Equiripple filter. But in some situations, other type of filter (for example: half-band filter) may achieve better performance.

The first limitation has been discussed in the above Chapter 5. The second and third limitations will be discussed below with the examples.

According to Chapter 4, the optimal filter decimation rate set can be obtained using the look-up table method. Therefore, the remaining optimization work is focused on the filter design. Based on the above Section 5.3.2, the optimization directions can be summarized as:

1. Use half-band instead of equiripple filter design (decimation rate is 2).
2. Adjust the passband ripple.
3. Widen the transition bandwidth.

As the example, if we set the filter design requirements to the common high-level $\Sigma$-$\Delta$ Modulation based audio ADC's design requirements which are:

Table 6.1: Filter Design Requirements

| | |
|---|---|
| Filter Stage ($K$) | 3 |
| Oversampling Rate ($D$) | 64 |
| Sampling Frequency ($Fs$) | 48kHz |
| Passband Edge ($f_p$) | 20kHz |
| Stopband Edge ($f_s$) | 24kHz |
| Passband Ripple ($A_p$) | 0.0001dB |
| Stopband Attenuation ($A_{st}$) | 120dB |

From the look-up table method in Chapter 4, the cost efficient filter design's decimation rate should be: [8 4 2]. However, because of the half-band filter's characteristics are not suitable for the tiny transition bandwidth filter design, the half-band structure filter can not be implemented to the last stage. Thus, within the possible low latency cost efficient filter designs listed by look-up table method, [8,2,4] becomes the most feasible option. Hence, the following example will take [8 4 2] and [8 2 4] as the filter's decimation rate for each stage.

According to Crochiere and Rabiner 1975 (1975), the passband ripple of each stage should split the design requirement equally, passband edge of each stage should be consistent with the filter design requirement. With the same argument, the stopband attenuation should be the same value of filter design requirement as well. The sampling frequency and stopband edge of each stage could be calculated based on the Figure 2.12 in Chapter 2. Therefore,

the design parameters for each stage can be calculated as:

Table 6.2: Filter Design Requirements for each stage

| | [8 4 2] | | | [8 2 4] | | |
|---|---|---|---|---|---|---|
| | Stage-1 | Stage-2 | Stage-3 | Stage-1 | Stage-2 | Stage-3 |
| $D_i$ | 8 | 4 | 2 | 8 | 2 | 4 |
| $F_s$ | 3072kHz | 384kHz | 96kHz | 3072kHz | 384kHz | 192kHz |
| $f_{st}$ | 360kHz | 72kHz | 24kHz | 360kHz | 168kHz | 24kHz |
| $f_p$ | 20kHz | | | 20kHz | | |
| $A_p$ | 0.0001dB/3 | | | 0.0001dB/3 | | |
| $A_{st}$ | 120dB | | | 120dB | | |

Where:
$D_i$ is oversampling rate.
$F_s$ is input sampling frequency.
$f_{st}$ is stopband edge of current stage.
$f_p$ is passband edge.
$A_p$ is passband ripple.
$A_{st}$ is stopband attenuation.

## 6.2 Cost Efficient Filter Design Using Half-band Filter

Before investigating into the influence of the half-band filter on the filter's costs and groupdelay, the characteristics of halfband filter should be figure out. According to Mintzer (1982) and Goodman and Carey (1977) the halfband filter is a special FIR filter, and the following characteristics are mentioned by Vaidyanathan and Nguyen (1987) and Bellanger (1977).

1. Number of filter's order is even and length of the filter coefficients is odd. Even number of filter coefficients are 0 except the middle one of the filter coefficients which is 0.5 as Figure 6.1 demonstrated.

2. Filter's passband ripple $\delta_p$ and stopband attenuation $\delta_s$ are same.

3. Filter's passband edge and stopband edge are symmetrical with half output sampling frequency which is $f_s/4$(2x decimation).

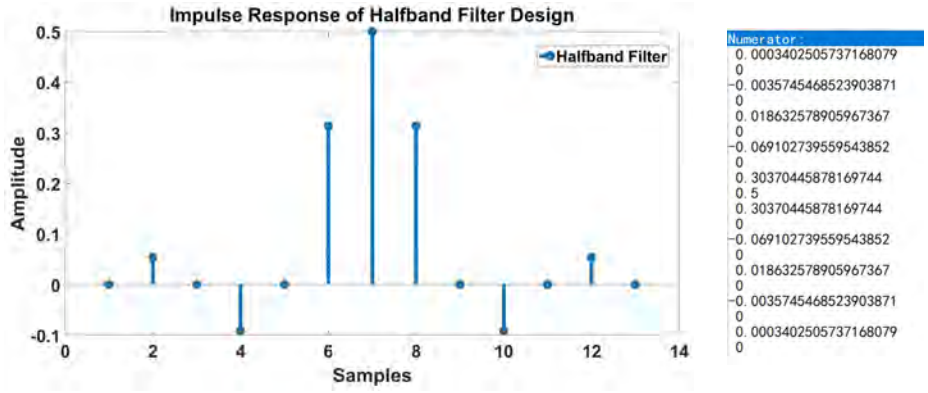



Figure 6.1: Impulse Response and Coefficients of Halfband Filter Design

Therefore, according to the above characteristics, advantages and disadvantages can be summarized as:

1. Cost efficient due to the 0 coefficients.
2. Has flat passband ripple when the stopband attenuation is large.
3. The overshoot may be occurred due to the larger stopband edge.

The first two advantages can be understand intuitively, the third disadvantage is a little bit abstract. Hence, an example is given below to explain what is overshoot and why it may occur in the filter design using halfband.

According to Jørgensen, Pracný, and Bruun (2013), it can be known the stopband edge for each stage has been set to $f_{ri} - f_s$ which means the transition bandwidth has been limited to $f_{ri} - f_s - f_p$. Like the following Figure 6.2 shows.

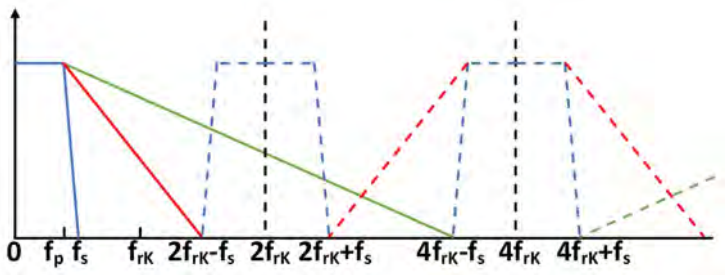



Figure 6.2: Schematic Diagram of Multi-stage Multi-rate Filter Design

However, the half-band filter's transition bandwidth has been defined as: $f_{ri} - f_p - f_p$ which means its transition bandwidth is larger than the limitation set by Crochiere and Rabiner (1975). Hence, the filter's magnitude response curve may become to the above Figure 6.3 shows.

As following Figure 6.3 expressed, the overshoot occurred between $2^{nd}$ stage filter and $3^{rd}$ stage filter. If we just have these two stages, the designed filter will not achieve the design requirements due to the overshoot problem. But if we look back into the diagram it can be seen the $1^{st}$ stage filter has already attenuated a little bit. Therefore, if we can control the overshoot signal not larger than the $1_{st}$ stage filter's attenuation at that frequency range the final filter design still can achieve the design requirements and the filter order and costs can be reduced. However, if the design still cannot meet the design requirements, stopband attenuation can be adjusted to make sure the cascaded filter can achieve the design requirements.

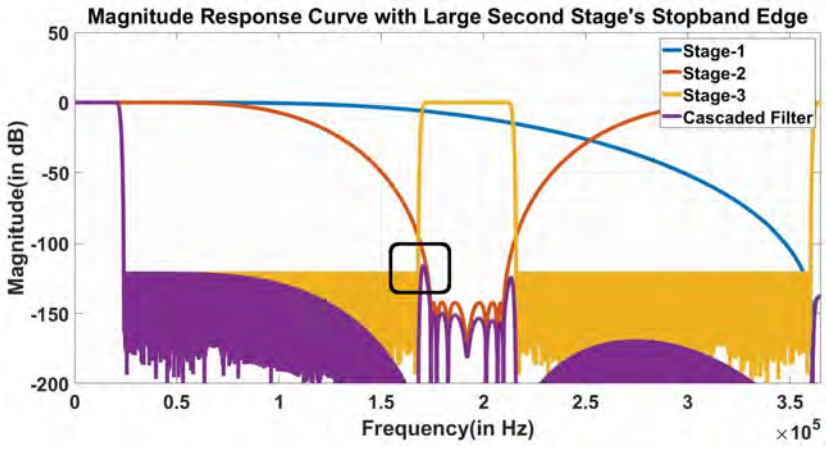



Figure 6.3: Diagram of Overshoot Problem in Multi-stage Filter Design

Fortunately, base on the above Table 6.1, 6.2 mentioned filter design requirements and following Figure 6.4, the overshoot problem didn't happened due to the superposition of first stage's transition bandwidth curve, second and third stage's stopband edge and attenuation. And the absolute value of passband ripple is only around 0.00004dB. Therefore, it can be said that the design requirements are fully achieved.

Table 6.3: Designed [8 2 4] Filter's Costs and Groupdelay

| | Normal | Halfband | Stage-1 | Stage-2 | Stage-2(HF) | Stage-3 |
|---|---|---|---|---|---|---|
| $NO.M$ | 429 | 424 | 65 | 16 | 11 | 348 |
| $NO.A$ | 426 | 421 | 64 | 15 | 10 | 347 |
| $MpIS$ | 14.5625 | 14.25 | 8.125 | 8 | 5.5 | 87 |
| $ApIS$ | 14.3594 | 14.0469 | 8 | 7.5 | 5 | 86.75 |
| $GD_{Sample}$ | 2868 | 2880 | 32 | 7.5 | 9 | 173.5 |

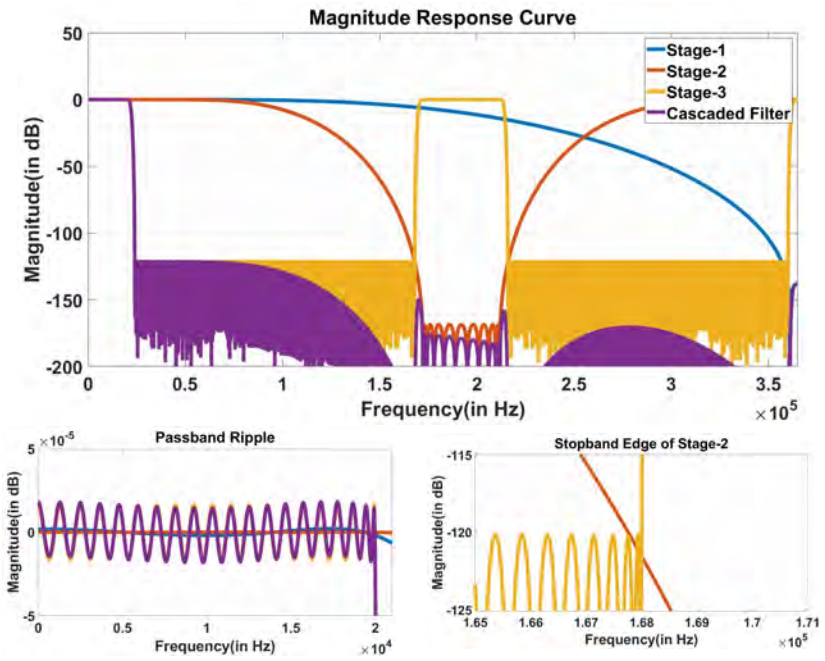



Figure 6.4: Magnitude Response Curve of Designed [8 2 4] Filter(Halfband)

From the above parameters and curves it can be found that due to the characteristics of halfband filter, the second stage's passband ripple is particularly small which is not even the same order of magnitude with the third stage's passband ripple. Therefore, a big gap is occurred between the design requirements and the designed filter. In other words, the designed filter does not make full use of the design requirements, which means that the filter design still has room for improvement. Nevertheless, the costs of designed filter are still somewhat reduced compared with the original design. If the comparison is only between the second stage, it can be said that the filter costs are reduced significantly. However, the groupdelay did increase slightly with the reduction of filter costs.

However, the implementation of the above optimization methods cannot be simply designed based on experience or equation. This is because:

1. The filter cost curve is stepped, and it has no one-to-one correspondence with the design parameters.
2. The overshoot problem is affected by different stages' magnitude response curve. But the response curves are hard to be calculated with only design requirements.
3. The filter response curve is an overall expression of the filter coefficients, it is hard to adjust only part of the curve without affecting the overall filter response curve.
4. Although the overall performance of the filter can be improved by adjust some filter design parameters, it is difficult for us to find out the direction of adjustment because there are too many parameters are involved.

Thus, an optimal solution search algorithm is needed for this project. This algorithm can be used to solve the optimization direction finding problem. And finally, the simulated annealing algorithm is selected and the following section explains how it works.

## 6.3 Search the Optimal Solution Using Simulated Annealing Algorithm

According to the Steinbrunn, Moerkotte, and Kemper (1997), Ma et al. (2018) and Zhang et al. (2019), The earliest idea of Simulated Annealing (SA) was proposed by Metropolis et al. (1953). As Ma, Hu, and Wang (2018) mentioned Kirkpatrick et al. (1983) successfully introduced annealing ideas into the combinatorial optimization field. it is a stochastic optimization algorithm based on the Monte-Carlo iterative solution strategy. It is based on the similarity between the annealing process of solid matter and the general combinatorial optimization problem. The simulated annealing algorithm starts from a certain hot initial temperature, with the continuous decreasing of the temperature and the probability of jump feature to find the global optimal solution randomly within the certain range.

According to Rutenbar (1989), the annealing algorithm needs 4 basic components:

1. Configurations: a model of what a legal placement (configuration) is. These represent the possible problem solutions over which we will search a good answer.

2. Move set: a set of allowable moves that will permit us to reach all feasible configurations and one that is easy to compute. These moves are the computations we must perform to move from configuration to configuration as annealing proceeds.

3. Cost function: to measure how good any given placement configuration is.

4. Cooling schedule: to anneal the problem from a random solution to a good, frozen, placement. Specifically, we need a starting hot temperature (or a heuristic for determining a starting temperature for the current problem) and rules to determine when the current temperature should be dropped, how much the temperature should be dropped, and when annealing should be terminated.

Based on the above four basic components, the configurations need to be given firstly. The configurations are the way to search the possible optimal solutions. Searching direction can be determined by the stages. For $K$ stage multi-rate filter design, the overshoot problem is normally caused by the $K-1$ stage. Hence, the influence of other stages can be ignored. Meanwhile,

from the previous research and test, the passband ripple can affect the filter's performance a lot. Therefore, the key factors for this case are each stage's passband ripple, stopband attenuation and transition bandwidth.

Table 6.4: Matrix of Search Direction

| | |
|---|---|
| Passband Ripple(Stage-3) | Repeat following array $3^4$ times<br>0 1 -1 ... 0 1 -1 |
| Passband Ripple(Stage-2) | Repeat following array $3^3$ times<br>0 0 0 1 1 1 -1 -1 -1 |
| Passband Ripple(Stage-1) | Repeat following array $3^2$ times<br>0 ... 0 1 ... 1 -1 ... -1<br>$3^2$ times $3^2$ times $3^2$ times |
| Stopband Edge(Stage-2) | Repeat following array $3^1$ times<br>0 ... 0 1 ... 1 -1 ... -1<br>$3^3$ times $3^3$ times $3^3$ times |
| Stopband Attenuation(Stage-2) | 0 ... 0 1 ... 1 -1 ... -1<br>$3^4$ times $3^4$ times $3^4$ times |

According to discussion of above sections, there are 5 parameters are key factors need to be adjusted for the filter design specifications mentioned in Table 6.1 and 6.2 which are passband ripple (Ap) of each stages, second stage's stopband edge (Fst) and stopband attenuation (Ast). Therefore, the search direction of each factor can be set to [0 1 -1] where 0 means the current factor maintains the current optimal value, 1 means increasing and -1 means decreasing. In order to make sure the every column is unique, the above Table 6.4 has been created.

However, in practical applications, the first column need to be removed because of the all zeros combination. Also, this example is only a special case, for optimization, number of key factors is not a fixed value. Hence, a calculation method has been proposed for calculate search direction matrix. The nature of this method is the classic permutations problem as the following MATLAB code shows.

Table 6.5: MATLAB Code for Search Direction Matrix Calculation

```
Direction = [0, 1, -1];
KeyFactorNumber = x;
DirectionMatrix = zeros(x,length(Direction)x);

for i = 1:KeyFactorNumber
    tempelements = repmat(Direction',1,(length(Direction)^...
                                (KeyFactorNumber-i)));
    temparray = [tempelements(1,:),tempelements(2,:),tempelements(3,:)];
    DirectionMatrix(i,:) = repmat(temparray,1,(length(Direction)^(i-1)));
end

DirectionMatrix = DirectionMatrix(:,2:end);
```

Then the move set is the border in the search space (specification map). This can be obtained by measuring the border of passband ripple, stopband attenuation and transition bandwidth. These parameters can be obtained from the filter's magnitude response curve which can be calculated by filter coefficients and MATLAB freqz() command. The only problem needs to pay attention is the sampling frequency when reconstructing the magnitude response curve. Because the low resolution curve may cause error measure results.

The cost function is the function to measure the designed filter performance. And results will lead the direction of optimization. Therefore, in this case, filter's costs and latency are the key factors to be measured. And these factors will determine the reference filter design for future search. Finally, the cooling schedule is about the initial temperature $T$, temperature drop rate *delta* and the minimal temperature *Esp*.

**Algorithm 1** Low Latency Cost Efficient Filter Design Algorithm

**Input:** $F_p, F_{st}, A_p, A_{st}, Fs, O_{sr}, Stage, D_F$
**Output:** $Hd, Hd_O$

**function** FILTERDESIGN($F_p, F_{st}, A_P, A_{st}, Fs, O_{sr}, Stage, D_F$)
  *Calculate out related parameter:* $Fs_{in}, F_{sts}, A_{ps}$ *and so on*
  **for** $n = 1 \rightarrow Stage$ **do**
    $Hd\{n\} \leftarrow Design(F_p, F_{sts}(n), A_{ps}, A_{st}, Fs_{in}(n))$
  **end for**
  $Hd \leftarrow cascade(Hd\{1 \rightarrow Stage\})$
  $Best\_Cost\_Temp \leftarrow Cost(Hd)$
  **while** $Temperature\_Current > Esp$ **do**
    **for** $i = 1 \rightarrow length(DirectionMatrix(1,:))$ **do**
      *Adjust Factors*($A_{st}(Stage-1)$, $F_{sts}(Stage-1)$ *and every* $A_{pa}$)
      **if** $F_{sts}(Stage-1)$ *and* $A_{pa}(1:Stage)$ *are feasible* **then**
        **for** $j = 1 \rightarrow Stage$ **do**
          $Ht\_temp\{j\} \leftarrow Design(F_p, F_{sts}(j), A_{pa}\{j\}, A_{st}\{j\}, Fs_{in}(j))$
        **end for**
        $Ht \leftarrow cascade(Ht\_temp\{1 \rightarrow Stage\})$
        $ResponseCurve \leftarrow measure(Ht)$
        **if** *ResponseCurve achieve design requirements* **then**
          $Cost\_Temp\{i\} \leftarrow Cost(Ht)$
        **end if**
      **end if**
    **end for**
    $Best\_Cost\_Current \leftarrow min(Cost\_Temp)$
    **if** $Best\_Cost\_Current \leq Best\_Cost\_Temp$ **then**
      $Hd_O \leftarrow Ht$
      $Best\_Cost\_Temp \leftarrow Best\_Cost\_Current$
      *Assign the parameters of current best design to reference design*
    **end if**
    $Temperature\_Current \leftarrow Temperature\_Current * drop\ rate$
  **end while**
  **return** $Hd, Hd_O$
**end function**

where:

$F_p$ is the passband edge of cascaded filter.
$F_{st}$ is the stopband edge of cascaded filter.
$A_p$ is the passband ripple of cascaded filter.
$Ast$ is the stopband attenuation of cascaded filter.
$Fs$ is the sampling frequency of the system.
$O_{sr}$ is the oversampling rate.
$Stage$ is total stage number of the filter.
$D_F$ is the array for each stage's decimation rate.
$Hd$ is the reference filter design based on Crochiere and Rabiner(1975).
$Hd_O$ is the designed low latency cost efficient filter.
$Fs_{in}$ is the array for each stage's input sampling frequency.
$F_{sts}$ is the array for stopband edge of each stage.
$Aps$ is each stage's passband ripple which calculated by $Ap/Stage$.
$DirectionMatrix$ is the the matrix mentioned in Table 6.4 and 6.5.
$A_{pa}$ is the array for adjusted passband ripple of each stage.

From the previous discussion, the above low latency cost efficient filter design search method has been proposed. In general, this method can be divided into 3 parts: the first one is to get external input parameters, the second part is to sort out and calculate the necessary parameters and the final part is to find out one of the optimal solution using SA algorithm. In this case, the above algorithm can be divided into 3 parts:

1. Prepare the relevant variables and reference filter design.

2. Search the local optimal filter design.

3. Find out the low latency cost efficient filter design.

In the first part, the filter design specifications need to be calculated out and the reference filter design based on Crochiere and Rabiner (1975) need to be designed. The filter design specifications are used to design the reference filter design which can be calculated from the input filter design requirements. And the reference filter design is the starting point of the search algorithm and its costs and latency are the reference benchmark for subsequent filter designs.

In second part, some global variables need to be defined first which are: initial temperature $T = 100$, escape temperature $Esp = 1$, temperature drop rate $delta = 0.95$ and the search direction matrix mentioned above. Then, the search algorithm is ready to start. However, due to the huge amount

of possible search directions, the parallel for loop computation is needed to reduce the search time. Therefore, the relationship between previous result and current computation needs to be minimized. Hence, the designed filters' costs and latency will only be compared after all possible search directions are traversed at the current temperature as Algorithm 1 line 9-21 shown. Also the possible search directions may not meet the filter design requirements for example: filter's stopband edge $F_{sts}$ may less than the filter's passband edge $F_p$ or the the passband ripple is 0. Therefore, the condition check (line 11) before the filter design can reduce the error and unnecessary design significantly.

For the final part, according to the previous discussion and explanation, one current best design could be found if the search for current temperature is done. Thus, the low latency cost efficient filter design can be found out by comparing the current best design and previous best design.

## 6.4 Results and Analysis

As the above part mentioned, the example takes Table 6.1 as the design condition, optimization target are filter's computational cost (CC) and Groupdelay (GD). The following low latency cost efficient filters can be obtained. And as the control group, the filter designs based on Crochiere and Rabiner (1975) are listed below as well.

Table 6.6: Control Group for Low Latency Cost Efficient Filter Design

| | [16,2,2] | [8,4,2] | [8,2,4] | [4,8,2] | [4,4,4] |
|---|---|---|---|---|---|
| $NO.M$ | 394 | 293 | 429 | 310 | 413 |
| $NO.A$ | 346 | 290 | 426 | 307 | 410 |
| $MpIS$ | 12.8281 | 12.5156 | 14.5625 | 13.2969 | 14.9375 |
| $ApIS$ | 12.7188 | 12.3438 | 14.3594 | 13 | 14.6094 |
| $GD_{Sample}$ | 3050 | 3024 | 2868 | 3008 | 2860 |

| | [4,2,8] | [2,16,2] | [2,8,4] | [2,4,8] | [2,2,16] |
|---|---|---|---|---|---|
| $NO.M$ | 737 | 398 | 433 | 737 | 1411 |
| $NO.A$ | 734 | 395 | 430 | 734 | 1408 |
| $MpIS$ | 19.7344 | 14.8594 | 15.5625 | 20.2344 | 29.9531 |
| $ApIS$ | 19.3438 | 14.3125 | 14.9844 | 19.5938 | 29.1875 |
| $GD_{Sample}$ | 2814 | 3000 | 2854 | 2811 | 2791 |

Table 6.7: Best Results for Low Latency Cost Efficient Filter Design(CC)

| | [16,2,2] | [8,4,2] | [8,2,4] | [4,8,2] | [4,4,4] |
|---|---|---|---|---|---|
| $NO.M$ | 342 | 284 | 412 | 302 | 398 |
| $NO.A$ | 339 | 281 | 409 | 299 | 395 |
| $MpIS$ | 12.4219 | 12.1094 | 14.0313 | 12.8906 | 14.6563 |
| $ApIS$ | 12.3125 | 11.9375 | 13.8281 | 12.5638 | 14.3281 |
| $GD_{Sample}$ | 3039 | 2947 | 2751 | 2937.5 | 2746 |

| | [4,2,8] | [2,16,2] | [2,8,4] | [2,4,8] | [2,2,16] |
|---|---|---|---|---|---|
| $NO.M$ | 706 | 385 | 414 | 706 | 1350 |
| $NO.A$ | 703 | 382 | 411 | 703 | 1347 |
| $MpIS$ | 19.0313 | 14.0625 | 14.5938 | 19.5313 | 28.7656 |
| $ApIS$ | 18.6406 | 13.5156 | 14.0156 | 18.8906 | 28 |
| $GD_{Sample}$ | 2694 | 2912.5 | 2737.5 | 2693 | 2670 |

Table 6.8: Best Results for Low Latency Cost Efficient Filter Design(GD)

| | [16,2,2] | [8,4,2] | [8,2,4] | [4,8,2] | [4,4,4] |
|---|---|---|---|---|---|
| $NO.M$ | 342 | 285 | 412 | 304 | 398 |
| $NO.A$ | 339 | 282 | 409 | 301 | 395 |
| $MpIS$ | 12.6563 | 12.1563 | 14.0313 | 13.4063 | 14.6563 |
| $ApIS$ | 12.5469 | 11.9844 | 13.8281 | 13.1094 | 14.3281 |
| $GD_{Sample}$ | 2961.5 | 2939 | 2751 | 2924.5 | 2746 |

| | [4,2,8] | [2,16,2] | [2,8,4] | [2,4,8] | [2,2,16] |
|---|---|---|---|---|---|
| $NO.M$ | 706 | 385 | 414 | 706 | 1350 |
| $NO.A$ | 703 | 382 | 411 | 703 | 1347 |
| $MpIS$ | 19.0313 | 14.0625 | 14.5938 | 19.5313 | 28.7656 |
| $ApIS$ | 18.6406 | 13.5156 | 14.0156 | 18.8906 | 28 |
| $GD_{Sample}$ | 2694 | 2912.5 | 2737.5 | 2693 | 2670 |

If the results presented with figure, it would be:

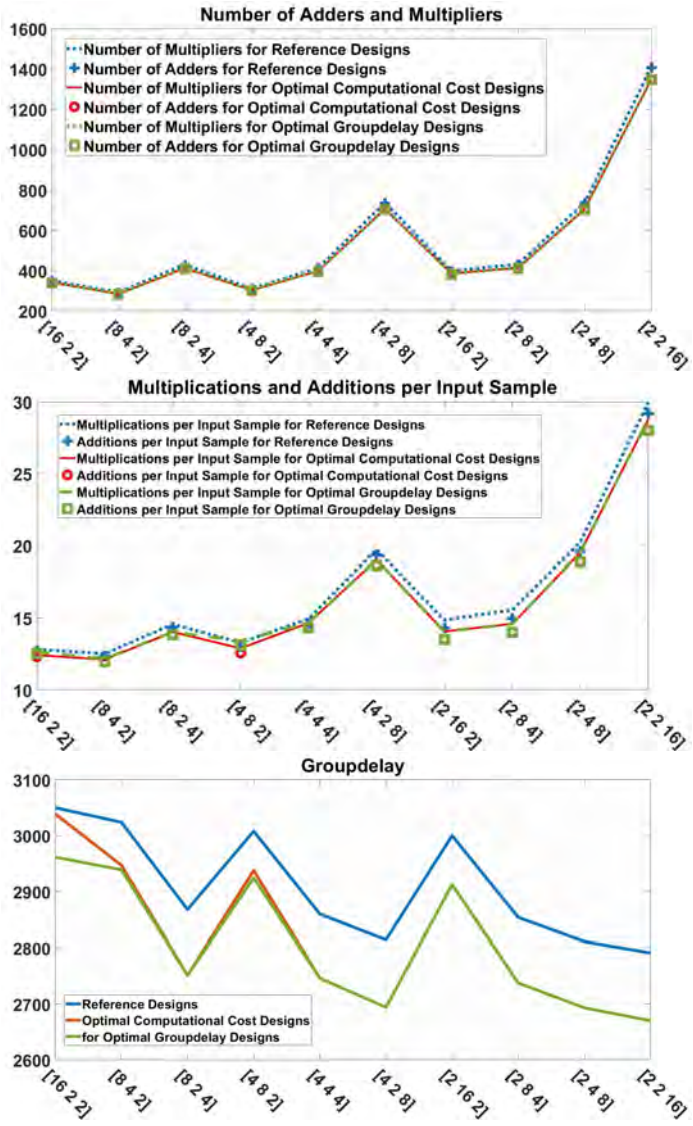



Figure 6.5: Improvement of Low Latency Cost Efficient Designs

As the above Figure 6.5 and Table 6.6, 6.7 and 6.8 presented, the filter designs have been optimized and the performance of the filter has been improved. Although the improvement seems not so obvious on the above Figure 6.5, it is still significant in terms of percentage as the following figure presented.

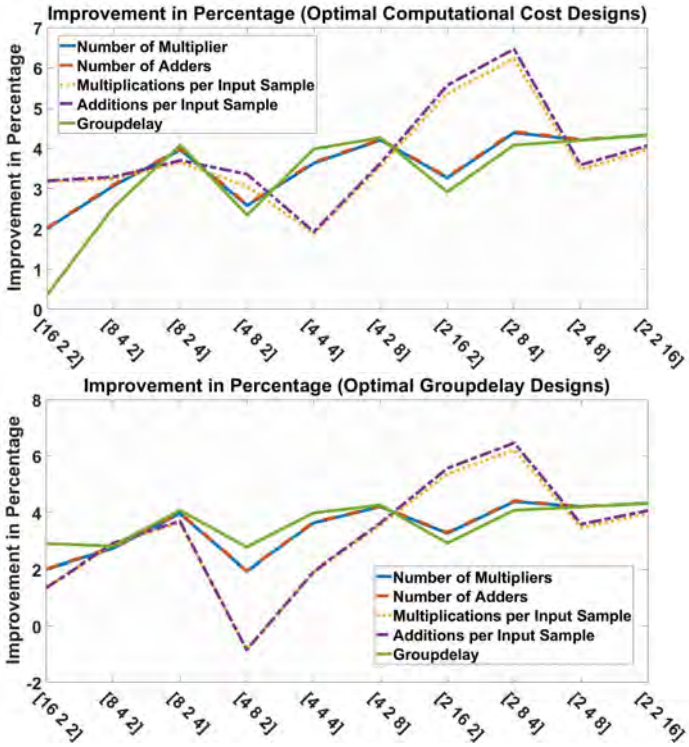



Figure 6.6: Improvement of Low Latency Cost Efficient Designs(Percentage)

From the above Figure 6.6, it can be obtained that in this case, the average improvement for different costs and latency is around 3% to 4% which is quite a significant improvement. However, there are negative optimizations for the multiplications and additions per input sample in the low latency filter design part due to the optimization target is groupdelay, majority of the designed filters have been improved in both costs and latency comparing with the reference filter design. Therefore, it can be said that it is feasible to use the Simulated Annealing algorithm to do the filter design optimization.

However, there are two problems need to be considered. One is the efficiency of the search algorithm. For the above examples, one low latency cost efficient filter design requires around 6 hours of running on a computer based on 6 core Intel i7 processor. The other one is that there is only one solution can be obtained by this search algorithm, but the filter design parameters and filter costs and latency are not one-to-one correspondence. Therefore, the optimal designs should be multiple solutions or interval solution.

## 6.5 MATLAB GUI Based Filter Design and Evaluation Framework

### 6.5.1 Brief Introduction

In order to evaluate and validate our filter design methods as well as to provide an integrated filter design framework within our research scope. We construct a graphic user interfaced (GUI) based framework prototype. This prototype combines the previous research outcomes mentioned in Chapter 4 to 6, which has the following characteristics:

1. Multiple optimization direction can be selected.
2. Flexible filter stage and decimation or interpolation rate selection.
3. Passband ripple and transition bandwidth optimization are added.
4. Halfband filter design can be involved while the filter design meet certain conditions.
5. The exported filter costs are more detailed.

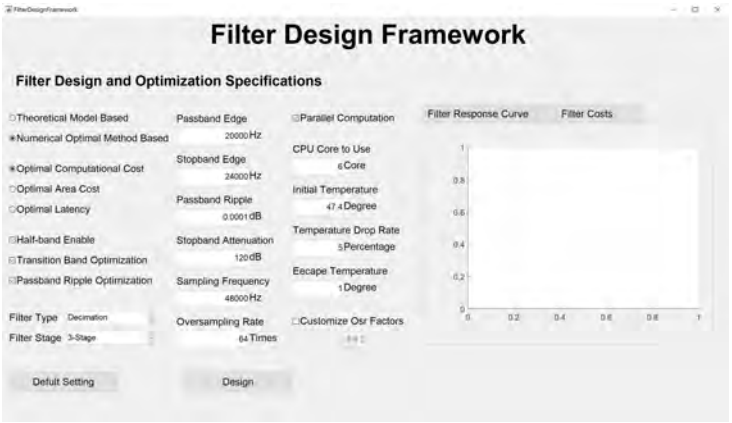



Figure 6.7: Interface of Filter Design and Evaluation Framework

As the above Figure 6.7 expressed, this interface is mainly constituted by two parts. The left part is for input filter design specifications, the right part is for express the details of designed filter.

In the left part, it is necessary to select whether we design the filter base on the mathematical calculation (mainly based on Crochiere and Rabiner (1975; 1981)) or numerical optimal method (based on Section 6.1 to 6.4). If the theoretical model based design method has been selected, the halfband filter can not be involved and the passband ripple and transition bandwidth optimization would be ignored. Otherwise, if the numerical optimal method based filter design method has been selected, it is possible to involve the halfband filter design as well as the passband and transition bandwidth optimization. Also the initial condition setting for simulated annealing algorithm are now available. Finally, what we need to do is fill in the filter design specifications and then design the filter.

When the filter design is completed, the filter's magnitude response curve and filter costs will appear in the right part of the interface like the following figure shows.

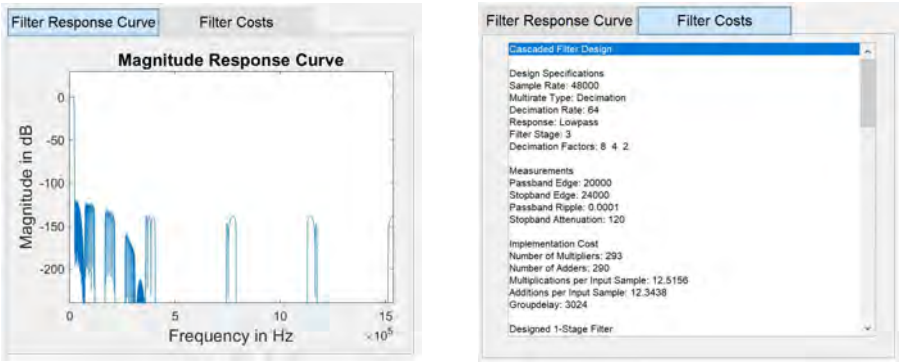


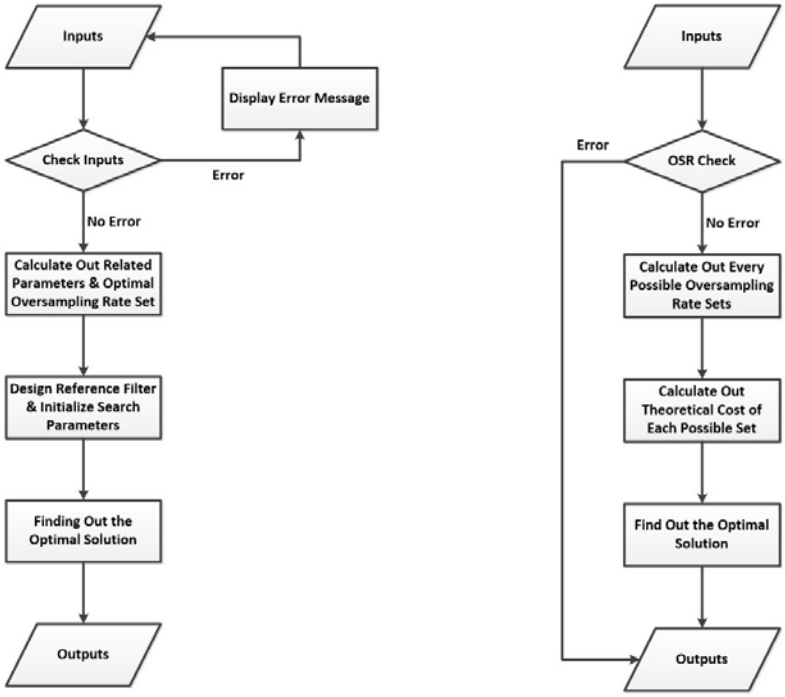

Figure 6.8: Interface of Output Magnitude Response Curve and Filter Costs

After the brief introduction of MATLAB GUI based filter design and evaluation framework, the next subsection will explain how it works.

### 6.5.2 GUI Working Principle

Like the following Figure 6.9 (a) shows, while the inputs are filled in and the design button has been clicked, the framework will check whether these inputs are appropriate. In case the conflict between different inputs are found, the framework will pop up a window which can indicate the error. If the inputs are appropriate, the software will move to data preparation stage which will calculate out the related intermedia parameters. For example: how many

stages the filter needs, possible decimation/interpolation rate for each stage, each stage's initial passband ripple, stopband edge and so on. While in this data preparation stage, the most important parameters are the filter stage number and decimation/interpolation rate for each stage. Because other related variables are calculated based on these two parameters.




(a) (b)

Figure 6.9: Flow Chart of Filter Design and Evaluation Framework 1

In order to calculate out these two parameters, the process shown in Figure 6.9 (b) has been proposed. the total oversampling rate (OSR) can be obtained from the input filter design specifications. After factorization and permutation, every possible decimation/interpolation rate sets can be found out. Hereafter, the theoretical costs of each possible sets could be calculated according to Equation 2.23 to 2.24, 2.28 to 2.29 and 5.19. Hence, the optimal decimation/interpolation rate sets can be found by comparing all results. Coming along with the optimal decimation/interpolation rate set, the number of filter stage can be obtained as well. Therefore, the other related parameters can be easily calculated.

With the calculated intermedia parameters and initial inputs, the theoretical model based optimal filter can be designed. If the 'Theoretical Model Based' option has been selected, the program will calculate, organize and export the designed filter's related information. However, if the 'Numerical Optimal Method Based' is selected, the program will continue to search for the better filter design. In this case, the program will jump into the numerical optimal method based optimal filter design search algorithm which is stated below in Figure 6.10.

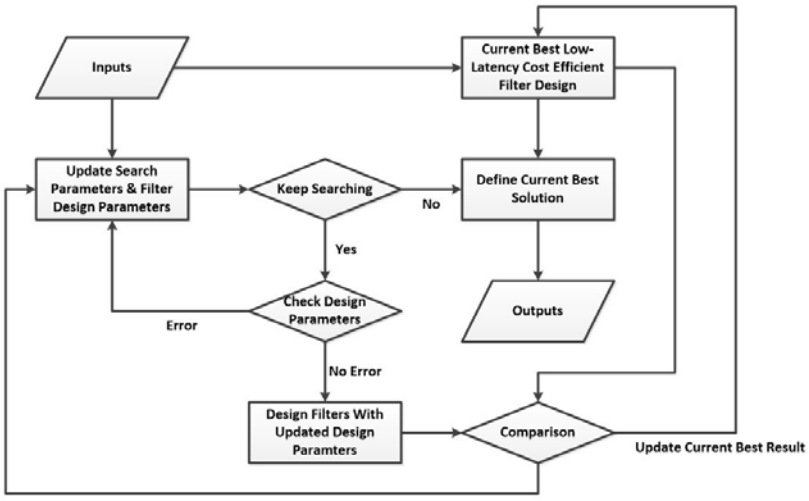



Figure 6.10: Flow Chart of Filter Design and Evaluation Framework 2

There are large number of variables and decision-making processes, the entire process is described the above flow chart 6.10 shows. According to the previous calculation, the reference filter design and related parameters are obtained. Combine with the input parameters, the initial search direction could be found out. As the simulated annealing algorithm mentioned in Section 6.3 while the current temperature is higher than the escape temperature, the program will continue update the search parameters to search the better filter design. Simultaneously, the results obtained in each loop will be compared with the current best filter design and update the current best results. When the current temperature reaches the escape temperature, the program stops running and export the searched optimal filter design.

When the optimal filter design has been found, the framework will calculate and derive the magnitude response curve, filter costs and related information of the designed filter by calculating and deriving the relevant parameters of the filter as the Figure 6.8 shows.

### 6.5.3 Conclusion

As the main outcome of the entire PhD thesis, this MATLAB GUI based filter design and evaluation framework incorporates most of the author's research which are:

1. The integer decimation/interpolation rate selection method.
2. The influence of using different filter structures on filter design.
3. Through the simulated annealing algorithm and subtle changes to each filter stage design specifications, a global optimal filter design under the given design requirements can be found.

Currently, this prototype produces a signal design, it can be further developed into a version that can output a range of designs that meet the cost objective. Although the operating speed can be further optimized, this framework can indeed help researchers, technology enthusiasts and engineers in this area to design and analysis the low-latency cost efficient filter designs. At the same time, during the period of my PhD study, my supervisor and I discovered that there are many potential breakthroughs such as the involvement of non-linear filters and so on. In the future, I hope that more and more research concepts could be added to this framework to help more people.

# Chapter 7

# Conclusion and Future Work

To conclude this thesis, we first summarize the findings and contributions made in the area of low latency cost efficient high level audio filter design. We then reflect upon possible improvements that can be made to improve the current high level audio filter design field. Finally we consider some potential avenues for future work.

## 7.1 Conclusion

In this thesis we have investigated the problems of low latency cost efficient filter evaluation and development techniques. The primary aim is to build up a filter development framework to help design the cost efficient low-latency $\Sigma$-$\Delta$ ADC/DAC for audio. There are three main contributions:

1. The new method for fast identifying the optimal decimation or interpolation factors for area and computational cost efficient multi-stage multi-rate filter design.
2. Formulation of new analytical latency estimation function for multi-stage multi-rate linear filters and the new discovery of the design factors that can be used for filter optimisation.
3. Using numeric methods to design optimal multi-stage multi-rate filter with both cost and low latency constraints including half-band filters.

Overall, we have shown that by using mathematical derivations, theoretical analysis, and optimization techniques, it is possible to generate the numerical optimal method based low latency cost efficient filter design method for audio engineers. it can improve the overall performance of $\Sigma$-$\Delta$ Modulation

based high level audio ADC.

Chapter 4 investigated the optimal computational cost and area cost filter design methods. We have found that the decimation rate distribution for the optimal computational cost and area cost have the certain disciplines which is the front stage always has larger decimation rate. Based on this finding, we have presented an improved optimal rate selection method which can reduce the computation complexities and export the implementable optimal solution sets directly. During this period, we have found that for a certain oversampling rate, the optimal decimation rate distribution does not change too much with the changing of transition bandwidth. Thus, the 3D database and knowledge based search for the optimal computational cost and area cost filter designs have been proposed. This method could increase the filter design efficiency significantly.

Chapter 5 explores the relationship between filter's latency and filter design parameters. We have proposed a latency estimation function for the optimal narrow band linear phase FIR filter design based on Crochiere and Rabiner (1975). By combining this estimation function with the previous 3-D database, we then have an efficient and balanced filter design and search method. Meanwhile, from the analysis we have deduced the decimation rate distribution for the optimal latency filter design in which the front stage always has smaller decimation rate and the decimation rate for the last stage must be the largest one. We also found that the latency can only be affected by the decimation rate distribution while the filter design requirements are determined. Hereafter, we have investigated into the relationship between computational cost, area cost and latency. Compared with the computational cost and area cost, the latency does not change too much with the changing of filter design requirements.

In Chapter 6, we have investigated into the numerical optimal method based low latency cost efficient filter design method. Based on the previous research outcomes, we first explored the relationship between the widening transition bandwidth and the overshoot problem. By investigating using half-band filter structure in low latency cost efficient filter design, we have found the passband ripple allocation affects the filter costs significantly. We then investigated how to combine these filter costs reduction methods together to design the low latency cost efficient filter. Accompanied by the using of simulated annealing algorithm, we finally proposed a numerical optimal method based low latency cost efficient filter design framework which has around 3%-4% improvement for different costs compared with different traditional

optimal filter design methods.

## 7.2 Critique and Future Work

Although the thesis proposes an improved low latency cost efficient filter design framework that is shown to provide better performance than the traditional optimal filter design methods. It could be potentially improved by adding more considerations.

1. Non-linear phase filter consideration.

   Fouda et al. (2009) and Chu and Burrus (1984) mentioned that the non-linear phase IIR filter has the advantage of low latency and cost compared with linear phase FIR filters. Also the flat passband is incomparable with linear phase FIR filter. But due to the varying latency and the feedback loop, these kind of filters have not been frequently used in the ADC/DAC hardware design. However, Hogenauer (1980, 1981) and Aboushady et al. (2001) stated that the Cascaded Integrator-Comb (CIC), Infinite Impulse Response (IIR) and minimum phase Finite Impulse Response (FIR) can achieve good performance with low cost.

   Due to the limitations of varying delay, filter performance or usage conditions, above listed filter haven't been introduced to $\Sigma$-$\Delta$ modulation based high-level audio ADC/DAC. However from the previous research, author found that the latency ascent rate of IIR filter is correlated with passband edge. And Ardalan and Paulos (1987) Claasen and Mecklenbrauker (1982) pointed out that the ascent rate is not fast at the beginning. therefore, if the IIR can be implemented to the front stages with larger passband edge, the influence of varying groupdelay can be minimized. thereby, the flat passband of IIR filter can provide more room for improvement.

   Following the above discussion, author will further research the characteristics of non-linear filter and the effect of varying latency on high-level audio signals. Also author will try to improve the existing optimization methods by adding non-linear filter to filter design and evaluation framework.

Currently, author's supervisor Prof. Wei Hu and his team in Wuhan University of Science and Technology are continuing to expand the different filter types into this method based on my research. For example, Hu et al. (2020) add CIC filters into this approach. In the future, it is interesting to add more considerations into this framework for example, hardware consideration.

2. Hardware consideration.

   According to the principle of Σ-Δ Modulation based ADC/DAC filter, the sampling frequency of input signal is very high. Therefore, the clock frequency must be very high as well. However, the sampling frequency of the signal will be declined rapidly due to the decimation process. And we also known that the last stage normally has the most adders and multipliers. Hence, if the adders or multipliers or can be used multiple times within one sample period, the area cost could be reduced significantly.

   Also, Blad and Gustafsson (2008) Summerfield, Kershaw, and Sandler (1994) and Shahein et al. (2012) have looked into further bit-level optimization of filter coefficients in a multi-stage architecture, and specific application cases were studied by Liu, Jiang, and Zhang, 2014, Mahesh and Vinod (2008) and Mottaghi-Kashtiban, Farazi, and Shayesteh, 2006. Hence, it would be interesting to see how the optimal filter can be applied to those cases.

3. Improvement of evaluation framework

   Current low latency cost efficient filter design framework can only provide one group of optimal solution, however, as Figure 4.9 demonstrated. the optimal solution should be a interval instead of just one point. Therefore, author will further research on how to find the boundary between optimal and non-optimal and perfect the proposed framework.

# Reference


Aboushady, H. et al. (2001). "Efficient polyphase decomposition of comb decimation filters in /spl Sigma//spl utri/ analog-to-digital converters". In: *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 48.10, pp. 898–903. DOI: 10.1109/82.974776.

Agarwal, S., V. K. Pavankumar, and R. Yokesh (2008). "Energy-Efficient, High Performance Circuits for Arithmetic Units". In: *21st International Conference on VLSI Design (VLSID 2008)*, pp. 371–376.

Ansari, R. and Bede Liu (1985). "A class of low-noise computationally efficient recursive digital filters with applications to sampling rate alterations". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 33.1, pp. 90–97. DOI: 10.1109/TASSP.1985.1164494.

Ardalan, S. and J. Paulos (1987). "An analysis of nonlinear behavior in delta - sigma modulators". In: *IEEE Transactions on Circuits and Systems* 34.6, pp. 593–603. DOI: 10.1109/TCS.1987.1086187.

Aziz, P.M., H.V. Sorensen, and J. vn der Spiegel (1996). "An overview of sigma-delta converters". In: *IEEE Signal Processing Magazine* 13.1, pp. 61–84. DOI: 10.1109/79.482138.

Bellanger, M (1977). "Computation rate and storage estimation in multirate digital filtering with half-band filters". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 25.4, pp. 344–346.

Bellanger, M., J. Daguet, and G. Lepagnol (1974). "Interpolation, extrapolation, and reduction of computation speed in digital filters". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 22.4, pp. 231–235. DOI: 10.1109/TASSP.1974.1162581.

Blad, Anton and Oscar Gustafsson (2008). "Bit-level optimized FIR filter architectures for high-speed decimation applications". In: *2008 IEEE International Symposium on Circuits and Systems*, pp. 1914–1917. DOI: 10.1109/ISCAS.2008.4541817.

Boser, B.E. and B.A. Wooley (1988). "The design of sigma-delta modulation analog-to-digital converters". In: *IEEE Journal of Solid-State Circuits* 23.6, pp. 1298–1308. DOI: 10.1109/4.90025.

Brandt, B.P. and B.A. Wooley (1994). "A low-power, area-efficient digital filter for decimation and interpolation". In: *IEEE Journal of Solid-State Circuits* 29.6, pp. 679–687. DOI: 10.1109/4.293113.

Candy, J. (1985). "A Use of Double Integration in Sigma Delta Modulation". In: *IEEE Transactions on Communications* 33.3, pp. 249–258. DOI: 10.1109/TCOM.1985.1096276.

— (1986). "Decimation for Sigma Delta Modulation". In: *IEEE Transactions on Communications* 34.1, pp. 72–76. DOI: 10.1109/TCOM.1986.1096432.

Candy, J. and O. Benjamin (1981). "The Structure of Quantization Noise from Sigma-Delta Modulation". In: *IEEE Transactions on Communications* 29.9, pp. 1316–1323. DOI: 10.1109/TCOM.1981.1095151.

Chae, Youngcheol and Gunhee Han (2009). "Low Voltage, Low Power, Inverter-Based Switched-Capacitor Delta-Sigma Modulator". In: *IEEE Journal of Solid-State Circuits* 44.2, pp. 458–472. DOI: 10.1109/JSSC.2008.2010973.

Chafe, Chris, Juan-Pablo Caceres, and Michael Gurevich (2010). "Effect of temporal separation on synchronization in rhythmic performance". In: *Perception* 39.7, pp. 982–992.

Chu, Shuni and C. Burrus (1984). "Multirate filter designs using comb filters". In: *IEEE Transactions on Circuits and Systems* 31.11, pp. 913–924. DOI: 10.1109/TCS.1984.1085447.

Claasen, T. and W. Mecklenbrauker (1982). "On stationary linear time-varying systems". In: *IEEE Transactions on Circuits and Systems* 29.3, pp. 169–184. DOI: 10.1109/TCS.1982.1085130.

Coffey, Mark W (2003). "Optimizing multistage decimation and interpolation processing". In: *IEEE Signal Processing Letters* 10.4, pp. 107–110.

— (2007). "Optimizing multistage decimation and interpolation processing—Part II". In: *IEEE Signal Processing Letters* 14.1, pp. 24–26.

Craven, Peter G. (2004). "Antialias Filters and System Transient Response at High Sample Rates". In: *journal of the audio engineering society* 52.3, pp. 216–242.

Crochiere, Ronald E and Lawrence R Rabiner (1975). "Optimum FIR digital filter implementations for decimation, interpolation, and narrow-band filtering". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 23.5, pp. 444–456.

— (1976). "Further considerations in the design of decimators and interpolators". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 24.4, pp. 296–311.

— (1981). "Interpolation and decimation of digital signals—A tutorial review". In: *Proceedings of the IEEE* 69.3, pp. 300–331.

Farner, Snorre et al. (2009). "Ensemble hand-clapping experiments under the influence of delay and various acoustic environments". In: *Journal of the Audio Engineering Society* 57.12, pp. 1028–1041.

Fouda, JS Armand Eyebe et al. (2009). *Toward a group delay reduction in digital filtering*. Vol. 19. 1. Elsevier, pp. 22–32.

Friedman, V. et al. (1989). "A dual-channel voice-band PCM codec using Sigma Delta modulation technique". In: *IEEE Journal of Solid-State Circuits* 24.2, pp. 274–280. DOI: 10.1109/4.18586.

Goodman, D. and M. Carey (1977). "Nine digital filters for decimation and interpolation". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 25.2, pp. 121–126. DOI: 10.1109/TASSP.1977.1162921.

Gray, R. (1987). "Oversampled Sigma-Delta Modulation". In: *IEEE Transactions on Communications* 35.5, pp. 481–489. DOI: 10.1109/TCOM.1987.1096814.

Haas, Helmut (1972). "The influence of a single echo on the audibility of speech". In: *Journal of the Audio Engineering Society* 20.2, pp. 146–159.

Hawksford, M.O. (1994). "An introduction to digital audio". In: *IEE Colloquium on Audio Engineering*, pp. 1/1–114.

Hejn, K., I. Kale, and A. Kurek (1993). "Comparative evaluation of delta-sigma modulator structures used in the oversampling A/D converters". In: *1993 IEEE Instrumentation and Measurement Technology Conference*, pp. 221–226. DOI: 10.1109/IMTC.1993.382646.

Herrmann, O, Lawrence R Rabiner, and DSK Chan (1973). "Practical Design Rules for Optimum Finite Impulse Response Low-Pass Digital Filters". In: *Bell Labs Technical Journal* 52.6, pp. 769–799.

Hogenauer, E. (1980). "A class of digital filters for decimation and interpolation". In: *ICASSP '80. IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 5, pp. 271–274. DOI: 10.1109/ICASSP.1980.1170846.

Hogenauer, Eugene (1981). "An economical class of digital filters for decimation and interpolation". In: *IEEE transactions on acoustics, speech, and signal processing* 29.2, pp. 155–162.

Hu, Wei et al. (2020). "A Novel Economical Design of Multistage Decimators for Low Latency Audio Applications". In: *2020 IEEE 22nd International Conference on High Performance Computing and Communications; IEEE 18th International Conference on Smart City; IEEE 6th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pp. 923–929. DOI: 10.1109/HPCC-SmartCity-DSS50907.2020.00123.

Huang, Der-Feng (2003). "The direct integer factorization approach to the Crochiere and Rabiner multistage FIR designs for multirate systems". In:

*Image and Signal Processing and Analysis, 2003. ISPA 2003. Proceedings of the 3rd International Symposium on*. Vol. 2. IEEE, pp. 1060–1065.

Huang, Der-Feng and Shau-Ren Hung (2009). "The optimum design of multistage multirate FIR filter for audio signal sampling rate conversion via a genetic algorithm approach". In: *Image and Signal Processing, 2009. CISP'09. 2nd International Congress on*. IEEE, pp. 1–5.

Ifeachor, E.C. and P.M. Smithson (1995). "A practical approach to digital signal processing". In: *IEE Colloquium on Teaching of Digital Signal Processing*. London, UK: IET., pp. 6/1–6/8. DOI: 10.1049/ic:19950210.

Johnson, C. (1984). "Adaptive IIR filtering: Current results and open issues". In: *IEEE Transactions on Information Theory* 30.2, pp. 237–250. DOI: 10.1109/TIT.1984.1056883.

Jørgensen, I. H. H., P. Pracný, and E. Bruun (2013). "Hardware-Efficient Implementation of Half-Band IIR Filter for Interpolation and Decimation". In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 60.12, pp. 892–896. DOI: 10.1109/TCSII.2013.2285975.

Kaiser, James F (1974). "Nonrecursive digital filter design using the I_0-sinh window function". In: *Proc. IEEE Int. Symp. Circuits & Syst., 1974*.

Kale, I. et al. (1995). "A high-fidelity decimator chip for the measurement of sigma-delta modulator performance". In: *IEEE Transactions on Instrumentation and Measurement* 44.5, pp. 933–939. DOI: 10.1109/19.414503.

Kester, Walt (2000). *Mixed-Signal and DSP Design Techniques*. Norwood, MA: Analog Devices.

— (2015). "A Brief History of Data Conversion: A Tale of Nozzles, Relays, Tubes, Transistors, and CMOS". In: *IEEE Solid-State Circuits Magazine* 7.3, pp. 16–37. DOI: 10.1109/MSSC.2015.2442371.

Kirkpatrick, S. et al. (1983). "Optimization by Simulated Annealing". In: *Science*.

Lapsley, Phil et al. (1997). "Digital Signal Processing and DSP Systems". In: *DSP Processor Fundamentals: Architectures and Features*. Wiley-IEEE Press, pp. 1–7. DOI: 10.1109/9780470544433.ch1.

Leslie, T.C. and B. Singh (1990). "An improved sigma-delta modulator architecture". In: *IEEE International Symposium on Circuits and Systems*, 372–375 vol.1. DOI: 10.1109/ISCAS.1990.112045.

Lesso, Paul and Anthony J. Magrath (2005). "An Ultra High Performance DAC with Controlled Time-Domain Response". In: *journal of the audio engineering society*.

Lester, Michael and Jon Boley (2007). "The effects of latency on live sound monitoring". In: *Audio Engineering Society Convention 123*. Audio Engineering Society.

Leung, B.H. et al. (1988). "Area-efficient multichannel oversampled PCM voice-band coder". In: *IEEE Journal of Solid-State Circuits* 23.6, pp. 1351–1357. DOI: 10.1109/4.90031.

Lipshitz, Stanley and John Vanderkooy (2001). "Towards a Better Understanding of 1-Bit Sigma-Delta Modulators". In: *Journal of the Audio Engineering Society*.

Liu, S., W. Jiang, and M. Zhang (2014). "Dual-channel multiplexing technology and its realization in interpolation filter in stereo audio sigma-delta DAC". In: *Analog Integrated Circuits and Signal Processing* 81.2, pp. 487–494.

Lyons, Richard G (2011). *Understanding Digital Signal Processing, (3rd Ed.)* Pearson Education India.

Ma, R. et al. (2018). "Optimum design of multistage half-band FIR filter for audio conversion using a simulated annealing algorithm". In: *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 74–78. DOI: 10.1109/ICIEA.2018.8397692.

Ma, Rongxuan, Wei Hu, and Yonghao Wang (2018). "Low Delay and Low Cost Audio Surveillance System via a New Optimum FIR Digtal Filter". In: *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, pp. 1541–1546. DOI: 10.1109/TrustCom/BigDataSE.2018.00220.

Madisetti, Vijay and D Williams (Jan. 1998). *The Digital Signal Processing Handbook*.

Mahesh, R. and A. P. Vinod (2008). "Coefficient decimation approach for realizing reconfigurable finite impulse response filters". In: *2008 IEEE International Symposium on Circuits and Systems*, pp. 81–84. DOI: 10.1109/ISCAS.2008.4541359.

McClellan, J., T. Parks, and L. Rabiner (1973). "A computer program for designing optimum FIR linear phase digital filters". In: *IEEE Transactions on Audio and Electroacoustics* 21.6, pp. 506–526. DOI: 10.1109/TAU.1973.1162525.

Metropolis, R. et al. (1953). "Simulated annealing". In: *The Journal of Chemical Physics* 21.161, pp. 1087–1092.

Mintzer, F. and Bede Liu (1977). "Aliasing error in the multirate implementation of narrowband filters". In: *ICASSP '77. IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 2, pp. 105–108. DOI: 10.1109/ICASSP.1977.1170269.

Mintzer, Fred (1982). "On half-band, third-band, and Nth-band FIR filters and their design". In: *IEEE Transactions on acoustics, speech, and signal processing* 30.5, pp. 734–738.

Mitra, S and J Kaiser (Jan. 1993). *Handbook for Digital Signal Processing.*

Mitra, Sanjit K (2011). *Digital signal processing: A computer based approach (4th Ed.)* McGraw Hill.

Mottaghi-Kashtiban, Mahdi, Saeed Farazi, and Mahrokh G. Shayesteh (2006). "Optimum Structures for Sample Rate Conversion from CD to DAT and DAT to CD Using Multistage Interpolation and Decimation". In: *2006 IEEE International Symposium on Signal Processing and Information Technology*, pp. 633–637. DOI: 10.1109/ISSPIT.2006.270877.

Nelson, G., L. Pfeifer, and R. Wood (1972). "High-speed octave band digital filtering". In: *IEEE Transactions on Audio and Electroacoustics* 20.1, pp. 58–65. DOI: 10.1109/TAU.1972.1162338.

Oppenheim, A. V. et al. (1976). "Digital Signal Processing and Theory and Application of Digital Signal Processing". In: *Physics Today* 29.1, pp. 63–64.

Oppenheim, Alan V., Ronald W. Schafer, and John R. Buck (1999). *Discrete-time Signal Processing (2nd Ed.)* Upper Saddle River, NJ, USA: Prentice-Hall, Inc. ISBN: 0-13-754920-2.

Parks, T. and J. McClellan (1972). "Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase". In: *IEEE Transactions on Circuit Theory* 19.2, pp. 189–194. DOI: 10.1109/TCT.1972.1083419.

PCM1840 (2019). *PCM1940 Quad Channel, 32-Bit, 192-kHz, Burr-Brown TM Audio ADC Datasheet.* URL: https://www.ti.com.cn/lit/ds/symlink/pcm1840.pdf?ts=1624250961849&ref_url=https%253A%252F%252Fwww.ti.com.cn%252Fproduct%252Fcn%252FPCM1840.

PCM4220 (2009). *High-Performance, 2-Channel, 24-Bit, 216kHz Sampling, Delta-Sigma A/D Converter Datasheet.* URL: https://www.ti.com.cn/cn/lit/ds/symlink/pcm4220.pdf?ts=1624250525025&ref_url=https%253A%252F%252Fwww.ti.com.cn%252Fproduct%252Fcn%252FPCM4220.

PCM56 (1993). *Serial Input 16-Bit Monolithic DIGITAL-TO-ANALOG CONVERTER.* URL: https://www.ti.com.cn/cn/lit/ds/symlink/pcm56.pdf?ts=1623679491141&ref_url=https%253A%252F%252Fwww.ti.com.cn%252Fproduct%252Fcn%252FPCM56.

Peled, A. and Bede Liu (1974). "A new hardware realization of digital filters". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 22.6, pp. 456–462. DOI: 10.1109/TASSP.1974.1162619.

Proakis, J.G. and Dimitris Manolakis (1996). *Digital Signal Processing : Principles, Algorithms, and Applications.* Prentice-Hall.

Rabiner, L. and R. Crochiere (1975). "A novel implementation for narrowband FIR digital filters". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 23.5, pp. 457–464. DOI: 10.1109/TASSP.1975.1162727.

Reiss, Joshua D. (2008). "Understanding Sigma-Delta Modulation: the Solved and Unsolved Issues". In: *Journal of the Audio Engineering Society* 56.1/2, pp. 49–64.

Rutenbar, Rob A (1989). "Simulated annealing algorithms: An overview". In: *IEEE Circuits and Devices magazine* 5.1, pp. 19–26.

Schreier, Richard, Gabor C Temes, and Steven R Norsworthy (1997). *Delta-Sigma Data Converters: Theory, Design, and Simulation*. IEEE.

Shahein, Ahmed et al. (2012). "A Novel Hybrid Monotonic Local Search Algorithm for FIR Filter Coefficients Optimization". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 59.3, pp. 616–627. DOI: 10.1109/TCSI.2011.2165409.

Shively, R. (1975). "On multistage finite impulse response (FIR)filters with decimation". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 23.4, pp. 353–357. DOI: 10.1109/TASSP.1975.1162692.

Smith, S (Jan. 1997). *The Scientist & Engineer's Guide to Digital Signal Processing*.

Smith, Steven W. (2003). *Digital signal processing: a practical guide for engineers and scientists*. English. London;Boston, [Mass.]; Newnes. ISBN: 0080477321;9780080477329;

Sozanski, Krzysztof (2013). "A linear-phase IIR filter for audio signal interpolator". In: *2013 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pp. 65–69.

Stein, Yaakov (Oct. 2001). "Digital Signal Processing: A Computer Science Perspective". In: ISBN: 9780471200598. DOI: 10.1002/047120059X.fmatter_indsub.

Steinbrunn, Michael, Guido Moerkotte, and Alfons Kemper (1997). "Heuristic and randomized optimization for the join ordering problem". In: *The VLDB Journal—The International Journal on Very Large Data Bases* 6.3, pp. 191–208.

Stojnic, Milena and Djordje Babic (2011). "Design of optimized multistage decimation filter chain based on optimal factorization of decimation ratio". In: *2011 10th International Conference on Telecommunication in Modern Satellite Cable and Broadcasting Services (TELSIKS)*. Vol. 1, pp. 273–276. DOI: 10.1109/TELSKS.2011.6112051.

Summerfield, S., S.M. Kershaw, and M.B. Sandler (1994). "VLSI design of a sigma-delta bitstream FIR filter". In: *1994 IEE Colloquium on Digital and Analogue Filters and Filtering Systems (Digest No. 1994/233)*, pp. 3/1–3/5.

Tan, Lizhe and Jean Jiang (2018). *Digital signal processing: fundamentals and applications*. Academic Press.

Trick and T. (1975). "Theory and application of digital signal processing". In: *Acoustics, Speech and Signal Processing, IEEE Transactions on*.

Vaidyanathan, P. and Truong Nguyen (1987). "A 'trick' for the design of FIR half-band filters". In: *IEEE Transactions on Circuits and Systems* 34.3, pp. 297–300. DOI: 10.1109/TCS.1987.1086124.

Vegte, Joyce Van De (2001). *Fundamentals of Digital Signal Processing with Cdrom*. 1st. USA: Prentice Hall PTR. ISBN: 0130160776.

Wang, Yonghao (2011). "Latency Measurements of Audio Sigma Delta Analog to Digital and Digital to Analog Converts". In: *Audio Engineering Society Convention 131*.

Wang, Yonghao and Joshua D. Reiss (2012). "Time Domain Performance of Decimation Filter Architectures for High Resolution Sigma Delta Analogue to Digital Conversion". In: *Audio Engineering Society Convention 132*.

Wang, Yonghao, Ryan Stables, and Joshua Reiss (2010). "audio latency measurement for desktop operating systems with onboard soundcards". In: *journal of the audio engineering society*.

Williams, Arthur B. and Fred J. Taylor (2006). *Electronic Filter Design Handbook, Fourth Edition*. New York: The McGRAW-HILL Companies, Inc.

Yang, Yuqing, Terry Sculley, and Jacob Abraham (2007). "A 216 khz 124 db Single Die Stereo Delta Sigma Audio Analog-to-Digital Converter". In: *Fournal of the Audio Engineering Society*.

Yuan, Bo and Hongxia Liu (2014). "Implementation Technology and Algorithm for Low Power Cost Fixed Point Decimal Multiplier(in Chinese)". In: *Journal of Central South University(Science and Technology)* 1, p. 141.

Zambreno, Joseph, David Nguyen, and Alok Choudhary (2004). "Exploring area/delay tradeoffs in an AES FPGA implementation". In: *International Conference on Field Programmable Logic and Applications*. Springer, pp. 575–585.

Zhang, J. et al. (2019). "Multi-Objective Optimization on Multistage Half-Band FIR Filter Design Using Simulated Annealing Algorithm". In: *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pp. 1101–1106. DOI: 10.1109/HPCC/SmartCity/DSS.2019.00156.

Zhu, Xiangyu and Yonghao Wang (2016). *Survey of current digital antialiasing/image filter design specifications of audio convertors*. URL: http://dmtlab.bcu.ac.uk/leowang/whitepapers/DFSurvey2016.pdf.

Zhu, Xiangyu et al. (2016). "Practical considerations on optimising multi-stage decimation and interpolation processes". In: *Digital Signal Processing (DSP), 2016 IEEE International Conference on*. IEEE, pp. 370–374.