

Design of Noise Shaping Filters in $\Delta\Sigma$ Modulators for Quantization Error Reduction

(量子化誤差削減のための $\Delta\Sigma$ 変調器の
ノイズ整形フィルタの設計)

by

Muhammad Rizwan Tariq
D151390

A thesis presented for the degree of
Doctor of Engineering

Department of System Cybernetics
Graduate School of Engineering
Hiroshima University

March, 2018

Contents

Acknowledgements	2
Abstract	3
1 Introduction	4
1.1 Research Background	4
1.2 Aims and Objectives	5
1.3 Thesis Outline	8
2 $\Delta\Sigma$ Modulator and Weighted Quantization Noise	10
2.1 $\Delta\Sigma$ Modulation	10
2.2 Generalized Model of a $\Delta\Sigma$ Modulator	12
2.3 Design Problem Formulation	14
2.3.1 Induced System Norms	15
2.3.2 Stability of $\Delta\Sigma$ Modulators	17
2.4 Conclusions	18
3 Design of FIR Noise Shaping Filters for $\Delta\Sigma$ Modulators	19
3.1 Design Based on H_2 System Norm	20
3.2 Design Based on H_∞ System Norm	21
3.3 Design Based on l_∞ Norm of Error	22

3.4	LMI for Stability Constraint	23
3.5	Design Examples	24
3.5.1	Lowpass $\Delta\Sigma$ Modulator with the 1st Order $H[z]$	25
3.5.2	Lowpass $\Delta\Sigma$ Modulator with the 4th Order $H[z]$	28
3.5.3	Bandpass $\Delta\Sigma$ Modulator with the 6th Order $H[z]$	31
3.5.4	Stability Under the l_∞ Norm Constraint	33
3.6	Conclusions	35
4	Design of IIR Noise Shaping Filters for $\Delta\Sigma$ Modulators	36
4.1	Design Based on the Extended LMI Technique	37
4.1.1	Numerical Design of IIR Noise Shaping Filters	38
4.1.2	Design Examples	44
4.2	Design Based on Approximation Techniques	61
4.2.1	The Yule-Walker Method	62
4.2.2	The Least-Squares (LS) Approximation	64
4.3	Design Based on the Hybrid Technique	71
4.3.1	Design Example	72
4.4	Design Based on the Iterative LMI Technique	76
4.4.1	Design Examples	79
4.5	Conclusions	84
5	Conclusions	86
	Bibliography	89
	Journal Publications	94
	Conference Papers	95

List of Abbreviations

A/D Analog-to-Digital

D/A Digital-to-Analog

SQNR Signal to Quantization Noise Ratio.

OSR Over Sampling Ratio

FIR Finite Impulse Response

IIR Infinite Impulse Response

NTF Noise Transfer Function

GKYP Generalized Kalman-Yakubovich-Popov

LMI Linear Matrix Inequality

BMI Bilinear Matrix Inequality

LTI Linear Time-Invariant

STF Signal Transfer Function

MSE Mean Squared Error

SNR Signal to Noise Ratio

LS Least-Squares

RF Radio Frequency

PA Power Amplifier

BPF Bandpass Filter

List of Figures

2.1	Delta modulation.	11
2.2	Block diagram of $\Delta\Sigma$ modulation.	11
2.3	A quantizer with error feedback filter and a system $H[z]$	12
3.1	The H_2 norm of $H[z]R[z]$ as a function of order of $R[z]$ for the first order lowpass weighting function, where $R[z]$ is designed based on the H_2 norm.	25
3.2	Frequency responses of filters designed by the proposed method and the referenced methods. The weighting function is of order unity. . .	26
3.3	Enlarged frequency response of our proposed filters in Fig. 3.2.	27
3.4	H_2 norm of $H[z]R[z]$ as a function of order of $R[z]$ for the fourth order lowpass weighting function, where $R[z]$ is designed based on the H_2 norm.	29
3.5	Frequency responses of filters designed by the proposed method and the referenced methods. The weighting function is of order 4.	29
3.6	Output and frequency spectrum plot of the lowpass $\Delta\Sigma$ modulator obtained by the proposed H_2 norm based design.	30
3.7	Frequency responses of filters designed by the proposed method and the conventional method. The weighting function is of order 6.	32

3.8	H_2 norm of $H[z]R[z]$ as a function of order of $R[z]$ for the sixth order bandpass weighting function, where $R[z]$ is designed based on the H_2 norm.	32
3.9	Enlarged frequency response of our proposed filters in Fig. 3.7.	33
3.10	Frequency response of the error feedback filter designed by minimizing the upper bound of the Lee coefficient under the constraint on the l_∞ norm of the weighted quantization noise.	34
4.1	Frequency responses of H_2 norm based noise shaping filters subjected to the variance of the feedback signal (Example 1).	44
4.2	MSE of optimal FIR noise shaping filters as a function of their order.	45
4.3	Frequency responses of H_2 norm based noise shaping filters designed by the extended and non-extended LMI techniques (Example 2).	47
4.4	MSE as a function of the Lee coefficient for the extended LMI design.	48
4.5	Empirical SNRs of H_2 norm based noise shaping filters designed by the extended and non-extended LMI techniques as functions of input frequency ω for $y_k = \sin(\omega k)$ (Example 2).	48
4.6	Empirical SNRs of H_2 norm based noise shaping filters designed by the extended and non-extended LMI techniques as functions of input amplitude $a(> 0)$ for $y_k = a \sin(k/100)$ (Example 2).	49
4.7	Output signals of quantizers designed by the extended and non-extended LMI techniques for $y_k = \sin(k/100)$ (Example 2).	50
4.8	Frequency responses of H_∞ norm based noise shaping filters designed by the extended and non-extended LMI techniques (Example 2).	51

4.9	Empirical SNRs of H_∞ norm based filters designed by the extended and non-extended LMI techniques as functions of the input frequency ω for $y_k = \sin(\omega k)$ (Example 2).	52
4.10	Empirical SNRs of H_∞ norm based filters designed by the extended and non-extended LMI techniques as functions of the input amplitude $a(> 0)$ for $y_k = a \sin(k/100)$ (Example 2).	52
4.11	Rotary inverted pendulum.	53
4.12	Frequency responses of noise shaping filters designed by the extended LMI technique with H_2 norm and the <code>synthesizeNTF</code> function (Example 3).	53
4.13	Empirical SNRs of noise shaping filters designed by the H_2 norm based extended LMI technique and the <code>synthesizeNTF</code> function as functions of input frequency ω for $y_k = \sin(\omega k)$ (Example 3).	55
4.14	Empirical SNRs of noise shaping filters designed by the H_2 norm based extended LMI technique and the <code>synthesizeNTF</code> function as functions of input amplitude $a(> 0)$ for $y_k = a \sin(k/20)$ (Example 3).	56
4.15	Empirical SNRs of noise shaping filters designed by the H_2 norm based extended LMI technique and the <code>synthesizeNTF</code> function as functions of input amplitude $a(> 0)$ for $y_k = a \sin(0.12k)$ (Example 3).	56
4.16	Output signals of the system with quantizers designed by the H_2 norm based extended LMI technique and the <code>synthesizeNTF</code> function for $y_k = \sin(k/20)$ (Example 3).	57
4.17	Output signals of the system with quantizers designed by the H_2 norm based extended LMI technique and the <code>synthesizeNTF</code> function for $y_k = \sin(0.12k)$ (Example 3).	57

4.18	MSEs of fixed-point $\Delta\Sigma$ modulators designed by the extended LMI technique with H_2 norm and the <code>synthesizeNTF</code> function as functions of the floating length (Example 3).	58
4.19	Frequency responses of fixed-point and floating-point noise shaping filters designed by the extended LMI technique with H_2 norm and the <code>synthesizeNTF</code> function (Example 3).	58
4.20	Empirical SNRs of fixed-point and floating point noise shaping filters designed by the H_2 norm based extended LMI technique and the <code>synthesizeNTF</code> function as functions of frequency ω of quantized sinusoidal input signals with unit amplitudes (Example 3).	59
4.21	MSE as a function of the FIR filter order.	62
4.22	Frequency responses of FIR filter of order 18, and IIR filter of order 3 for the $\Delta\Sigma$ modulator. The output filter is of order 4.	63
4.23	Frequency responses of FIR and IIR error feedback filters for the low-pass $\Delta\Sigma$ modulator.	66
4.24	Enlarged frequency responses of lowpass filters near the cutoff frequency.	66
4.25	l_2 norm error as a function of the order of the lowpass IIR filter obtained by using the LS method.	68
4.26	Convergence behavior of the iterative algorithm [13] for OSR=[16 32 64 128 256 512]	69
4.27	Frequency responses of FIR and IIR error feedback filters for the bandpass $\Delta\Sigma$ modulator.	69
4.28	Enlarged frequency responses of bandpass filters near cutoff frequencies.	70
4.29	l_2 norm error as a function of the order of the bandpass IIR filter obtained by using the LS method.	70

4.30	Block diagram of an audio amplifier with a $\Delta\Sigma$ modulator as an A/D converter.	73
4.31	Frequency responses of noise shaping IIR filters for a $\Delta\Sigma$ modulator.	73
4.32	Simulation of the lowpass $\Delta\Sigma$ modulator with a noise shaping IIR filter obtained by using the LS approximation method.	74
4.33	Frequency spectrum of the digital output of the lowpass $\Delta\Sigma$ modulator in Fig. 4.32	75
4.34	A simple block diagram of a $\Delta\Sigma$ modulator based RF transmitter.	79
4.35	Frequency responses of noise shaping IIR bandpass filters obtained by using three different design methods. Here, iterative LMI algorithm is initialized using the LS method.	81
4.36	Zeros “o” and poles “x” of noise shaping IIR filters obtained by using the iterative LMI algorithm, <code>synthesizeNTF</code> and hybrid design methods.	82
4.37	Noise Spectrum of the bandpass $\Delta\Sigma$ modulator.	82
4.38	Frequency responses of noise shaping IIR bandpass filters obtained by using three different design methods. Here, iterative LMI algorithm is initialized using the hybrid design.	83

List of Tables

3.1	$\ H[z]R[z]\ _2$, $\ H[z]R[z]\ _\infty$, and l_1 norm of the impulse response of $H[z]R[z]$ for the first order lowpass weighting function.	27
3.2	$\ H[z]R[z]\ _2$, $\ H[z]R[z]\ _\infty$, and l_1 norm of the impulse response of $H[z]R[z]$ for the fourth order lowpass weighting function.	30
3.3	$\ H[z]R[z]\ _2$, $\ H[z]R[z]\ _\infty$, and l_1 norm of the impulse response of $H[z]R[z]$ for the sixth order bandpass weighting function.	34
4.1	Comparison between different design methods	67
4.2	Comparison of $\ H[z]R[z]\ _2^2$ and $\ H[z]R[z]\ _\infty^2$ norm values of different design methods	74
4.3	Variance of the quantization noise ϵ , when iterative LMI is initialized using the LS method.	80
4.4	Variance of the quantization noise ϵ , when iterative LMI is initialized using the hybrid design.	84

Acknowledgements

I would like to express my deepest gratitude to my advisor Dr. Shuichi Ohno for his full support, expert guidance, understanding and encouragement throughout my study and research. Without his incredible patience and timely wisdom and counsel, my research and thesis work would have been a frustrating and overwhelming pursuit.

In addition, I express my appreciation to Professor Dr. Toru Yamamoto, Associate Professor Dr. Yoshifumi Zoka and Assistant Professor Dr. Nakamoto for their valuable insights and comments during seminars and presentations.

Also, I would like to thank all of my laboratory student members, especially the doctoral students Mr. Liao Yuntao and Mr. Guan Zhe for their continuous support and fruitful discussions.

Finally, I would like to thank my parents for their unconditional love and support during my doctoral studies; I would not have been able to complete my research and thesis without their continuous love and encouragement.

Abstract

Optimal finite impulse response (FIR) and infinite impulse response (IIR) noise shaping filters for delta-sigma ($\Delta\Sigma$) modulators are designed based on the system norms. We incorporate the weighting function, connected to the output of the $\Delta\Sigma$ modulator, into our design problem. Then, we minimize the weighted norms of the quantization noise in the output of a $\Delta\Sigma$ modulator, which corresponds to the minimization of the system norm. Three norms, the H_2 system norm, the H_∞ system norm, and the l_1 norm of the impulse response of the system, are adopted. The H_2 system norm can be used to calculate the mean squared error of quantization noise. On the other hand, the H_∞ system norm gives us the worst case gain, while the l_1 norm of the impulse response can minimize the maximum error. The optimization problems for three types of FIR noise shaping filters are evaluated by using linear matrix inequalities (LMIs) and then solved numerically via semi-definite programming. For IIR noise shaping filters, the design problem becomes non-convex, which is hard to solve numerically. To solve the non-convex optimization problem, we propose the extended LMI technique, FIR approximation techniques, the hybrid technique and an iterative LMI algorithm to obtain good IIR noise shaping filters. Design examples are provided to demonstrate the effectiveness of our proposed methods over the existing methods.

Chapter 1

Introduction

1.1 Research Background

Analog-to-digital (A/D) and digital-to-analog (D/A) data converters are some of the most important parts of electronic systems which act as the interfaces between the digital signal world and the real analog world. The performance of digital signal processing and communication systems is generally limited by the precision of the digital input signal which is achieved at the interface between analog and digital information. In A/D converters, the continuous-valued signals are discretized and quantized for transmission over wireline or wireless communication systems [1]. Quantization maps a continuous-valued signal to a discrete-valued signal. This usually introduces undesirable effects, which are resulted from quantization noise. The important aspect of these converters is their ability to determine whether and how much the conversion can correctly keep the important information of signals, while suppressing undesirable noises.

The conventional methods for analog to digital conversion use sampling and quantization to obtain a digital signal. The resolution of these converters depends upon the number of quantization levels. As we increase the quantization levels, the quantization noise decreases but the number of bits per sample also increases, which is not

suitable for applications with limited bandwidth. To solve this problem, we use the technique of delta-sigma ($\Delta\Sigma$) modulation.

Currently, the $\Delta\Sigma$ modulation is a popular technique for making high-resolution A/D and D/A converters [2, 3]. Modern $\Delta\Sigma$ converters offer several benefits including high resolution, low power consumption, and low cost, making them a reasonable choice for the A/D converter for many signal processing applications such as audio devices [4, 5]. These $\Delta\Sigma$ A/D converters are effective for converting analog signals over a wide range of frequencies, from DC to several megahertz.

The $\Delta\Sigma$ modulator mainly consists of a static uniform quantizer and an error feedback filter to shape quantization noise [6], which is called noise shaping filter. The input to the modulator is an oversampled signal which is to be digitized. In oversampling, the signal is sampled at a frequency much higher than the Nyquist frequency (twice the input bandwidth) which reduces the effect of the quantization noise in the frequency band carrying the information signal, while the total noise remains the same.

The high rate digital output of the modulator has two components, one is the signal which is located in the low frequency region and the other is the noise which has to be reduced.

1.2 Aims and Objectives

In the design of a $\Delta\Sigma$ modulator, the objective is to minimize the in-band quantization noise which as a result improves the signal to quantization noise ratio (SQNR) of the $\Delta\Sigma$ modulator. It has been observed that the technique of oversampling alone may not be enough to improve the SQNR in the band of interest, we need to exploit the noise shaping properties of the $\Delta\Sigma$ modulator to further reduce the in-band

quantization noise. This can be achieved by using a feedback filter which employs the noise shaping to obtain a high SQNR while keeping the oversampling ratio (OSR) not too high. Although, the overall quantization noise may not be changed by the noise shaping but SQNR is increased in the information signal frequency band of the frequency spectrum. Therefore, our objective is to design noise shaping filters in the feedback of $\Delta\Sigma$ modulators so that we can minimize the noise in the frequency region which constitutes our signal bandwidth.

Two types of digital filters are available to be designed and utilized as noise shaping filters in the feedback of a $\Delta\Sigma$ modulator: finite impulse response (FIR) and infinite impulse response (IIR) digital filters. As the terminology suggests, these classifications refer to the filter's impulse response. By varying the weights of the coefficients and the number of filter taps, virtually any frequency response characteristic can be realized with an FIR filter. FIR filters can achieve performance levels which are not possible with analog filter techniques (such as perfect linear phase response). However, high performance FIR filters generally require a large number of multiply-accumulates and therefore require fast and efficient digital signal processors (DSPs). On the other hand, IIR filters tend to mimic the performance of traditional analog filters and make use of feedback. Therefore their impulse response extends over an infinite period of time. Because of feedback, IIR filters can be implemented with fewer coefficients than FIR filters.

For FIR digital filters, several designs for feedback filters have been proposed which also use the noise spectrum shaping technique [7, 8]. The FIR error spectrum shaping filters have been proposed for recursive digital filters composed of cascaded second order section in [9]. The method in [10] is a min-max design of the FIR filter which optimizes the noise transfer function (NTF) via generalized Kalman-Yakubovich-Popov (GKYP) lemma. This approach minimizes the worst case gain of

the NTF over the signal frequency band and is shown to be able to improve the overall SQNR of $\Delta\Sigma$ modulators as well. In [11], the filter connected to the $\Delta\Sigma$ modulator is incorporated into the design of the NTF. By using a truncated impulse response, the H_2 system norm is minimized to reduce the in-band quantization noise. The NTF has been designed in [12] by using the weighted noise spectrum under the so called white noise assumption.

Unlike FIR noise shaping filters, there are very few design examples of IIR noise shaping filters in $\Delta\Sigma$ modulators. In the method proposed by [13], the noise shaping filter is assumed to have an IIR which is converted to a minimization problem by virtue of GKYP lemma and solved by using an iterative algorithm. The method in [13] can only minimize the worst-case system gain for obtaining the IIR filter, and also, it does not incorporate the behavior of the non-ideal filter at the output.

We design both FIR and IIR noise shaping filters in the feedback of $\Delta\Sigma$ modulators. The FIR design problem can be formulated as convex optimization problems with linear matrix inequalities (LMIs) [14], which can be solved efficiently. However, the IIR design problems becomes non-convex which consists of bilinear matrix inequalities (BMIs). To solve the non-convex IIR design problem with BMIs, we propose an extended LMI technique [15]. The extended LMI technique assumes that the order of the IIR noise shaping filter is identical to the non-ideal output filter or the weighting function. We also introduce two approximation techniques [16, 17] which can be used to obtain an IIR noise shaping filter without directly solving the non-convex IIR design problem. Also, a hybrid technique [18] is utilized to obtain a stable IIR filter which utilizes the FIR numerator coefficients. Moreover, we also propose an iterative LMI algorithm [19] to obtain IIR noise shaping filters which converts the BMIs into LMIs using the iterative algorithm.

To keep $\Delta\Sigma$ modulators versatile, we utilize the weighting function to design

$\Delta\Sigma$ modulators. We minimize the weighted quantization noise in the output of the $\Delta\Sigma$ modulator. Three norms are adopted to measure the quantity of the weighted quantization noise. One is the variance of the weighted quantization noise when the quantization errors at different time are assumed to be independent of each other. The others are the l_2 and the l_∞ norms of the weighted quantization noise. They correspond to the minimization of the H_2 system norm, the H_∞ system norm, and the l_1 norm of the impulse response of a system, respectively. The stability condition of $\Delta\Sigma$ modulators is also incorporated into our design of FIR and IIR noise shaping filters.

By providing several design examples and performing software simulations, we demonstrate the effectiveness of our proposed methods over the existing methods. All the simulation results are obtained by using MATLAB, while semi-definite programming (SDP) problems are solved by using CVX tool [20], which is an effective solver for convex optimization problems. Throughout this thesis, we also compare and analyze the performances of our proposed noise shaping filters with the existing noise shaping filters.

1.3 Thesis Outline

The organization of this thesis is as follows:

Chapter 2 introduces the $\Delta\Sigma$ modulator and its approximated linearized model with error feedback noise shaping filter. Then, we derive the expression of the weighted quantization noise at the output of the $\Delta\Sigma$ modulator. Three types of different norms are also introduced here. The stability of the $\Delta\Sigma$ modulator is also discussed.

In **Chapter 3**, we propose our optimal design of FIR noise shaping filters for a $\Delta\Sigma$ modulator. Three types of FIR noise shaping filters are designed based on three

kinds of system norms.

Chapter 4 discusses the non-convex nature of the design problem of IIR noise shaping filter. We utilize extended LMI technique, FIR approximation techniques, the hybrid technique and an iterative LMI algorithm to obtain IIR noise shaping filters by solving the non-convex design problem.

Finally, **Chapter 5** provides the conclusions to our thorough study of $\Delta\Sigma$ modulators.

Chapter 2

$\Delta\Sigma$ Modulator and Weighted Quantization Noise

In this Chapter, we introduce the generalized model of a $\Delta\Sigma$ modulator and derive the expression of the weighted quantization noise present in its' output. We formulate our design problem for the minimization of the weighted quantization noise subject to the stability constraint. We also define induced system norms and consider three types of system norms.

2.1 $\Delta\Sigma$ Modulation

$\Delta\Sigma$ modulation was developed as an extension to the well established Delta modulation [21]. Let us review the delta modulation structure for the analog to digital conversion process. Fig. 2.1 shows a block diagram of the Delta modulator. Delta modulation is based on quantizing the change in the signal from sample to sample rather than the absolute value of the signal at each sample. Since the output of the integrator in the feedback loop of Fig. 2.1 tries to predict the input $x(t)$, the integrator works as a predictor. The prediction error term, $x(t) - \bar{x}(t)$, in the current prediction is quantized and used to make the next prediction. The quantized prediction error (delta modulation output) is integrated in the receiver just as it is in the feedback

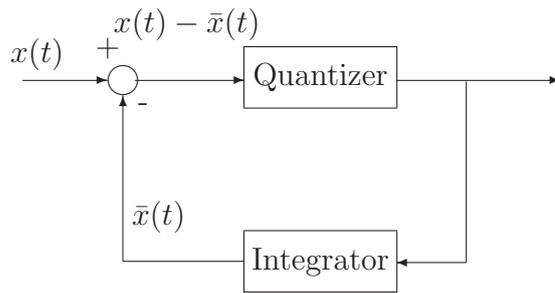


Figure 2.1: Delta modulation.

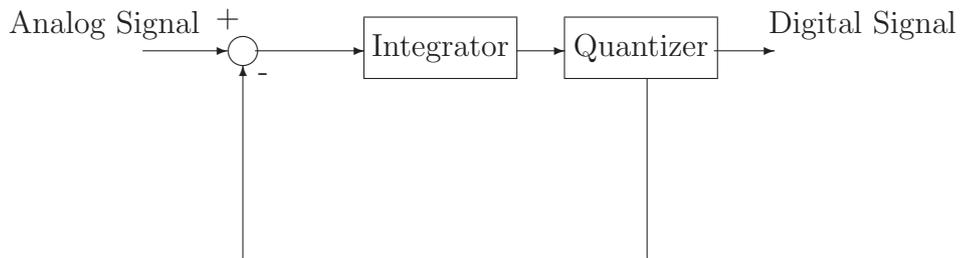


Figure 2.2: Block diagram of $\Delta\Sigma$ modulation.

loop.

The arrangement shown in Fig. 2.2 is called a $\Delta\Sigma$ modulator. The name $\Delta\Sigma$ modulator comes from putting the integrator (sigma) in front of the Delta modulator. Sometimes, the $\Delta\Sigma$ modulator is referred to as an interpolative coder. The quantization noise characteristic (noise performance) of such a coder is frequency dependent in contrast to delta modulation. Like delta modulators, the $\Delta\Sigma$ modulators use a simple coarse quantizer (comparator). However, unlike delta modulators, $\Delta\Sigma$ modulators encode the integral of the signal itself and thus their performance is insensitive to the rate of change of the signal.

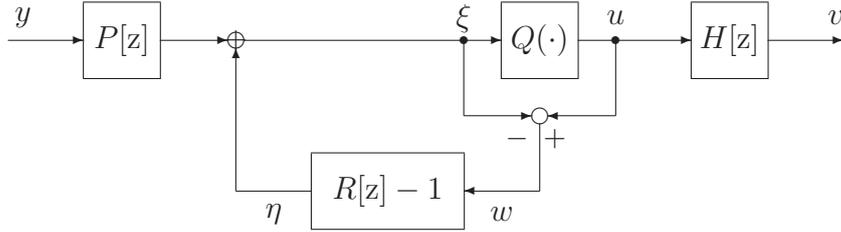


Figure 2.3: A quantizer with error feedback filter and a system $H[z]$.

2.2 Generalized Model of a $\Delta\Sigma$ Modulator

Let us consider a general linearized model of a $\Delta\Sigma$ modulator for analyzing the noise shaping characteristics and designing of the optimal noise shaping filter. We only consider the discretized single-input/single-output system with discrete-time signals. Let us denote the z -transform of a sequence $f = \{f_k\}_{k=0}^{\infty}$ as $F[z] = \sum_{k=0}^{\infty} f_k z^{-k}$ and express the output (sequence) b of the linear time-invariant (LTI) system $F[z]$ to the input $a = \{a_k\}_{k=0}^{\infty}$ as $b = F[z]a$.

Fig. 2.3 shows the error feedback configuration of a $\Delta\Sigma$ modulator. The input to the modulator is y , while the output is u . The filter $P[z]$ acts as a pre-filter to shape the frequency response of the input signal and $Q(\cdot)$ is our static quantizer. The quantization error w is filtered by $R[z] - 1$ and is fed back to y . We assume that $\lim_{z \rightarrow \infty} R[z] = 1$, i.e., the zeroth coefficient of the impulse response of $R[z]$ is 1, which implies $R[z] - 1$ is strictly causal. We also assume that

$$P[z], R[z] \in S, \quad (2.1)$$

where S denotes the set of all stable, causal, and rational transfer functions with real coefficients.

The static uniform quantizer can be described by two parameters, the quantization interval $d \in \mathbb{R}_+$ and the maximum quantization level $L \in \mathbb{Z}_+$. Here, \mathbb{R} and \mathbb{Z} denote

set of real numbers and set of integers respectively. For the continuous-valued input ξ , let the output of the static uniform quantizer be

$$Q(\xi) = \begin{cases} id, & \xi \in ((i - \frac{1}{2})d, (i + \frac{1}{2})d) \text{ and } |\xi| \leq L \\ L, & \xi > L \\ -L, & \xi < -L \end{cases}, \quad (2.2)$$

where d is the quantization interval, and i is an integer. We assume that the maximum quantization level is sufficiently large to avoid the saturation.

The difference between the input and the output of the static quantizer Q is known as a quantization error, which is denoted at time k as

$$w_k = u_k - \xi_k. \quad (2.3)$$

The quantization error is filtered by the noise shaping filter and added to the input to the static quantizer. Then, the input to the static quantizer is expressed as

$$\xi[z] = P[z]Y[z] + (R[z] - 1)W[z]. \quad (2.4)$$

Then, we have

$$U[z] = W[z] + \xi[z] = P[z]Y[z] + R[z]W[z]. \quad (2.5)$$

Here, $\xi[z]$, $Y[z]$, $W[z]$, and $U[z]$ are z -domain representations of the signals ξ, y, w and u , respectively. The gain from the input y to the output of the modulator u is known as signal transfer function (STF), while the gain between the quantization error w and the modulator output u is commonly known as NTF. In our setting, the STF and NTF for the $\Delta\Sigma$ modulator are $P[z]$ and $R[z]$, respectively.

The feedback loop acts in such a way that the quantization noise is shifted away from a certain frequency band. If the input to the modulator lies within this certain frequency band, then most of the noise due to quantization lies outside the frequency band of interest.

To design the noise shaping filter, we utilize a weighting function $H[z]$. More specifically, we consider the weighted quantization noise ϵ defined as

$$\mathcal{E}[z] = H[z]R[z]W[z], \quad (2.6)$$

where $H[z] \in S$ and $\mathcal{E}[z]$ is z-transform of the weighted quantization noise ϵ . Without loss of generality, we normalize the maximum magnitude of $H[z]$ to be unity. The weighting function is selected to reduce the effect of the quantization noise in the passband of the y . For example, when the passband of y is $[-\omega_p, \omega_p]$, we will use the weighting filter that meets $H[e^{j\omega}] \approx 1$ for $\omega \in [-\omega_p, \omega_p]$ and $|H[e^{j\omega}]|$ is small enough outside the passband to let most of the noise be outside the passband.

The output of our $\Delta\Sigma$ modulator u is connected to a system $H[z]$ whose output is denoted by v . Then, we have

$$V[z] = H[z]U[z], \quad (2.7)$$

where $V[z]$ is a z-transform of the signal v . Substituting (2.5) into (2.7), we get

$$V[z] = H[z]P[z]Y[z] + H[z]R[z]W[z]. \quad (2.8)$$

2.3 Design Problem Formulation

Our objective is to obtain the optimal filter $R[z]$ in (2.8) for a given $H[z]$ that minimizes $\mathcal{E}[z] = H[z]R[z]W[z]$ in a sense. Mathematically, we can formulate our design problem as

$$\min_{R[z] \in S} \|\epsilon\|_p \quad (2.9)$$

for a fixed pair (p, r) and a bounded input $\|w\|_r = c(> 0)$ subject to $R[\infty] = 1$. The p norm can be defined as

$$\|\epsilon\|_p = \left[\sum_{k=0}^{\infty} |\epsilon_k|^p \right]^{\frac{1}{p}}. \quad (2.10)$$

Then, using induced norms, we have

$$\|\epsilon\|_p \leq \|H[z]R[z]\|_{(p,r)} \|w\|_r. \quad (2.11)$$

Instead of directly minimizing $\|\epsilon\|_p$, we minimize the upper bound of $\|\epsilon\|_p$. For $\|w\|_r = 1$, we have $\|\epsilon\|_p \leq \|H[z]R[z]\|_{(p,r)}$. All we have to do is to find $R[z]$ that minimizes the (p, r) induced norm $\|H[z]R[z]\|_{(p,r)}$.

The signal w which is the difference between the input and the output of the static uniform quantizer satisfies

$$|w_k| \leq \frac{d}{2}. \quad (2.12)$$

Since the transfer function from w to ϵ is linear, we can put $d = 2$ without loss of generality so that $|w_k| \leq 1$ and hence $|w_k|^2 \leq 1$.

For simplification, we assume $\|w\|_r$ to have a value of unity in our calculations and simulations.

2.3.1 Induced System Norms

We consider three types of induced norms, the H_2 norm, H_∞ norm and l_1 norm of the impulse response of the system. The H_2 norm relates to the variance of the error, while the H_∞ norm corresponds to the worst-case error. The l_1 norm of the impulse response can minimize the maximum error.

H_2 Norm

The quantization error w may be modeled as a uniform random variable with zero mean and variance σ_w^2 . For analysis and synthesis, the errors at different time

are often assumed to be independent of each other, which is called the white noise assumption.

Under the white noise assumption, the mean squared error can be expressed as

$$\sigma_\epsilon^2 = \|H[z]R[z]\|_2^2 \sigma_w^2, \quad (2.13)$$

where σ_ϵ^2 denotes the variance of the weighted quantization noise ϵ , σ_w^2 is the variance of w , and $\|H[z]R[z]\|_2$ is the H_2 norm which can be defined as

$$\|G[z]\|_2 = \left[\sum_{k=0}^{\infty} |g_k|^2 \right]^{\frac{1}{2}}. \quad (2.14)$$

H_∞ Norm

The $(2, 2)$ induced norm of a system $G[z]$ is known as the H_∞ norm, which is defined as

$$\|\epsilon\|_2 \leq \|H[z]R[z]\|_\infty \|w\|_2 \quad (2.15)$$

To minimize the worst-case gain, we minimize the upper bound of $\|\epsilon\|_2$. The H_∞ norm of the function $G[z]$ can be defined as

$$\|G[z]\|_\infty = \max_{-\pi \leq \omega \leq \pi} |G[e^{j\omega}]|. \quad (2.16)$$

l_1 Norm

The l_∞ norm of the error can be also a requirement for the system. The (∞, ∞) induced norm of a system $G[z] = \sum_{k=0}^{\infty} g_k z^{-k}$ is given by

$$\|G[z]\|_{imp} = \sum_{k=0}^{\infty} |g_k|, \quad (2.17)$$

which is the l_1 norm of the impulse response of the system. Then, the l_∞ norm $\|\epsilon\|_\infty$ is bounded as

$$\|\epsilon\|_\infty = \|H[z]R[z]w\|_\infty \leq \|H[z]R[z]\|_{imp} \|w\|_\infty \quad (2.18)$$

$$= \|H[z]R[z]\|_{imp}. \quad (2.19)$$

We can reduce $\|\epsilon\|_\infty$ by minimizing $\|H[z]R[z]\|_{imp}$.

2.3.2 Stability of $\Delta\Sigma$ Modulators

Another important factor which is considered in the design of the $\Delta\Sigma$ modulator is its stability. The stability of the $\Delta\Sigma$ modulator can be ensured by limiting the amount of the feedback signal η to the quantizer. The input to the quantizer should be limited to a certain value which can avoid overloading the quantizer. The z- transform of the feedback signal η can be defined as

$$\eta[z] = (R[z] - 1)W[z] \quad (2.20)$$

By limiting the norm $\|R[z] - 1\|_q$ in (2.20), where $q \geq 1$ be a real number, to some constant value γ , we can avoid overloading the quantizer which can make the $\Delta\Sigma$ modulator unstable. The q norm of a system $G[z]$ can be defined as

$$\|G[z]\|_q = \left[\sum_{k=0}^{\infty} |g_k|^q \right]^{\frac{1}{q}}. \quad (2.21)$$

The Lee criterion [6, 22] is often utilized for the stability of the $\Delta\Sigma$ modulator. It limits the maximum magnitude of the frequency response to avoid overloading of the quantizer such that

$$\|R[z] - 1\|_\infty < \gamma. \quad (2.22)$$

The peak value of the magnitude response of $R[z]$ must be bounded to some constant value γ , where the value of γ depends on the number of quantization levels. For the case of binary quantizers, the value of γ is usually set as 1.5. However, higher order modulators are often more unstable causing the value of γ to be reduced further below 1.5 to avoid the unstable behavior of the modulator. It is observed that

reducing the value of γ also reduces the effectiveness of the noise shaping behavior of the modulator.

2.4 Conclusions

We have obtained the mathematical expression for the weighted quantization noise at the output of the $\Delta\Sigma$ modulator. The design problem is formulated for three types of most commonly used system norms. The stability of the modulator is also considered by limiting the magnitude of the input of the quantizer.

Chapter 3

Design of FIR Noise Shaping Filters for $\Delta\Sigma$ Modulators

An FIR filter of order N can be denoted as

$$R[z] = \sum_{n=0}^N r_n z^{-n}, \quad r_0 = 1. \quad (3.1)$$

Let us denote the matrices of a state-space realization of $R[z]$ by $(A_R, B_R, C_R, 1)$, where

$$A_R = \begin{bmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & 0 \end{bmatrix}, \quad B_R = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (3.2)$$

$$C_R = [r_N, r_{N-1}, \dots, r_1]. \quad (3.3)$$

It is noted that A_R and B_R are constants. Our design parameter is

$$r = [r_1, \dots, r_N], \quad (3.4)$$

which defines C_R above.

The weighted quantization noise ϵ in (2.6) to be minimized is characterized by the the composite system $H[z]R[z]$, which has to be internally stable.

Let $H[z]$ be a proper rational function, whose (A, B, C, D) matrices of a state-space realization is (A_H, B_H, C_H, D_H) . Then, one can express the state-space realization of

$H[z]R[z]$ as

$$x_{k+1} = Ax_k + Bw_k \quad (3.5)$$

$$\epsilon_k = Cx_k + Dw_k \quad (3.6)$$

where

$$A = \begin{bmatrix} A_R & B_R C_H \\ \mathbf{0} & A_H \end{bmatrix}, \quad B = \begin{bmatrix} B_R D_H \\ B_H \end{bmatrix},$$

$$C = [C_R \quad C_H], \quad D = D_H. \quad (3.7)$$

3.1 Design Based on H_2 System Norm

First of all, let us consider the minimization of the variance σ_ϵ^2 of the weighted quantization error under the white noise assumption. It is sufficient to minimize the H_2 norm of $H[z]R[z]$ to minimize σ_ϵ^2 given by (2.13).

For FIR $R[z]$, $\|H[z]R[z]\|_2^2$ can be expressed as a quadratic function of $r = [r_1, \dots, r_N]$ by using inverse Fourier transform of $|H[e^{j\omega}]|^2$ [8], which requires numerical integrations. On the other hand, a truncated impulse response of $H[z]R[z]$ is utilized in [11], where the order of some parameters is scaled by the length of the truncated impulse response. Here we adopt LMIs to evaluate the H_2 norm.

We will design optimal FIR error feedback filters, using the techniques developed for the optimal H_2 controllers in control theory. The next lemma assures that the H_2 norm can be evaluated by LMIs.

Lemma 1. ([23]) *Let $G[z]$ be a proper stable rational function, whose state-space realization is (A, B, C, D) . Then, A is Schur and*

$$\|G[z]\|_2^2 < \mu_2 \quad (3.8)$$

if and only if there exist positive definite matrices P and Z which satisfy

$$APA^T - P + BB^T \prec 0 \quad (3.9)$$

$$Z - DD^T - CPC^T \succ 0 \quad (3.10)$$

$$\text{trace}(Z) < \mu_2. \quad (3.11)$$

Using the Schur complement, one can show that (3.9) holds true if and only if

$$\begin{bmatrix} P & PA & PB \\ A^T P & P & \mathbf{0} \\ B^T P & \mathbf{0} & 1 \end{bmatrix} \succ 0. \quad (3.12)$$

Similarly, since our system has a single input and a single output, Eq. (3.10) for (A, B, C, D) can be expressed as

$$\begin{bmatrix} \mu_2 & C & D \\ C^T & P & \mathbf{0} \\ D^T & \mathbf{0} & 1 \end{bmatrix} \succ 0. \quad (3.13)$$

3.2 Design Based on H_∞ System Norm

The $(2, 2)$ induced norm of a system $G[z]$ is known as the H_∞ norm as defined in (2.15).

To minimize the worst-case gain, we utilize the bounded real lemma that provides us an LMI to evaluate the gain.

Lemma 2. ([24]) *Let $G[z]$ be a proper stable rational function, whose state-space realization is (A, B, C, D) . Then, A is Schur and*

$$\|G[z]\|_\infty^2 < \mu_\infty \quad (3.14)$$

if and only if there exists a positive definite matrix P which satisfies

$$\begin{bmatrix} A^T P A - P + C^T C & A^T P B + C^T D \\ B^T P A + D^T C & B^T P B + D^T D - \mu_\infty I \end{bmatrix} \prec 0. \quad (3.15)$$

By using the Schur complement, (3.15) can be converted into an LMI given by

$$\begin{bmatrix} -P & PA & PB & \mathbf{0} \\ A^T P & -P & \mathbf{0} & C^T \\ B^T P & \mathbf{0} & -\mu_\infty & D^T \\ \mathbf{0} & C & D & -1 \end{bmatrix} \prec 0. \quad (3.16)$$

3.3 Design Based on l_∞ Norm of Error

Unlike the H_2 norm and the H_∞ norm, only upper bounds of the H_{imp} norm are available. In [25, 26], an upper bound based on the invariant set of a discrete-time system has been utilized to design IIR error feedback filters for dynamic quantizers. The invariant set of a discrete-time system is defined as follows [27]:

Definition 1. Let $x_k \in \mathbb{R}^N$ be the state vector of the LTI system given by

$$x_{k+1} = Ax_k + Bw_k \quad (3.17)$$

where $A \in \mathbb{R}^{N \times N}$, $B \in \mathbb{R}^{N \times M}$ and $w_k \in \mathbb{R}^M$. A set X that satisfies $x_{k+1} \in X$ if $x_k \in X$ and $w_k^T w_k \leq 1$ is called an invariant set of the system given by (3.17).

The following lemma describes how to obtain an ellipsoid which is an invariant set of the system (3.17).

Lemma 3. ([27]) Let $\mathcal{E}(P)$ be the ellipsoid defined by an $N \times N$ real symmetric matrix $P \succ 0$ as $\mathcal{E}(P) = \{x \in \mathbb{R}^N : x^T P x \leq 1\}$.

The ellipsoid $\mathcal{E}(P)$ is an invariant set of the system (3.17) if and only if there exists a scalar $\alpha \in [0, 1 - \rho^2(A)]$ which satisfies

$$\begin{bmatrix} A^T P A - (1 - \alpha)P & A^T P B \\ B^T P A & B^T P B - \alpha I \end{bmatrix} \preceq 0 \quad (3.18)$$

where $\rho(A)$ is the spectrum radius of A .

If $x_k \in \mathcal{E}(P)$, then

$$\sup_{x_k \in \mathcal{E}(P)} |Cx_k|^2 = CP^{-1}C^T. \quad (3.19)$$

It follows from $|\epsilon_k| = |Cx_k + Dw_k| \leq |Cx_k| + |Dw_k|$ that

$$\|H[z]R[z]\|_{imp} \leq |CP^{-1}C^T|^{\frac{1}{2}} + |D|. \quad (3.20)$$

Thus, we can conclude that $|CP^{-1}C^T|^{\frac{1}{2}} + |D|$ is an upper bound of the norm.

Since D is constant, we minimize $CP^{-1}C^T$ with respect to α and C_R . It should be also remarked that we can assume that $\alpha \neq 0$ since our B matrix is not zero. Similarly, using the Schur complement we can express (3.18) with (A, B, C, D) as

$$\begin{bmatrix} (1-\alpha)P & \mathbf{0} & A^T P \\ \mathbf{0} & \alpha & B^T P \\ PA & PB & P \end{bmatrix} \succeq 0. \quad (3.21)$$

Moreover, we can express $CP^{-1}C^T \leq \mu$ as an LMI given by

$$\begin{bmatrix} P & C^T \\ C & \mu \end{bmatrix} \succeq 0. \quad (3.22)$$

For a fixed α , the minimization of μ is a semidefinite program, which can be numerically solved by existing optimization packages, e.g., CVX [20]. Then, all we have to do is to find α which gives the minimum. Since A is our design parameter, a line search for $\alpha \in (0, 1)$ is required to obtain the minimum. The optimal (A, B, C, D) is given by the arguments corresponding to the optimal α .

3.4 LMI for Stability Constraint

Not only the objective function but also the condition (2.22) on the stability can be described by LMIs. For example, as shown in [10], it follows from Lemma 2 that the Lee criterion

$$\|R[z] - 1\|_{\infty} < \gamma \quad (3.23)$$

is satisfied if and only if there exists a positive definite matrix P_R which meets

$$\begin{bmatrix} -P_R & P_RA_R & P_RB_R & \mathbf{0} \\ A_R^T P_R & -P_R & \mathbf{0} & C_R^T \\ B_R^T P_R & \mathbf{0} & -\gamma^2 & 1 \\ \mathbf{0} & C_R & 1 & -1 \end{bmatrix} \prec 0. \quad (3.24)$$

Thus, if one would like to design the FIR noise shaping filter that minimizes σ_ϵ^2 under the Lee criterion, it suffices to solve the following convex optimization problem

$$\min_{r_1, \dots, r_N} \mu_2 \quad (3.25)$$

subject to (3.12), (3.13), and (3.24).

In summary, our unified approach enables the design of the FIR noise shaping filter to minimize the H_2 , the H_∞ , or the l_1 system norm under the H_2 , the H_∞ , or the l_1 norm constraint. Moreover, since norms are described by LMIs, different types of problems can be solved numerically. For example, some signal processing applications may require us to design an error feedback filter for a $\Delta\Sigma$ modulator by adding a constraint that limits the magnitude of the weighted quantization noise to a certain value. Then, our design objective is to design the noise shaping filter that attains the optimal value of the stability threshold γ under the maximum weighted quantization noise constraint. If we adopt the Lee criterion, we can obtain a stable error feedback filter by minimizing (3.23) subject to $\|\epsilon\|_\infty \leq c$, where c is the maximum bound on the weighted quantization noise ϵ , by using LMIs in (3.21), (3.22), and (3.24).

3.5 Design Examples

In this section, simulations for lowpass and bandpass $\Delta\Sigma$ modulators have been shown by using the proposed design method based on H_2 , H_∞ , and l_1 system norms. For the design of a conventional $\Delta\Sigma$ modulator by NTF zero optimization method [6], the DELSIG toolbox [28] is utilized to obtain the frequency response of an IIR noise

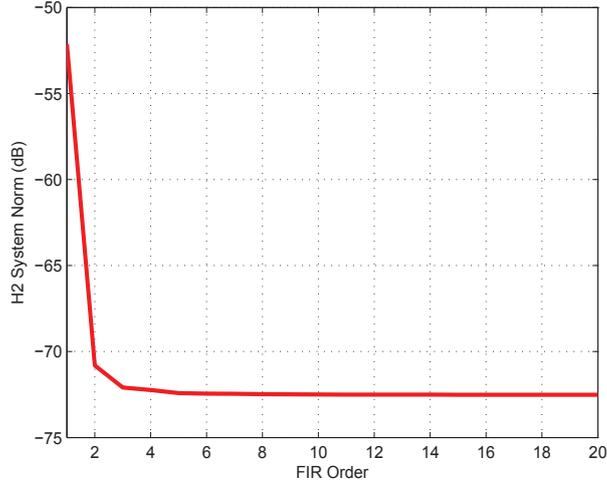


Figure 3.1: The H_2 norm of $H[z]R[z]$ as a function of order of $R[z]$ for the first order lowpass weighting function, where $R[z]$ is designed based on the H_2 norm.

shaping filter with `synthesizeNTF` MATLAB function. The frequency response and the noise shaping characteristics of the FIR feedback filter proposed in [10] are also compared with our designed filters.

3.5.1 Lowpass $\Delta\Sigma$ Modulator with the 1st Order $H[z]$

Now, let us design a lowpass $\Delta\Sigma$ modulator by using a first order lowpass Butterworth filter as our weighting function $H[z]$. The first order Butterworth filter provides us the maximum flat response in the passband at the expense of a wide transition band as the filter changes from the passband to stopband. The input signal y to the lowpass $\Delta\Sigma$ modulator is assumed to be oversampled with the an OSR of 512. Then, the cut-off frequency of the first order Butterworth filter is set at $\pi/\text{OSR} \approx 0.0061$ in the normalized angular frequency interval $[0, \pi]$.

For the stability of the $\Delta\Sigma$ modulator, we assume the value of the Lee coefficient γ to be 1.5 which is equivalent to 3.52 in decibels (dB), however, the value of γ can be increased further as long as the $\Delta\Sigma$ modulator remains stable.

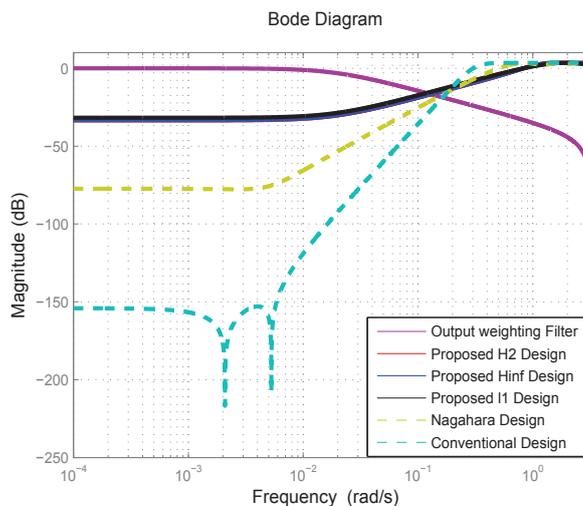


Figure 3.2: Frequency responses of filters designed by the proposed method and the referenced methods. The weighting function is of order unity.

The order of the FIR feedback filter $R[z]$ is chosen based on the convergence behavior of the objective function. Fig. 3.1 shows that the H_2 norm of $H[z]R[z]$ reaches a value as we keep on increasing the order of FIR filter. Above the FIR order 8, the norm of the weighted quantization noise remains almost constant in terms of the H_2 norm. In this example, the FIR feedback filter $R[z]$ for noise shaping is set to be 8.

Fig. 3.2 depicts the frequency responses of H_2 , H_∞ , and l_1 norm based filters compared with the referenced methods in [6] and [10]. The order of FIR feedback filter in [10] is also chosen to be 8, while the order of IIR feedback filter for conventional design [6] is set to be 4. Our designed FIR filters have almost the same frequency response. It can be observed that the frequency responses of our designed FIR filters have uniform attenuation in the low frequency region of frequency spectrum, while the conventional design shows a peak in the magnitude response near the cut-off frequency.

To precisely see the difference between magnitude responses of our designed filters,

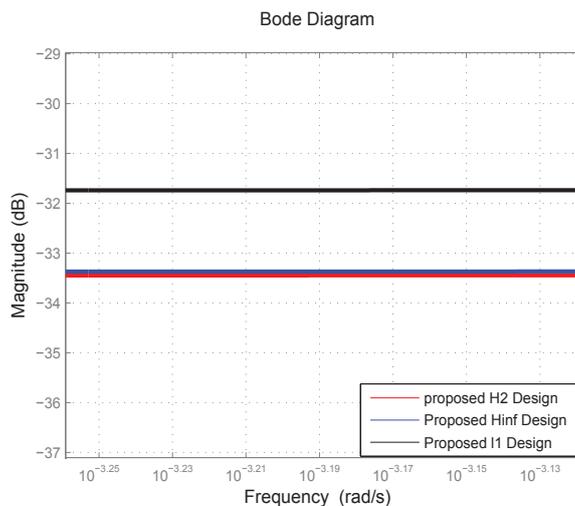


Figure 3.3: Enlarged frequency response of our proposed filters in Fig. 3.2.

Table 3.1: $\|H[z]R[z]\|_2$, $\|H[z]R[z]\|_\infty$, and l_1 norm of the impulse response of $H[z]R[z]$ for the first order lowpass weighting function.

	H_2 norm	H_∞ norm	l_1 norm
H_2 norm design	1.54×10^{-2}	2.19×10^{-2}	2.62×10^{-2}
H_∞ norm design	1.54×10^{-2}	2.16×10^{-2}	2.59×10^{-2}
l_1 norm design	1.63×10^{-2}	2.59×10^{-2}	2.59×10^{-2}
Nagahara design [10]	1.92×10^{-2}	3.82×10^{-2}	4.89×10^{-2}
Conventional design [6]	2.61×10^{-2}	6.92×10^{-2}	11×10^{-2}

the enlarged view of Fig. 3.2 is shown in Fig. 3.3. The FIR filters based on H_2 and H_∞ norms exhibit almost equivalent attenuation and similar behavior, while l_1 based design exhibits slightly lower attenuation as compared to H_2 and H_∞ based designs.

The method in [10] designs the FIR noise shaping filter based on the weighted H_∞ norm of $R[z]$. Near the cut-off frequency, the magnitude response of the FIR filter in [10] increases rapidly showing the high steepness in the transition band, while all of our proposed filters exhibit good performance, matching the steepness of the weighting function. Note that, the maximum magnitude value of all filters are bounded to 3.52 dB approximately due to stability constraint which utilizes the Lee coefficient $\gamma = 1.5$.

Table 3.1 lists the H_2 norm $\|H[z]R[z]\|_2$, the H_∞ norm $\|H[z]R[z]\|_\infty$, and the l_1 norm of the impulse response of $H[z]R[z]$ for our designed FIR filters compared with the referenced designs in [6] and [10]. All three designed filters have less H_2 , H_∞ , and l_1 norms as compared with optimal feedback filters in [6] and [10]. Although the referenced designs have lower gains in the passband as observed in Fig. 3.2, our designed filters have better performance in the weighted norms. This is because the referenced designs only take into account the passband, while our design does the whole band. Indeed, if an ideal lowpass filter can be used as our weighting function, our H_∞ norm based filter is equivalent to the weighted H_∞ norm based filter in [10]. Since any ideal lowpass filter is not available in practice, it is important to consider the noise in the stopband. Our method can trade off the properties of the noise shaping filter in the passband and the stopband using an appropriate weighting function.

The H_∞ and l_1 norm designs exhibit an equivalent l_1 norm, while H_2 and H_∞ norm designs have an equivalent H_2 norm. This may be partially due to the implementation and the numerical errors in our numerical optimization. It should be noted that we minimize the upper bounds, which implies that we can not guarantee that the quantizer designed based on a norm is optimal in the sense of the norm.

3.5.2 Lowpass $\Delta\Sigma$ Modulator with the 4th Order $H[z]$

Now let us introduce a higher order lowpass Butterworth filter of order 4 as our weighting function, where OSR is 32. The maximum magnitude of NTF is limited to 3.52 dB by using the Lee coefficient $\gamma = 1.5$. The fourth order Butterworth filter with a cut-off frequency of $\pi/\text{OSR} \approx 0.0098$ has better stopband attenuation than the first order Butterworth filter by increasing the steepness of passband to stopband transition at the cost of reduced passband flatness.

For this lowpass $\Delta\Sigma$ modulator, Fig. 3.4 shows the convergence behavior of the

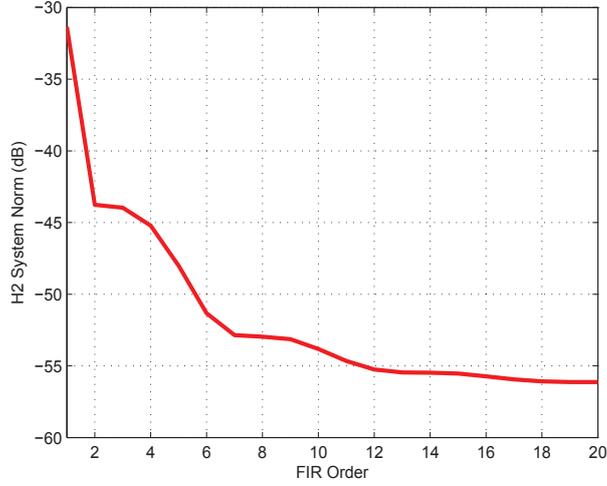


Figure 3.4: H_2 norm of $H[z]R[z]$ as a function of order of $R[z]$ for the fourth order lowpass weighting function, where $R[z]$ is designed based on the H_2 norm.

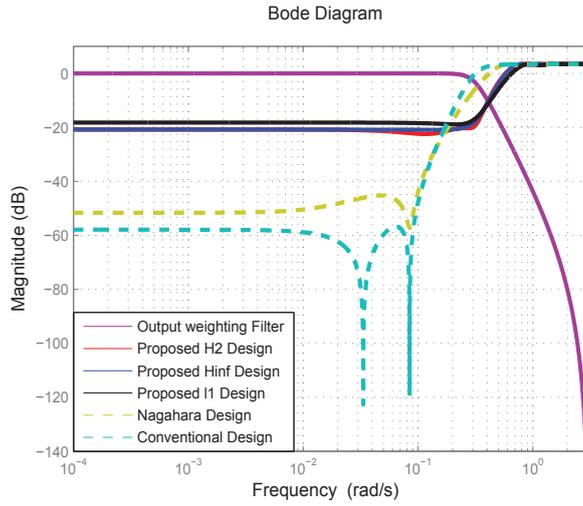


Figure 3.5: Frequency responses of filters designed by the proposed method and the referenced methods. The weighting function is of order 4.

H_2 norm of $H[z]R[z]$ for the H_2 norm based design. From this, the FIR feedback filter of order 20 is chosen for proposed designs and referenced design in [10], while the IIR feedback filter for conventional design [6] is of order 4.

In Fig. 3.5, we give the frequency responses of proposed H_2 , H_∞ , and l_1 norm

Table 3.2: $\|H[z]R[z]\|_2$, $\|H[z]R[z]\|_\infty$, and l_1 norm of the impulse response of $H[z]R[z]$ for the fourth order lowpass weighting function.

	H_2 norm	H_∞ norm	l_1 norm
H_2 norm design	3.95×10^{-2}	9.71×10^{-2}	1.40×10^{-1}
H_∞ norm design	4.07×10^{-2}	9.09×10^{-2}	1.24×10^{-1}
l_1 norm design	4.43×10^{-2}	1.22×10^{-1}	1.23×10^{-1}
Nagahara design [10]	9.18×10^{-2}	3.53×10^{-1}	4.74×10^{-1}
Conventional design [6]	1.49×10^{-1}	6.69×10^{-1}	9.01×10^{-1}

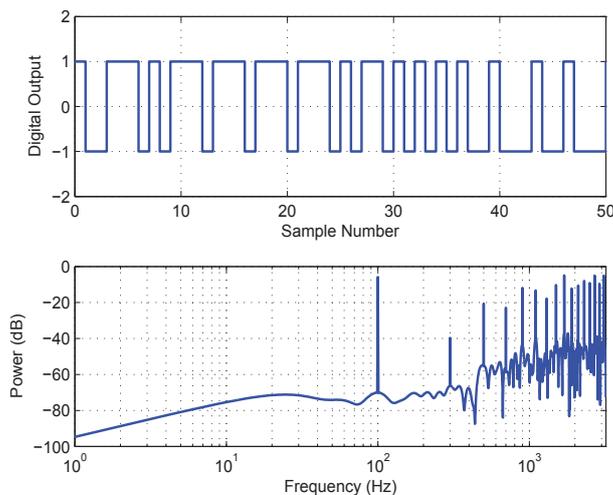


Figure 3.6: Output and frequency spectrum plot of the lowpass $\Delta\Sigma$ modulator obtained by the proposed H_2 norm based design.

based filters compared with referenced methods. Our proposed designs show better performance by providing uniform attenuation in the low frequency region, and exhibiting better magnitude responses near the cut-off frequency as compared to the referenced methods in [10] and [6].

Table 3.2 shows the H_2 norm, the H_∞ norm, and the l_1 norm of the impulse response of $H[z]R[z]$ for our designed FIR filters compared with the referenced designs in [6] and [10]. It can be observed that all three designed filters have less H_2 , H_∞ and, l_1 norms than the optimal feedback filters in [6] and [10]. The H_2 , H_∞ and, l_1 norm designs have the least H_2 , H_∞ and l_1 norms, respectively.

To assess the performance of the lowpass $\Delta\Sigma$ modulator with an error feedback filter obtained by our proposed H_2 norm based design, the MATLAB function `simulateDSM` in DELSIG toolbox [28] is used to simulate the $\Delta\Sigma$ modulator for obtaining the digital output. The input to the $\Delta\Sigma$ modulator is a sinusoidal wave with frequency 100 Hz and amplitude 0.5. We assume a uniform quantizer with maximum quantization levels $L = 2$ and quantization interval $d = 2$.

The output of this uniform quantizer is a digital signal which is represented by using +1 and -1 volts for binary 0 and 1 respectively, which is shown in the upper part of Fig. 3.6. The lower part of Fig. 3.6 is the frequency spectrum of the digital output, which gives the performance of our lowpass $\Delta\Sigma$ modulator. Our lowpass $\Delta\Sigma$ modulator attenuates the quantization noise in the frequency region which contains the information signal. The frequency notch for the input signal appears at 100 Hz, which is the same with the sinusoidal wave, and the magnitude of quantization noise is low in the passband. Our proposed H_2 filter efficiently shifts the quantization noise towards the high frequency region which does not carry much information. Similar results can be found for H_∞ and l_1 norm based designs, which are omitted.

3.5.3 Bandpass $\Delta\Sigma$ Modulator with the 6th Order $H[z]$

Finally, we adopt a 6th order bandpass Butterworth filter as our weighting function, whose frequency response is found in Fig. 3.7.

The input to the modulator is assumed to have the center frequency $\omega_o = \pi/2$ and bandwidth parameter $\Omega = \pi/16$. For the passband $\omega \in [\pi/2 - \pi/16, \pi/2 + \pi/16]$, we use the bandpass Butterworth filter that meets $H[e^{j\omega}] \approx 1$ for $\omega \in [\omega_o - \Omega, \omega_o + \Omega]$ and $|H[e^{j\omega}]|$ is small enough outside the passband to let most of the noise be outside the passband. For the conventional design [6], OSR is set to be 16.

As illustrated in Fig. 3.8 the H_2 norm of $H[z]R[z]$ for H_2 norm based design

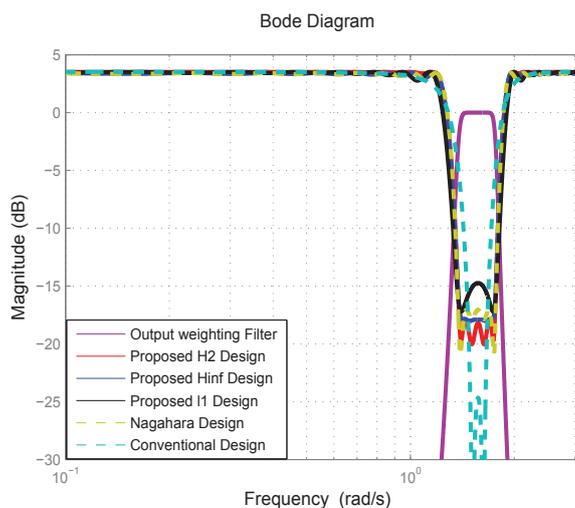


Figure 3.7: Frequency responses of filters designed by the proposed method and the conventional method. The weighting function is of order 6.

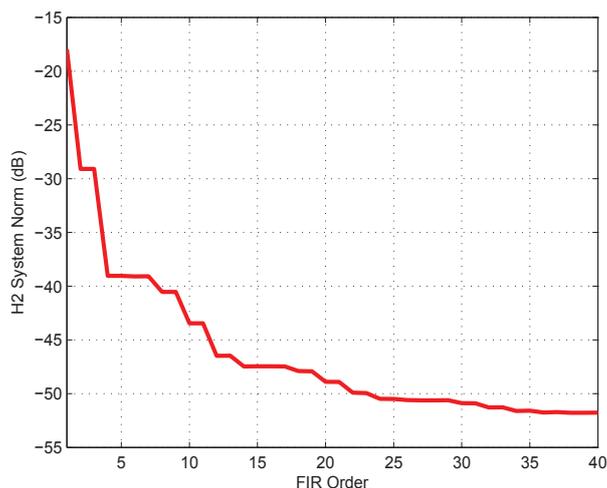


Figure 3.8: H_2 norm of $H[z]R[z]$ as a function of order of $R[z]$ for the sixth order bandpass weighting function, where $R[z]$ is designed based on the H_2 norm.

converges slowly compared to the previous examples. A longer order is required to adjust to the 6th order bandpass Butterworth filter. Thus, the order of proposed FIR feedback filters $R[z]$ is chosen to be 40. The order of FIR feedback filter in [10] is also set to be 40. For the conventional bandpass $\Delta\Sigma$ modulator [6], the order of IIR

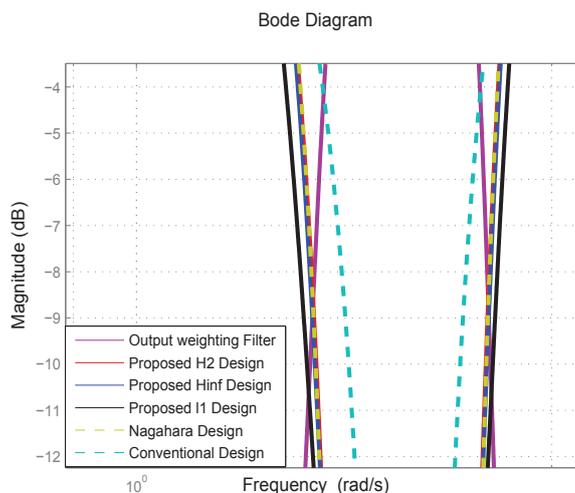


Figure 3.9: Enlarged frequency response of our proposed filters in Fig. 3.7.

feedback filter is 4, whereas, the center frequency is $f_o = 1/4$.

We compare the frequency responses of our proposed FIR feedback filters for the bandpass $\Delta\Sigma$ modulator with the referenced designs in [6] and [10]. Fig. 3.7 shows that the magnitude responses of proposed H_2 and H_∞ design FIR filters have higher attenuation levels as compared to the method proposed in [10]. Again, the magnitude responses of our proposed design filters are uniformly attenuated over the passband, while the conventional design shows a peak near the edges of the band which can be observed in Fig. 3.9.

Table 3.3 gives the H_2 norm, the H_∞ norm, and the l_1 norm of the impulse response of $H[z]R[z]$ for our designed FIR filters compared with the referenced designs in [6] and [10]. Again, our proposed H_2 , H_∞ and, l_1 norm designs have the least H_2 , H_∞ and l_1 norms, respectively.

3.5.4 Stability Under the l_∞ Norm Constraint

Here, to obtain the most stable error feedback filter for a lowpass $\Delta\Sigma$ modulator, we minimize (3.23) under the l_∞ norm constraint on the weighted quantization noise

Table 3.3: $\|H[z]R[z]\|_2$, $\|H[z]R[z]\|_\infty$, and l_1 norm of the impulse response of $H[z]R[z]$ for the sixth order bandpass weighting function.

	H_2 norm	H_∞ norm	l_1 norm
H_2 norm design	5.08×10^{-2}	1.385×10^{-1}	2.094×10^{-1}
H_∞ norm design	5.38×10^{-2}	1.277×10^{-1}	1.916×10^{-1}
l_1 norm design	6.08×10^{-2}	1.833×10^{-1}	1.858×10^{-1}
Nagahara design [10]	5.45×10^{-2}	1.408×10^{-1}	2.222×10^{-1}
Conventional design [6]	10.19×10^{-2}	4.253×10^{-1}	5.461×10^{-1}

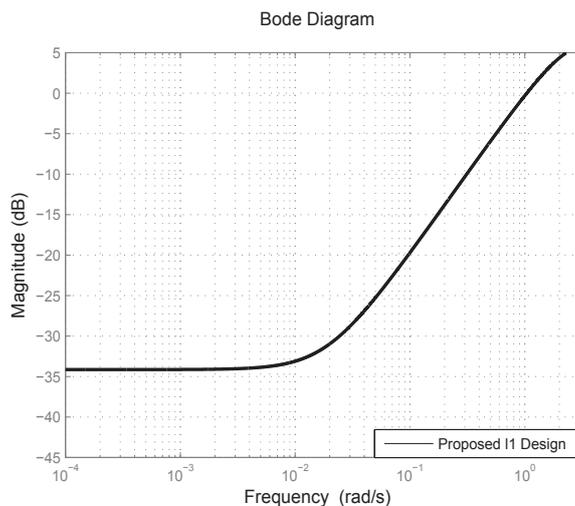


Figure 3.10: Frequency response of the error feedback filter designed by minimizing the upper bound of the Lee coefficient under the constraint on the l_∞ norm of the weighted quantization noise.

such that $\|\epsilon\|_\infty = 1.96 \times 10^{-2}$. We use the same first order Butterworth filter in Section 3.5.1.

The minimum magnitude value of the in-band quantization noise is -34.2 dB. The obtained upper bound of the Lee criterion is $\gamma = 1.92$, which is equivalent to 5.7 dB. It is larger than 1.5 used the l_1 norm design in Table 3.1, since we impose a slight tighter constraint on the $\|\epsilon\|_\infty = 1.96 \times 10^{-2}$ than 2.59×10^{-2} in Table 3.1. The frequency response of the designed feedback filter is illustrated in Fig. 3.10.

3.6 Conclusions

We have proposed a unified design method of FIR noise shaping filters of $\Delta\Sigma$ modulators based on H_2 , H_∞ , and l_1 norms. The minimization of the norm of the weighted quantization error is cast into a convex optimization problem by using LMIs, which can be efficiently and numerically solved. To ensure the stability of a $\Delta\Sigma$ modulator, we have also included LMI constraints which subsumes the Lee criterion. Our results show that the frequency response of our filters exhibits good performance throughout the low frequency region providing uniform attenuation and matching the weighting function. Also, our proposed H_2 , H_∞ , and l_1 norm designed error feedback filters are shown to provide us with minimum H_2 , H_∞ , and l_1 norms of weighted quantization error, respectively, which shows the effectiveness of our proposed design methods.

Chapter 4

Design of IIR Noise Shaping Filters for $\Delta\Sigma$ Modulators

The design problem for an IIR filter becomes non-convex, and the solution for this non-convex problem is not guaranteed to be optimal unlike FIR design problem in the previous chapter. In the method proposed in [13], the filter is assumed to have an IIR which is converted to a minimization problem by virtue of generalized GKYP lemma and solved by using an iterative algorithm. The limitation to this method is that it only addresses the H_∞ norm based merit factor and does not take into account the non-ideal behavior of the filter at the output of a $\Delta\Sigma$ modulator.

In this chapter, we propose several design methods which can solve the non-convex design problem for obtaining the IIR noise shaping filter. In Section 4.1, we utilize the extended LMI technique to obtain IIR noise shaping filters. The order of the obtained IIR noise shaping filters is constrained to be identical to the order of the weighting function.

Section 4.2 introduces the two commonly used approximation techniques to obtain sub-optimal IIR noise shaping filters in the feedback of the $\Delta\Sigma$ modulator. The Yule-Walker method and the least-squares (LS) approximation are utilized to approximate the high order FIR filters with the lower order IIR filters.

In Section 4.3, the hybrid technique is adopted which utilizes the FIR coefficients in the numerator and the `synthesizeNTF` denominator coefficients in its denominator to obtain IIR noise shaping filters.

Finally in Section 4.4, we propose an iterative LMI technique which can outperform the existing method to obtain a near optimal IIR noise shaping filter in the feedback of a $\Delta\Sigma$ modulator.

We will perform simulations and compare the proposed methods with the existing methods to show the effectiveness of each proposed method.

4.1 Design Based on the Extended LMI Technique

Again, let us consider the the linearized error feedback model of a $\Delta\Sigma$ modulator as shown in Fig. 2.3. Now, our objective is to design a stable noise shaping IIR filter that minimizes the effect of quantization, which is expressed by (2.6). To avoid overloading the quantizer and/or maintain the stability of the quantizer, we also impose a constraint on the error feedback signal η in (2.20).

The transfer function of an IIR filter $R[z]$ can be written as

$$R[z] = \frac{B[z]}{A[z]}, \quad (4.1)$$

where the numerator $B[z]$ is given by

$$B[z] = \sum_{n=0}^{N_B} b_n z^{-n}, \quad (4.2)$$

and the denominator $A[z]$ is

$$A[z] = \sum_{n=0}^{N_A} a_n z^{-n}. \quad (4.3)$$

The order of the IIR filter is the highest power of either the denominator or the numerator polynomial, $N = \max(N_B, N_A)$. To ensure at-least one sample delay in

the feedback filter $R[z] - 1$ of the $\Delta\Sigma$ modulator for practical realization, we have to choose the value of the first coefficient of both the numerator and denominator as $b_0 = a_0 = 1$.

Let us denote the state-space matrices of IIR filter $R[z]$ by $(A_R, B_R, C_R, 1)$, where

$$A_R = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_N & -a_{N-1} & -a_{N-2} & \dots & -a_1 \end{bmatrix}, \quad B_R = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (4.4)$$

$$C_R = [d_N, d_{N-1}, \dots, d_1], \quad (4.5)$$

with $d_i = b_i - a_i$.

Since the quantization noise consists of a composite system $H[z]R[z]$, the state-space matrices (A, B, C, D) of this composite system can be expressed using (3.7).

4.1.1 Numerical Design of IIR Noise Shaping Filters

We consider the minimization of the weighted quantization noise in (2.6) under the stability constraint (2.22). More specifically, our minimization problem can be defined as follows:

$$\min_{R[z] \in \mathcal{S}} \gamma_\epsilon \quad (4.6)$$

subject to $R[\infty] = 1$ and

$$\|H[z]R[z]\|_{(p,r)} < \gamma_\epsilon \quad (4.7)$$

$$\|R[z] - 1\|_\infty < \gamma_\eta. \quad (4.8)$$

We consider two important induced norms to design IIR filters, the H_2 norm and H_∞ norm, which can be evaluated by using state-space expressions and LMIs. As defined in Chapter 2, the H_2 norm relates to the variance of the error, while the H_∞ norm corresponds to the worst-case error.

H_2 Norm

Under this white noise assumption [2], the variance of the error ϵ (2.13) can be evaluated by using

$$\|H[z]R[z]\|_2^2 = \sum_{k=0}^{\infty} \|CA^k B\|_2^2 + DD^T. \quad (4.9)$$

If A is Schur [29], i.e., all the eigenvalues of A lie in the unit circle, then there exists a positive semi-definite solution P of the discrete Lyapunov equation defined as

$$P = A^T P A + B B^T \quad (4.10)$$

and the squared H_2 norm is given by

$$\|H[z]R[z]\|_2^2 = C P C^T + D D^T. \quad (4.11)$$

The squared H_2 norm (4.11) is expressed in matrix inequalities (3.12),(3.13) using Lemma 1 in Chapter 3.

For IIR filters, the constraint (3.12) is a BMI, since it contains the products of the variables and Lyapunov matrix. On the other hand, the constraint (3.13) is an LMI, which is convex. In general, BMIs are not convex and NP-hard to solve numerically. However, by using the change of variables proposed independently in [30] and [31], we can convert a BMI to a convex LMI that can be evaluated numerically.

The variance of the feedback signal η is given under the white noise assumption by

$$E\{|\eta|^2\} = \|R[z] - 1\|_2^2, \quad (4.12)$$

which can be expressed as

$$\sum_{k=1}^{\infty} \|\tilde{C} A^k B\|_2^2 \quad (4.13)$$

where

$$\tilde{C} = [\mathbf{0} \quad C_R]. \quad (4.14)$$

Then, $\|R[z] - 1\|_2 < \gamma_\eta$ if and only if there exists a positive definite matrix P that satisfies (3.9) and

$$\begin{bmatrix} \mu_\eta & \tilde{C} \\ \tilde{C}^T & P \end{bmatrix} \succ \mathbf{0} \quad (4.15)$$

where

$$\mu_\eta = \gamma_\eta^2. \quad (4.16)$$

Since the objective function and constraint can be evaluated by LMIs, we can state the following lemma:

Lemma 4. *Consider a quantizer with a noise shaping filter as shown in Fig 2.3. The quantization error w of the static quantizer $Q(\cdot)$ is assumed to be white and independent of the input y . Then, the optimal filter $R[z]$ that minimizes the variance of ϵ given in (2.13) under a constraint on the variance of the error feedback signal can be found by solving a convex optimization problem.*

The proof of this lemma is as follows:

Let us see the change of variables, following the notations in [30].

Let the order of $H[z]$ be N . The set of $N \times N$ positive definite matrices is denoted as $PD(N)$. We define the following matrices $\{P_f, P_g, W_f, W_g, W_h, L\}$, where $P_f \in PD(N)$, $P_g \in PD(N)$, $W_f \in \mathbb{R}^{1 \times N}$, $W_g \in \mathbb{R}^{N \times 1}$, $W_h \in \mathbb{R}$, $L \in \mathbb{R}^{N \times N}$, with P_f and P_g . Let us also define matrices from $\{P_f, P_g, W_f, W_g, W_h, L\}$ as

$$P^{-1} = \begin{bmatrix} P_f & S_f \\ S_f & S_f \end{bmatrix} \quad (4.17)$$

$$U = \begin{bmatrix} P_f & I_N \\ S_f & \mathbf{0} \end{bmatrix} \quad (4.18)$$

$$P_g = (P_f - S_f)^{-1} \quad (4.19)$$

and the matrices $\{M_A, M_B, M_C, M_P\}$ as

$$M_A = \begin{bmatrix} A_p P_f + B_p W_f & A_p \\ L & P_g A_p \end{bmatrix} \quad (4.20)$$

$$M_B = \begin{bmatrix} B_p \\ W_g \end{bmatrix} \quad (4.21)$$

$$M_C = [C_p P_f + D_p W_f \quad C_p] \quad (4.22)$$

$$M_P = \begin{bmatrix} P_f & I_n \\ I_n & P_g \end{bmatrix} \quad (4.23)$$

We can observe that if the matrices $\{A_R, B_R, C_R\}$ are given by

$$A_R = [B_p P_f - P_g^{-1}(L - P_g A_H P_f)] P_f^{-1} \quad (4.24)$$

$$B_R = B_p - P_g^{-1} W_g \quad (4.25)$$

$$C_R = W_f P_f^{-1} \quad (4.26)$$

then $\{A, B, C\}$ satisfy

$$M_A = U^T P A U \quad (4.27)$$

$$M_B = U^T P B \quad (4.28)$$

$$M_C = C U \quad (4.29)$$

$$M_P = U^T P U. \quad (4.30)$$

Multiplying both sides of (4.12) with the transformation matrix $\Phi = \text{diag}(U, U, 1)$ from the right hand side and Φ^T from the left hand side leads to

$$\begin{bmatrix} M_P & M_A & M_B \\ M_A^T & M_P & 0 \\ M_B^T & 0 & 1 \end{bmatrix} \succ \mathbf{0}. \quad (4.31)$$

Similarly, with $\text{diag}(1, U, 1)$, (3.13) can be transformed into

$$\begin{bmatrix} \mu_\epsilon & M_C & D^T \\ M_C^T & M_P & 0 \\ D & 0 & 1 \end{bmatrix} \succ \mathbf{0}, \quad (4.32)$$

where $\mu_\epsilon = \gamma_\epsilon^2$. Moreover, since

$$\tilde{C}U = [C_R S_f \quad \mathbf{0}] = [W_f \quad \mathbf{0}] := M_{\tilde{C}}, \quad (4.33)$$

the constraint (4.15) is converted by the transformation matrix $\text{diag}(1, U)$ into

$$\begin{bmatrix} \gamma_\epsilon^2 & M_{\tilde{C}} \\ M_{\tilde{C}}^T & M_P \end{bmatrix} \succ \mathbf{0} \quad (4.34)$$

Since (4.31), (4.32), and (4.34) are convex LMIs, the minimization of γ_ϵ^2 subject to (4.31), (4.32), and (4.34) is a convex optimization. Thus, there exists one global minimum. From the optimal solution, we can reconstruct (A_R, B_R, C_R) with (4.24), (4.25), and (4.26), which concludes the proof.

H_∞ Norm

The worst-case norm is often utilized to capture the effect of a deterministic error. The maximum (worst-case) gain is the H_∞ norm of the system when a finite-energy signal is applied to it. The H_∞ norm of the system is characterized into a BMI (3.16) using the Lemma 2 in Chapter 3.

Now let us consider the H_2 and the H_∞ norm for (4.7) and (4.8). Each of them can be characterized by a BMI of the variables and a Lyapunov matrix. In this case, we cannot apply an identical change of variables to different BMIs since the change of variables depends on the Lyapunov matrix and the BMIs do not share an identical Lyapunov matrix in general. If we force all the Lyapunov matrices to be identical, then BMIs can be converted into LMIs but the global solution cannot be obtained because of the additional constraint on the Lyapunov matrices. This is known as the Lyapunov sharing paradigm [31].

To obtain better designs, we utilize the extended H_2 and H_∞ norm characterizations with an extra matrix G , which is described as an extended LMI technique in [32].

Lemma 5 (Extended H_2 characterization[32]). *Let $F[z]$ be a proper stable rational function whose state space expression is (A, B, C, D) . Then, the inequality $\|F[z]\|_2^2 < \mu_2$ holds if and only if there exists a matrix G and a symmetric matrix P such that*

$$\begin{bmatrix} G + G^T - P & GA & GB \\ A^T G & P & \mathbf{0} \\ B^T G & \mathbf{0} & 1 \end{bmatrix} \succ 0 \quad (4.35)$$

and

$$\begin{bmatrix} \mu_2 & C & D \\ C^T & P & \mathbf{0} \\ D^T & \mathbf{0} & 1 \end{bmatrix} \succ 0. \quad (4.36)$$

Lemma 6 (Extended H_∞ characterization([32])). *Let $F[z]$ be a proper stable rational function whose state space expression is (A, B, C, D) . Then, the inequality $\|F[z]\|_\infty^2 < \mu_\infty$ holds if and only if there exists a matrix G and a symmetric matrix P such that*

$$\begin{bmatrix} G + G^T - P & GA & GB & \mathbf{0} \\ A^T G & P & \mathbf{0} & C^T \\ B^T G & \mathbf{0} & 1 & D^T \\ \mathbf{0} & C & D & \mu_\infty \end{bmatrix} \succ 0. \quad (4.37)$$

The substitution of the state-space expression in (4.35) and (4.37) will give us BMIs as the products of system matrices and instrumental variables P and G . Similar to the BMIs in (3.12) and (3.16), the BMIs in (4.35) and (4.37) can be converted into LMIs by using a change of variables as developed in [32].

After the change of variables, if we allow the Lyapunov matrices of the two LMIs to be identical, then the problem becomes convex and can be solved numerically. The performance of the filter with the augmented variable G is expected to be better than the filter designed with the original LMIs without G . It should be noted that if one puts $G = P$ in (4.35) and (4.37), then (4.35) and (4.37) are reduced to (3.12) and (3.16). This implies that at least in theory, the design based on the extended LMIs never produces results worse than the design based on the original LMIs. However, even with the extended LMIs, global optimality cannot be guaranteed. It should

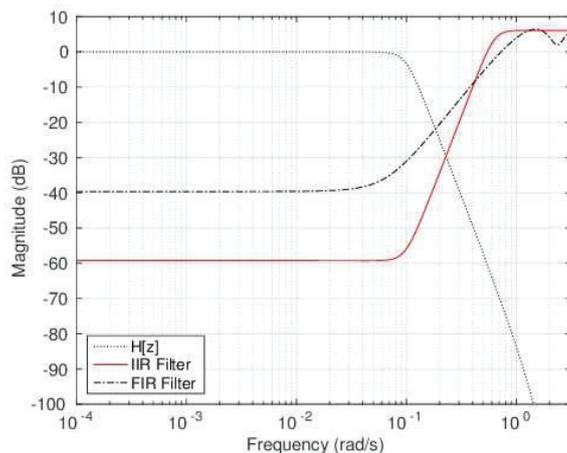


Figure 4.1: Frequency responses of H_2 norm based noise shaping filters subjected to the variance of the feedback signal (Example 1).

also be noted that the numerical solution to the optimization problem using the extended LMIs is not necessarily better than or equal to the numerical solution to the optimization problem using the conventional LMIs due to numerical errors.

4.1.2 Design Examples

In this section, we provide design examples to demonstrate the effectiveness of our proposed H_2 and H_∞ norm based IIR noise shaping filters. The order of the IIR noise shaping filter $R[z]$ is set to be equal to the order of the system $H[z]$.

Example 1: H_2 norm based noise shaping filter constrained by the feedback error variance

To begin with, let us consider the H_2 norm based design constrained by the variance of the feedback signal, which can be cast into a convex optimization problem as stated in Lemma 4.

The system $H[z]$ is a low pass Butterworth filter of order 4 with a cutoff frequency of $\pi/32$. Since the lowpass Butterworth filter has zeros at -1 , we cannot use its inverse for the noise shaping filter.



Figure 4.2: MSE of optimal FIR noise shaping filters as a function of their order.

We minimize the variance of the error ϵ under a constraint on the variance of the feedback signal given by (4.12). Since the variance σ_w^2 of the quantization error w of the static uniform quantizer is just a scalar in our optimization, we put $\sigma_w^2 = 1$. Then, the constraint on the variance of the feedback signal is $\|R[z] - 1\|_2 < 1.5$.

Our aim is to obtain an IIR filter that minimizes μ_ϵ under (3.12), (3.13), and (4.15) with $\mu_\eta = (1.5)^2$. As shown in the proof of Lemma 5, the design problem is equivalent to the minimization of μ_ϵ under the LMIs (4.31), (4.32), and (4.34). We numerically solve the problem by CVX to find the optimal (P_f, P_g, W_f, W_g, L) in (M_A, M_B, M_C, M_P) given by (4.20), (4.21), (4.22), and (4.23). Then, we reconstruct the matrices (A_R, B_R, C_R) of the IIR filter from the optimal (P_f, P_g, W_f, W_g, L) by using (4.24), (4.25), and (4.26).

On the other hand, the optimal FIR noise shaping filter can be designed directly from (3.12), (3.13) and (4.15) [14] since A and B are constants, hence the inequality (3.12) becomes an LMI.

Fig. 4.1 depicts the frequency responses of the designed IIR and FIR filters under the same constraint, where the order of the FIR filter is chosen to be 4 so that it is

identical to the order of the IIR filter. The designed IIR filter has a lower response than the designed FIR filter at the passband of the system. Then, the MSE of our designed IIR filter is calculated to be -66.12 dB, while the designed FIR filter gives us -49.29 dB. The designed IIR filter performs better than the designed FIR filter having the same order. This highlights the importance of the design of the IIR filter.

If the optimal filter is an IIR filter, then it is expected that the MSE of the optimal FIR noise shaping filter decreases as its order increases and then converges to the MSE of the optimal IIR filter.

Fig. 4.2 shows the MSE of the designed FIR filters with different orders. The MSE of the designed FIR filter monotonically decreases as a function of its order and converges. The MSE of the designed FIR filter of order 15 is at most -58 dB, which is larger than the MSE of the designed IIR filter.

Example 2 : Extended vs. Non-extended

Next, we compare the extended LMI-based design with the conventional non-extended LMI-based design that imposes an additional constraint by forcing the different Lyapunov matrices in the BMIs to be identical. Let us consider two types of error feedback filters to minimize the H_2 or the H_∞ norm of the error of the signal-of-interest subjected to the Lee criterion for ensuring the stability of the quantizer.

The Lee criterion is the H_∞ norm of the noise shaping filter as described in (3.23). We can evaluate this with (3.24) for the non-extended LMI-based design and (4.37) for the extended LMI-based design. Using the change of variables, the BMI (3.24) is converted into an LMI as

$$\begin{bmatrix} M_P & M_A & M_B & \mathbf{0} \\ M_A^T & M_P & \mathbf{0} & M_{\tilde{C}} \\ M_B^T & \mathbf{0} & 1 & 1 \\ \mathbf{0} & M_{\tilde{C}}^T & 1 & \gamma_\eta^2 \end{bmatrix} \succ 0. \quad (4.38)$$

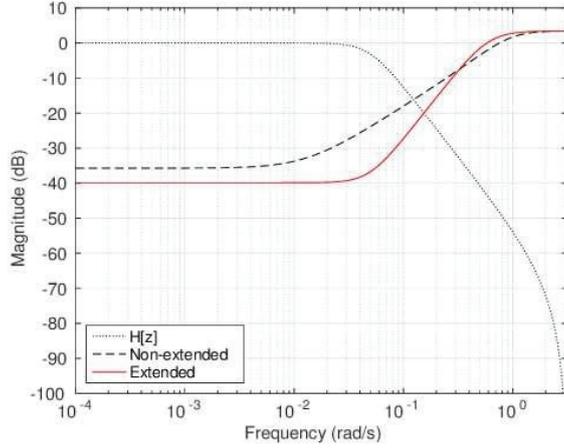


Figure 4.3: Frequency responses of H_2 norm based noise shaping filters designed by the extended and non-extended LMI techniques (Example 2).

Then, we numerically minimize γ_ϵ with respect to (P_f, P_g, W_f, W_g, L) subject to (4.31), (4.32), and (4.38). Finally, we obtain the matrices (A_R, B_R, C_R) from (4.24), (4.25), and (4.26) with the computed (P_f, P_g, W_f, W_g, L) .

Similarly, we can design the noise shaping filter by using the extended LMIs with change of variables as described in [32].

The system $H[z]$ is a low pass Butterworth filter of order 2 whose cutoff frequency is $\pi/64$ and the order of the noise shaping filter is 2.

H_2 norm based filter subjected to the Lee criterion

Fig. 4.3 presents the frequency responses of the H_2 norm based IIR filters obtained by using the extended and non-extended LMI techniques. The value of the Lee coefficient is set to be $\gamma_\eta = 1.5$.

Contrary to the IIR filter based on the non-extended LMI technique, the extended LMI-based filter matches the steepness of the output low pass filter and gives us uniform attenuation in the low frequency region. The MSE for the extended LMI design is -47.16 dB, while the MSE for the non-extended LMI-based design is -41.73 dB.

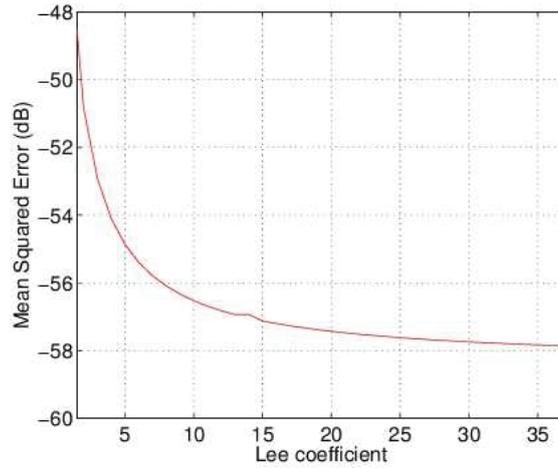


Figure 4.4: MSE as a function of the Lee coefficient for the extended LMI design.

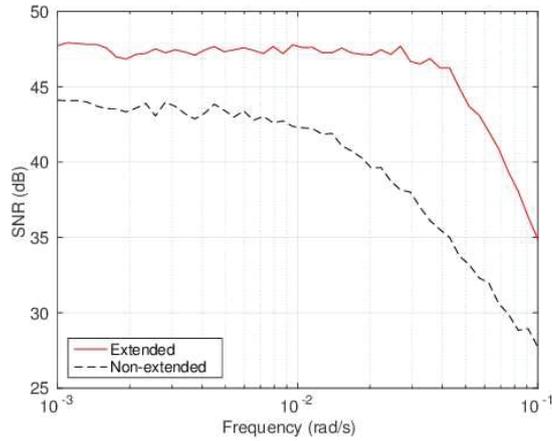


Figure 4.5: Empirical SNRs of H_2 norm based noise shaping filters designed by the extended and non-extended LMI techniques as functions of input frequency ω for $y_k = \sin(\omega k)$ (Example 2).

Fig. 4.4 illustrates the MSE of extended LMI-based filters for different values of the Lee coefficient. We can observe that the MSE monotonically converges to a set value as the value of the Lee coefficient increases to much higher values. On the other hand, this monotonic convergence behavior is not observed for non-extended LMI-based IIR filter in this example, which may be the result of the instability of

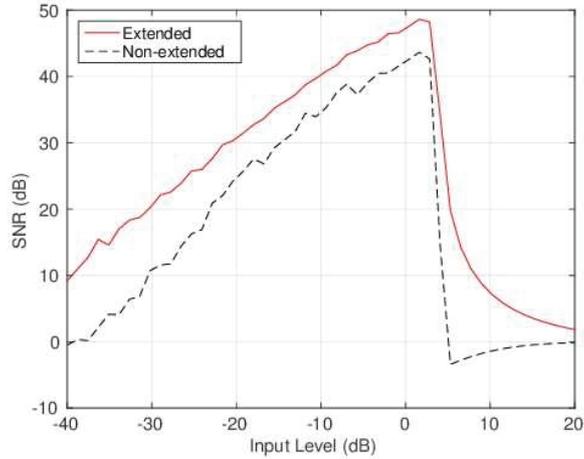


Figure 4.6: Empirical SNRs of H_2 norm based noise shaping filters designed by the extended and non-extended LMI techniques as functions of input amplitude $a(> 0)$ for $y_k = a \sin(k/100)$ (Example 2).

numerical optimization.

We also minimize the MSE without the Lee coefficient, which can be formulated as a convex optimization since the minimization can be described only by one BMI given by (3.12) and the LMI (3.13). The resultant MSE is -64.69 dB, which is close to the theoretical limit $20 \log_{10} |h_0| \approx -64.7$ dB.

To assess the performance of the designed quantizers, we evaluate empirical SNRs for sinusoidal signals with different frequencies and amplitudes. The input to a quantizer is converted into a binary signal whose values are either $L_o(> 0)$ or $-L_o$, where the value of L_o is determined by the noise shaping filter. It is noted that the maximum magnitude that can be quantized without an excessive overloading error is $2L_o$.

If the quantization error w of a two-level static uniform quantizer is a uniform random variable independent of the wide-sense stationary input y , then the variance

of the input to the static uniform quantizer is given by

$$\sigma_u^2 = \sigma_y^2 + \|R[z] - 1\|_2^2 \frac{L_o^2}{3}. \quad (4.39)$$

Let the loading factor [33], which is the ratio between $2L_o$ and the standard deviation σ_u of the input to the static uniform quantizer, be L_f . Then, we have from $L_f = 2L_o/\sigma_u$ and (4.39)

$$L_o = \frac{L_f \sigma_y}{2 \sqrt{1 - \frac{L_f^2 \|R[z] - 1\|_2^2}{12}}}. \quad (4.40)$$

For $L_f = 2$, the output levels L_o of the static uniform quantizer are given by 1.5260 and 1.4504 for the quantizers designed by the extended and non-extended LMI techniques, respectively.

We generate $N = 2 \cdot 10^3$ samples of $y_k = a \sin(\omega k)$ for $k = 0, 1, \dots, N-1$ as inputs to the designed quantizers connected to the system $H[z]$. Without quantization, the output of the system is $z_k = \sum_l h_l y_{k-l}$, where $\{h_l\}$ is the impulse response of $H[z]$. If we connect the quantizer output v to the system $H[z]$, the output of the system is $\hat{z}_k = \sum_l h_l v_{k-l}$ and the quantization error at the output is $\epsilon_k = \hat{z}_k - z_k$. From these signals, we compute the empirical SNR at the output of the system as

$$\frac{\sum_{k=0}^{N-1} |z_k|^2}{\sum_{k=0}^{N-1} |\epsilon_k|^2}. \quad (4.41)$$

Fig. 4.5 depicts the empirical SNRs for the input $y_k = \sin(\omega k)$ with angular frequencies from 10^{-3} to 10^{-1} . For every frequency, the extended LMI-based design has a greater empirical SNR than the non-extended LMI-based design.

For $w = 10^{-2}$, Fig. 4.6 shows the empirical SNRs for $y_k = a \sin(\omega k)$ having different amplitudes. The extended LMI-based design outperforms the non-extended LMI-based design. Although the extended LMI-based design is slightly more robust to quantizer overloading at higher input values than the non-extended design, both

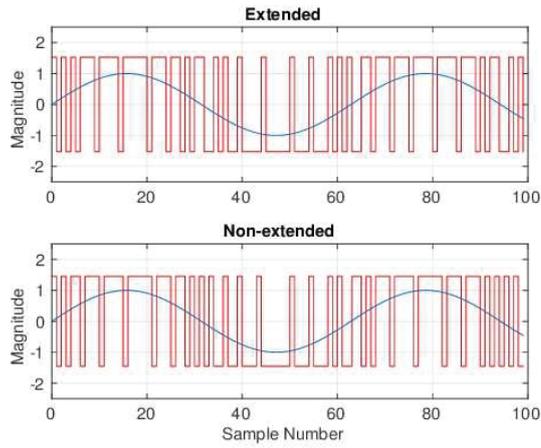


Figure 4.7: Output signals of quantizers designed by the extended and non-extended LMI techniques for $y_k = \sin(k/100)$ (Example 2).

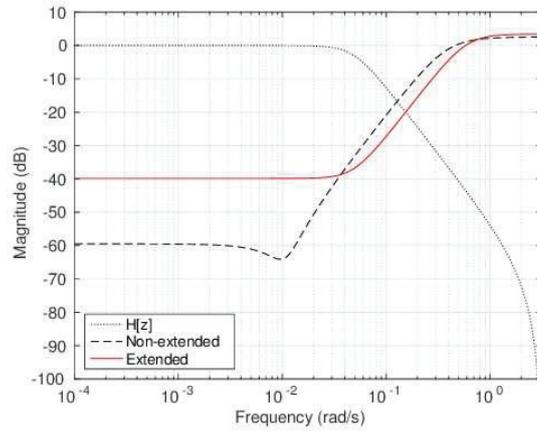


Figure 4.8: Frequency responses of H_∞ norm based noise shaping filters designed by the extended and non-extended LMI techniques (Example 2).

the designs suffer from excessive overloading errors for inputs whose levels are higher than 3 dB.

Fig. 4.7 shows the first 100 samples of the outputs of the two quantizers for the input $y_k = \sin(k/100)$. It can be seen that there are no significant differences between them.

H_∞ norm based filter subjected to the Lee criterion

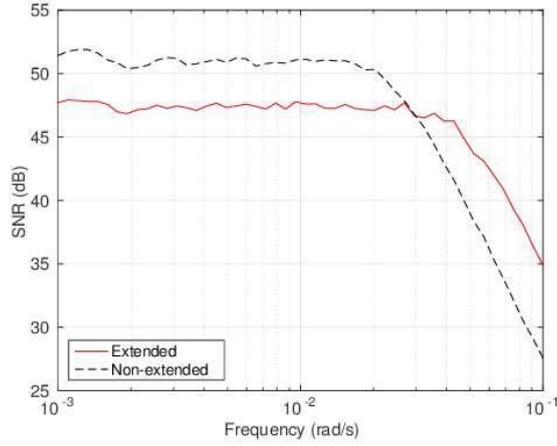


Figure 4.9: Empirical SNRs of H_∞ norm based filters designed by the extended and non-extended LMI techniques as functions of the input frequency ω for $y_k = \sin(\omega k)$ (Example 2).

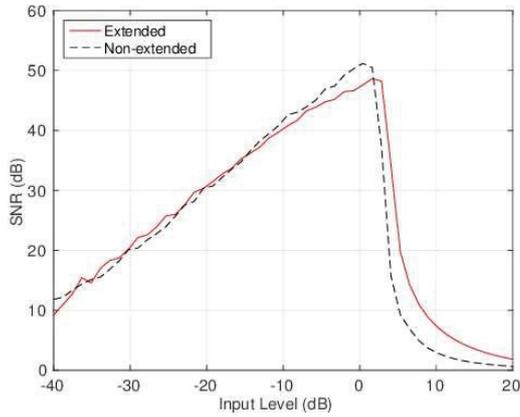


Figure 4.10: Empirical SNRs of H_∞ norm based filters designed by the extended and non-extended LMI techniques as functions of the input amplitude $a(> 0)$ for $y_k = a \sin(k/100)$ (Example 2).

We design noise shaping filters based on the H_∞ norm by using the extended and non-extended LMI techniques, where the value of the Lee coefficient is $\gamma_\eta = 1.5$.

Fig. 4.8 compares the frequency responses of the designed IIR filters. It is observed that our proposed extended LMI-based filter provides better results for the H_∞ norm of $H[z]R[z]$ with the value of -41.50 dB, whereas the H_∞ norm for the non-extended

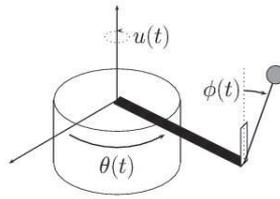


Figure 4.11: Rotary inverted pendulum.

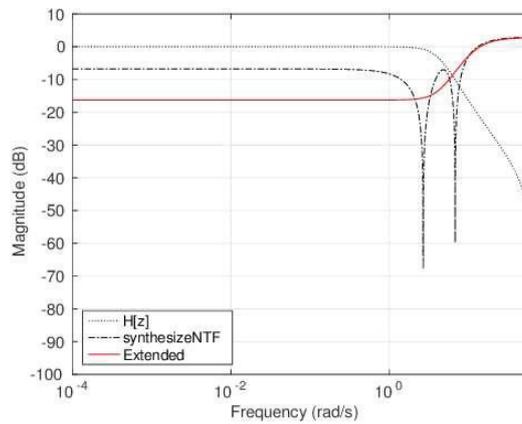


Figure 4.12: Frequency responses of noise shaping filters designed by the extended LMI technique with H_2 norm and the `synthesizeNTF` function (Example 3).

LMI design provides -33.50 dB.

For the loading factor $L_f = 2$, the output level L_o of the quantizer designed by the extended LMI technique is 1.5260, whereas the output level of the quantizer designed by the non-extended LMI technique is 1.4092.

As shown in Fig. 4.9, the non-extended LMI-based design has larger SNRs than the extended LMI-based design at low frequencies. This may be due to the fact that the noise shaping filter of the non-extended LMI-based design has smaller responses at low frequencies.

In Fig. 4.10, the empirical SNRs for the input $y_k = a \sin(\omega k)$ with different

amplitudes are presented. Although the extended LMI design is slightly more robust to the quantizer overloading, the two designs have almost the same performance on the input level.

Example 3 : Extended vs. `synthesizeNTF`

In the last example, we compare our designed H_2 norm based IIR filter with the conventional IIR filter for the $\Delta\Sigma$ modulator. The conventional IIR filter is obtained by using the `synthesizeNTF` function in the DELSIG toolbox [28], which does not utilize the knowledge about the connected system $H[z]$ to minimize the quantization error unlike our proposed design.

The system is a rotary inverted pendulum, which is controlled based on the observation signals.

Fig. 4.11 is the rotary inverted pendulum for our design Example 3.

The pendulum connected at the end of the rotary arm is controlled by rotating the main body in the horizontal plane. The yaw angle of the arm is $\theta(t)$. The pendulum freely swings about a pitch angle $\phi(t)$ in the vertical plane to the arm. The torque $u(t)$ is applied to actuate the pendulum. If $\phi(t) = 0$, then the pendulum is balanced in the inverted position.

We define the state of the rotary inverted pendulum as

$$x^T(t) = [\phi(t), \theta(t), \dot{\phi}(t), \dot{\theta}(t)].$$

With the sampling period $T_s = 0.05$, the linearized continuous system is discretized. The (A_H, B_H, C_H) matrices of the discrete-time linearized system are given

by

$$\begin{aligned}
A_H &= \begin{bmatrix} 1.1420 & 0 & 0.0523 & 0.0019 \\ -0.0083 & 1 & -0.0001 & 0.0491 \\ 5.8075 & 0 & 1.1420 & 0.0768 \\ -0.3362 & 0 & -0.0083 & 0.9631 \end{bmatrix} \\
B_H &= \begin{bmatrix} -0.0109 \\ 0.0054 \\ -0.4451 \\ 0.2137 \end{bmatrix} \\
C_H &= [0 \ 1 \ 0 \ 0].
\end{aligned}$$

Assuming that all of the state variables are available at the controller, we adopt the state feedback control with gain $K = [K_1, K_2, K_3, K_4] = [34.1, 2.59, 3.52, 1.67]$ to stabilize this system.

We assume that the yaw angle $\theta(t)$ is quantized by our designed quantizer. Then the discrete transfer function from the quantization error to the output of the system is given by $C(zI - A - BK)^{-1}BK_2$, which leads to (4.42).

The output of the quantizer passes through the fourth order system $H[z]$ given by

$$H[z] = \frac{-0.01387z^3 + 0.01751z^2 + 0.01696z - 0.0137}{z^4 - 2.678z^3 + 2.548z^2 - 0.9734z + 0.1097}. \quad (4.42)$$

The Lee coefficient is set at $\gamma_\eta = 1.4$.

Fig. 4.12 shows the frequency responses of our H_2 norm based IIR filter designed using the extended LMI technique and the conventional IIR filter designed using the `synthesizeNTF` function.

The MSE of our extended LMI technique is -23.55 dB, whereas the MSE of the conventional IIR filter is -20.72 dB. Our filter has a smaller MSE compared with the conventional IIR filter, which justifies the design with the knowledge of the system $H[z]$ for minimizing the error. However, the smaller MSE does not guarantee a better performance at every frequency.

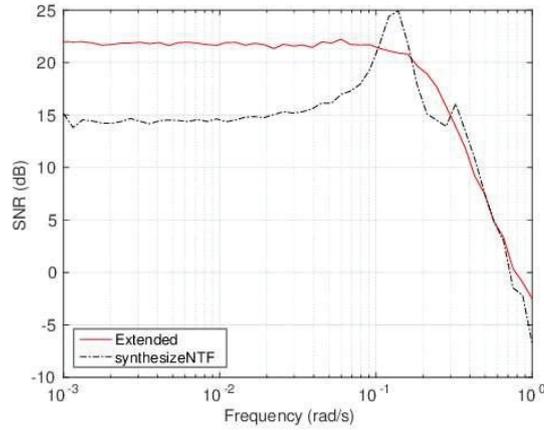


Figure 4.13: Empirical SNRs of noise shaping filters designed by the H_2 norm based extended LMI technique and the `synthesizerNTF` function as functions of input frequency ω for $y_k = \sin(\omega k)$ (Example 3).

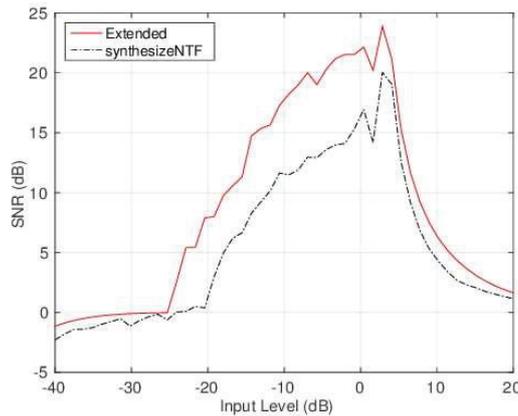


Figure 4.14: Empirical SNRs of noise shaping filters designed by the H_2 norm based extended LMI technique and the `synthesizerNTF` function as functions of input amplitude $a(> 0)$ for $y_k = a \sin(k/20)$ (Example 3).

The output levels L_o for the quantizers designed by the extended LMI technique and the `synthesizerNTF` function are respectively 1.3720 and 1.4050 for $L_f = 2$.

Fig. 4.13 compares the empirical SNRs for our designed and conventional quantizers. Up to 10^{-1} , our designed quantizer exhibits better SNR performance than the conventional quantizer. The SNRs of our designed quantizer are almost the same up

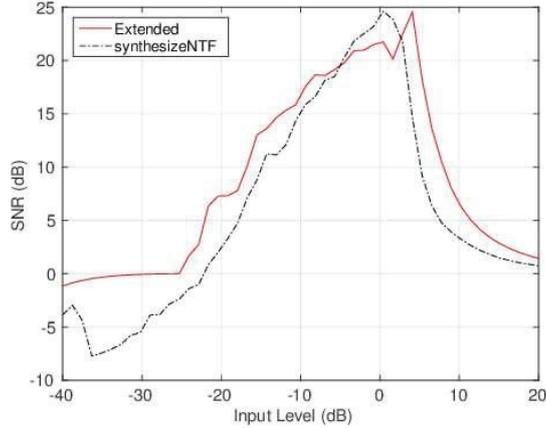


Figure 4.15: Empirical SNRs of noise shaping filters designed by the H_2 norm based extended LMI technique and the `synthesizeNTF` function as functions of input amplitude $a(> 0)$ for $y_k = a \sin(0.12k)$ (Example 3).

to 10^{-1} , whereas the SNRs of the conventional quantizer are different.

For the unit sinusoidal signal at frequency $1/20$, our designed quantizer outperforms the conventional quantizer for practical values of the input level as illustrated in Fig. 4.14. For the unit sinusoidal signal at frequency 0.12 , as depicted in Fig. 4.15, our designed quantizer is slightly inferior to the conventional quantizer at some input levels around 0 dB. However, our design is superior to the conventional design for most of the input values.

Fig. 4.16 and Fig. 4.17 demonstrate the outputs of the system $H[z]$ for the inputs $y_k = \sin(k/20)$ and $y_k = \sin(0.12k)$. For $y_k = \sin(k/20)$, our designed quantizer has smaller errors compared with the conventional quantizer, while for $y_k = \sin(0.12k)$, there is no significant difference between the designed quantizer and the conventional quantizers.

Fixed-point hardwares are often utilized in practice. To assess performances of $\Delta\Sigma$ modulators in the fixed-point arithmetic, we conduct simulations with fixed-point arithmetic by MATLAB Fixed-Point Designer [34].

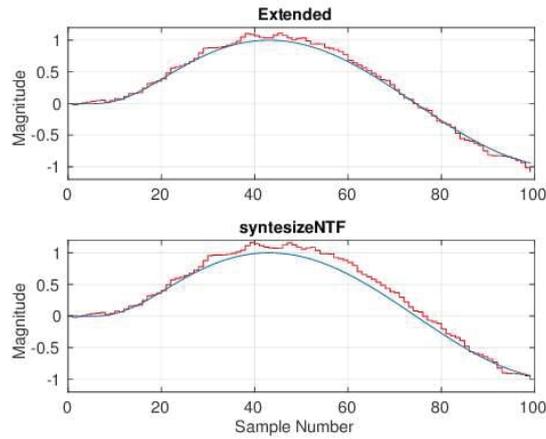


Figure 4.16: Output signals of the system with quantizers designed by the H_2 norm based extended LMI technique and the `syntesizeNTF` function for $y_k = \sin(k/20)$ (Example 3).

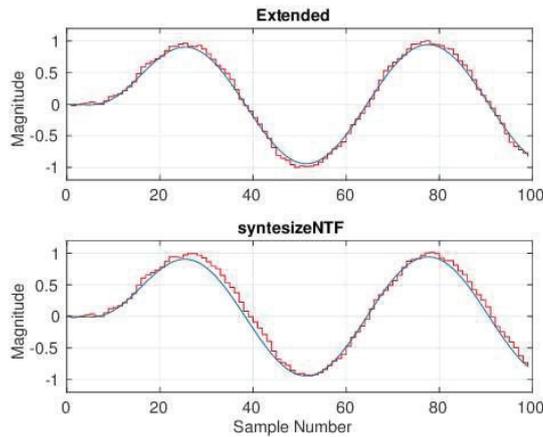


Figure 4.17: Output signals of the system with quantizers designed by the H_2 norm based extended LMI technique and the `syntesizeNTF` function for $y_k = \sin(0.12k)$ (Example 3).

In a simulation, all the values in a $\Delta\Sigma$ modulator are represented in a signed fixed-point binary number format having a specified word length and fraction length. A numerical value is rounded to the nearest fixed-point binary number and is wrapped around if an overflow occurs. Addition and multiplication of two numbers are stored in the fixed-point binary number with the word length and fraction length of the

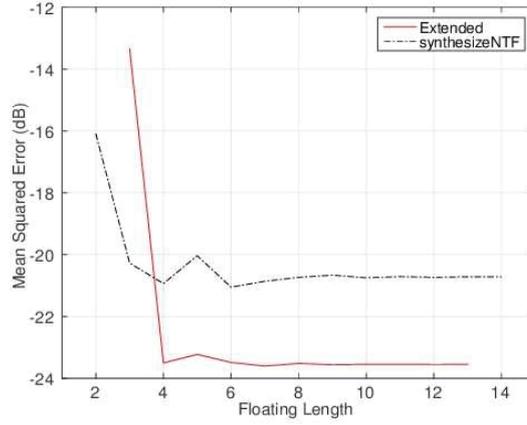


Figure 4.18: MSEs of fixed-point $\Delta\Sigma$ modulators designed by the extended LMI technique with H_2 norm and the `synthesizeNTF` function as functions of the floating length (Example 3).

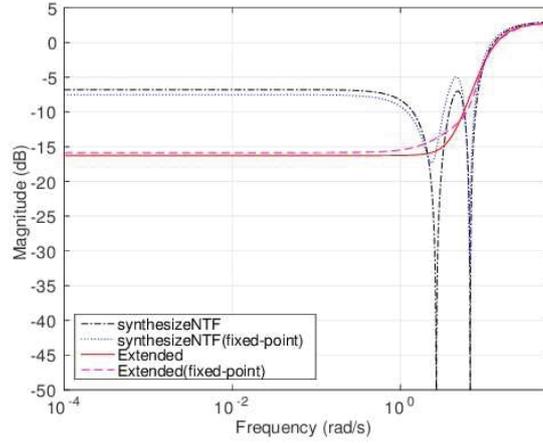


Figure 4.19: Frequency responses of fixed-point and floating-point noise shaping filters designed by the extended LMI technique with H_2 norm and the `synthesizeNTF` function (Example 3).

operands.

There are a large number of realizations of a digital filter. The performance of a fixed-point implementation heavily depends on its realization. Here we adopt two realizations: One is based on the space-state realization which is given directly by CVX for the extended LMI design or by `synthesizeNTF` for the conventional IIR

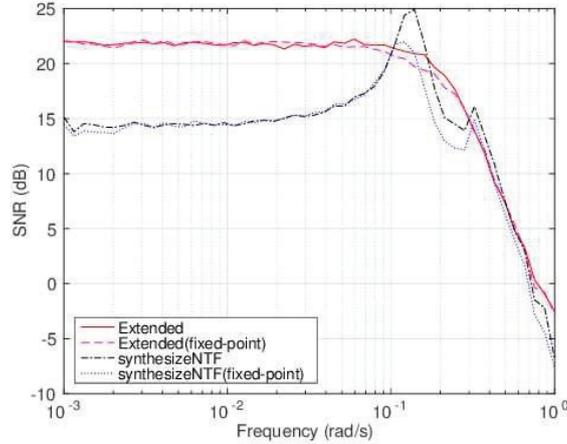


Figure 4.20: Empirical SNRs of fixed-point and floating point noise shaping filters designed by the H_2 norm based extended LMI technique and the `synthesizeNTF` function as functions of frequency ω of quantized sinusoidal input signals with unit amplitudes (Example 3).

design. The other is the space-state realization in the controllable canonical form.

From the (A, B) matrices of a state-space realization of $R[z]$, we solve the Lyapunov equation $K = A_R K A_R^T + B_R B_R^T$ and construct a transformation matrix $T = \text{diag}(K_{11}, K_{22}, \dots, K_{nn})$, where n is the order of $R[z]$ and K_{ii} is the i th diagonal entry of K . Then, we apply the so-called l_2 scaling to obtain the new scaled state-space realization $(T^{-1}A_R T, T^{-1}B_R, C_R T)$. The l_2 scaling reduces the possibility of the overflows.

To see which realization is better and how many bits should be assigned to the integer part and the fractional part, we fix the word length to be 16 and change the fraction length from 1 to 15. For the two type of realizations, we have found that the l_2 scaled realizations give smaller MSEs. We have also observed that the controllable canonical form is better for the extended LMI design, whereas the realization given by `synthesizeNTF` is for the conventional IIR design. Thus, in the followings, we will only show the results using the controllable canonical form for the extended LMI

design and the realization given by `synthesizeNTF` for the conventional IIR design.

Fig. 4.18 compares MSEs of $\Delta\Sigma$ modulators implemented using fixed-point number representations, which we call fixed-point $\Delta\Sigma$ modulators, as functions of the floating length, where only finite values are shown. To attain -20dB , at least four bits are necessary for the fractional part of the extended LMI design, whereas three bits are for fractional part of the conventional IIR design. Fig. 4.18 shows that the integer length of the extended LMI design requires at least three bits, whereas the integer length of the conventional IIR design requires two bits.

Next, we set the word length to be 8 and the floating length to be 5. Fig. 4.19 presents the frequency responses of fixed-point $\Delta\Sigma$ modulators obtained by the extended LMI design and by the conventional IIR design.

At low frequencies, the fixed-point extended LMI design suffers from a small performance loss, whereas the fixed-point conventional IIR design enjoys a small performance gain. The MSE of the fixed-point extended LMI design is -23.50 dB , which is slightly worse than the MSE of the floating-point extended LMI design. On the other hand, the MSE of the fixed-point conventional IIR design is -20.94 dB , which is slightly better than the MSE of the floating-point conventional IIR design.

For different frequencies, we generate a sinusoidal signal $\sin(\omega t)$ and convert its values into 8-bit signed fixed-point binary numbers with floating length being 5. For the fixed-point sinusoidal signal, we evaluate the empirical SNRs. Fig. 4.20 compares the empirical SNRs of the fixed-point extended LMI design (dashed curve) and the fixed-point conventional IIR design (dotted curve) are compared with the floating-point extended LMI design (solid curve) and the floating-point conventional IIR design (dashed-dotted curve).

We can conclude from Fig. 4.19 and Fig. 4.20 that when the word length is 8 and the floating length is 5, the fixed-point implementation does not exhibit a significant

performance loss both for extended LMI design and the conventional IIR design in this example.

4.2 Design Based on Approximation Techniques

In the previous section, we obtain IIR noise shaping filters for $\Delta\Sigma$ modulators using the extended LMI technique [26]. The IIR design problem proposed in the previous section utilizes the non-ideal output filter to minimize the weighted quantization noise at the output of the $\Delta\Sigma$ modulator. However, the order of the IIR noise shaping filter obtained is constrained to be identical to the non-ideal output filter. In this section, we introduce two well-known approximation techniques that can be used to obtain IIR noise shaping filters for $\Delta\Sigma$ modulators without any constraint on the IIR noise shaping filter order.

4.2.1 The Yule-Walker Method

In the first approximation technique, we will obtain an IIR noise shaping filter for a $\Delta\Sigma$ modulator by using the Yule-Walker method. First, we obtain an optimal FIR filter using the method proposed in Chapter 3, then we use the Yule-Walker method to obtain the IIR filter by approximating the frequency response of the optimal FIR filter. The Yule-Walker method for the filter design is based on recursion and the criterion used for the approximation is based on the least-squares (LS) method. More specifically, the Yule-Walker method leads to generation of an IIR filter by fitting a specified frequency response.

Now, we provide a design example to show the effectiveness of our proposed H_2 norm based IIR filter based on the Yule-Walker method. For the approximation of the FIR by an IIR digital filter using the Yule-Walker method, we use the function `yulewalk` in MATLAB. Similarly, H_∞ norm IIR filter can be designed by using the

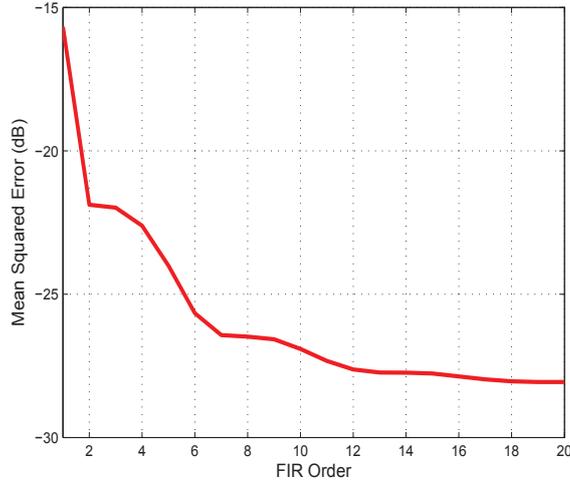


Figure 4.21: MSE as a function of the FIR filter order.

Yule-Walker approximation. Here, we omit the design of H_∞ norm based filter.

Let us consider a lowpass $\Delta\Sigma$ modulator with an OSR of 32. A fourth order lowpass Butterworth filter $H[z]$ is assumed to be connected at the output of our $\Delta\Sigma$ modulator to recover the digital information. Here, any other type of lowpass filter can be used as our weighting function, but we confine our attention to Butterworth filter only in this example.

First, we obtain an optimal FIR filter using the method proposed in [14]. The order of FIR is chosen based on the convergence of the MSE as the order of the FIR filter is increased to higher values. Fig. 4.21 shows us that the optimal order of FIR filter should be 18 based on the convergence of the variance of the quantization noise.

Fig. 4.22 depicts frequency responses of FIR and IIR noise shaping filters for the $\Delta\Sigma$ modulator. To achieve a good approximation of the FIR filter based on the MSE performance, an IIR filter of order 3 results in almost the same MSE value as that of the FIR filter of order 18. Through our calculations, we note that the MSE of the FIR filter is -28.04 dB, while the IIR filter gives us -28.90 dB. With the order-reduction

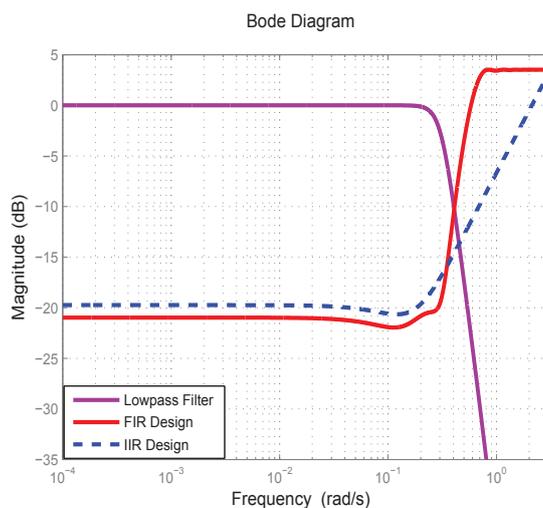


Figure 4.22: Frequency responses of FIR filter of order 18, and IIR filter of order 3 for the $\Delta\Sigma$ modulator. The output filter is of order 4.

of 83.3%, a low-order IIR filter obtained by using an effective approximation method can achieve almost the same performance as that of the FIR filter.

4.2.2 The Least-Squares (LS) Approximation

The LS method is another well-known approximation technique which can be used to design IIR filters. We take a simple approach to design IIR filters for $\Delta\Sigma$ modulators. First, we design an optimal FIR filter to satisfy the required specifications. It is expected that the optimal FIR filter having a sufficiently large order can well approximate the theoretically optimal IIR filter. Then, we approximate the optimal FIR filter with a lower-order IIR filter. If we achieve a good approximation, then the obtained IIR filter can be considered as a good approximation of the theoretically optimal IIR filter. The approximation of FIR by IIR filter has been well studied and some recent design methods have been summarized in [35] which give satisfactory results. As an example, we consider the H_∞ optimal IIR filter for our design criterion

and provide a design example to show the effectiveness of IIR error feedback filters for the $\Delta\Sigma$ modulator designed by our proposed method.

Problem Formulation

We utilize the design of the optimal FIR filter in [10], and then approximate the FIR filter by an IIR filter. If we consider the weighting function $H[z]$, then the FIR filter can also be designed using the method proposed in Chapter 3.

Let the transfer function of the FIR filter be denoted as $R_F[z]$. Now, we would like to approximate the designed $R_F[z]$ by using a reduced-order IIR filter $R[z]$ with a transfer function as expressed in (4.1), (4.2), (4.3).

The approximation error function, defined as

$$E[z] = R_F[z] - R[z], \quad (4.43)$$

should be minimized.

The LS method proposed in [36] tries to find $A[z]$ in (4.3) and $B[z]$ in (4.2) that minimizes the squared error given by $\int_{-\pi}^{\pi} |E[e^{j\omega}]|^2 d\omega$. In the LS method, we first determine the denominator coefficients a_n in (4.3) using an iterative procedure which requires the solution of an over-determined set of linear equations and some operations. Then, we use the denominator coefficients a_n to calculate the numerator coefficients b_n in (4.2).

Design Examples

In this section, we provide design examples to show the effectiveness of proposed IIR error feedback filters for a $\Delta\Sigma$ modulator. We obtain low-order IIR filters by approximation of the optimal FIR noise shaping filter proposed in [10]. The optimization problem presented in [13] is solved by using LMI toolbox [37], while we utilize CVX tool [20] to obtain all simulation results.

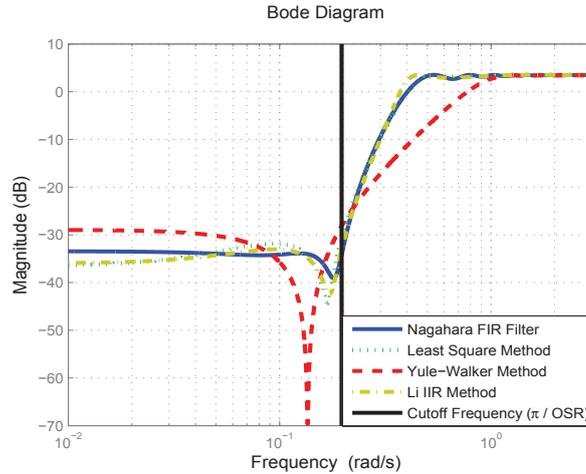


Figure 4.23: Frequency responses of FIR and IIR error feedback filters for the lowpass $\Delta\Sigma$ modulator.

Lowpass $\Delta\Sigma$ Modulator

First, we consider the design of a lowpass $\Delta\Sigma$ modulator with a FIR error feedback filter. We assume that the FIR error feedback filter of order 28 is chosen to achieve the desired performance. The input to the $\Delta\Sigma$ modulator is a lowpass signal with an OSR of 16 and the signal bandwidth defined as $\Omega_x = [-\frac{\pi}{OSR}, \frac{\pi}{OSR}]$. The optimal FIR noise shaping filter can be obtained by solving the convex optimization problem based on the H_∞ norm as proposed in [10]. Then, we approximate the optimal FIR noise shaping filter with IIR digital filter which is obtained by using the LS approximation method [36].

Fig. 4.23 depicts frequency responses of noise shaping filters which are obtained by minimizing the H_∞ norm of $H[z]R[z]$ in (2.15) subject to (2.22). The Lee criterion is used to limit the maximum magnitude to 3.52 dB.

Here, our objective is to approximate the FIR noise shaping filter of order 28 by a low-order IIR digital filter. The maximum value of $R[z] - 1$ in signal band occurs near the cutoff frequency for all the filters, which can be seen in the enlarged frequency

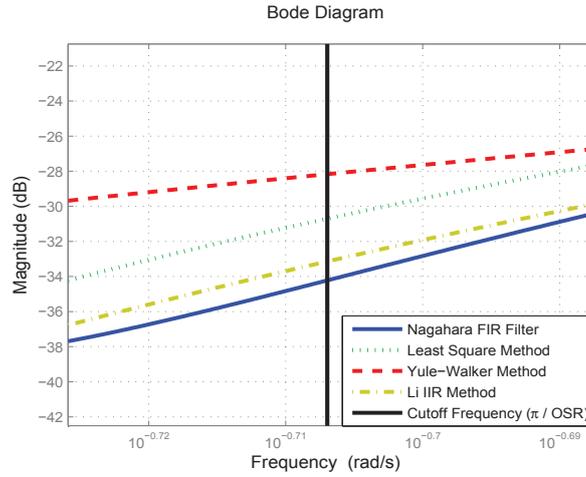


Figure 4.24: Enlarged frequency responses of lowpass filters near the cutoff frequency.

response in Fig. 4.24.

Table 4.1 gives us the comparison between different methods used to obtain FIR and IIR digital filters. For the 4th order IIR filter, the LS method results in H_∞ norm of -30.79 dB, while the Yule-Walker method results in -28.18 dB. Between these two approximation methods, LS method gives us better approximation $R[z]$ of the optimal FIR noise shaping filter $R_F[z]$.

The difference between the H_∞ norm of the LS method and [13] is not too large. Therefore, our LS design can achieve almost the same performance as the design proposed in [13].

Fig. 4.25 shows us the l_2 norm of the error $R_F[z] - R[z]$ as a function of the order of the IIR filter obtained by using the LS method. For the 4th order IIR filter, the LS gives us l_2 norm of -30.92 dB, while the Yule-Walker method gives us -6.44 dB.

Fig. 4.26 shows the convergence of the objective function for the first 20 iterations of the algorithm for $OSR=[16\ 32\ 64\ 128\ 256\ 512]$. The important point to note here is that, the iterative algorithm method utilizes LMI toolbox with some special settings

Table 4.1: Comparison between different design methods

Design Methods	Filter Type	Filter Order	H_∞ Norm (dB)
Nagahara Design [10]	FIR	28	-34.19
LS Method [36]	IIR	4	-30.79
Yule-Walker Method	IIR	4	-28.18
Li's Method [13]	IIR	4	-33.12

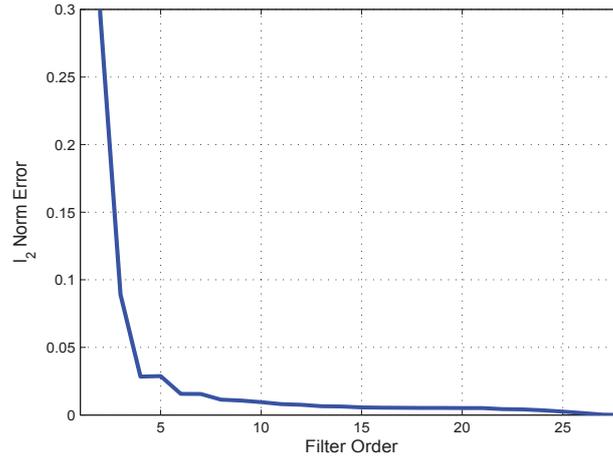


Figure 4.25: l_2 norm error as a function of the order of the lowpass IIR filter obtained by using the LS method.

to ensure the convergence, while we have utilized CVX toolbox without any special settings for simulating their iterative algorithm.

Bandpass $\Delta\Sigma$ Modulator

Next, we consider the design of a bandpass modulator. The noise shaping filter is a bandpass filter which suppresses the noise between two specific frequencies in the signal band. The input to the modulator is an oversampled signal with OSR of 16. The signal bandwidth is defined as $[\frac{\pi}{2} - \frac{\pi}{OSR}, \frac{\pi}{2} + \frac{\pi}{OSR}]$, with a center frequency $\frac{\pi}{2}$. We assume that the desired performance of the bandpass $\Delta\Sigma$ modulator is achieved by using a FIR noise shaping filter of order 32.

After the optimal design of the bandpass FIR error feedback filter using the

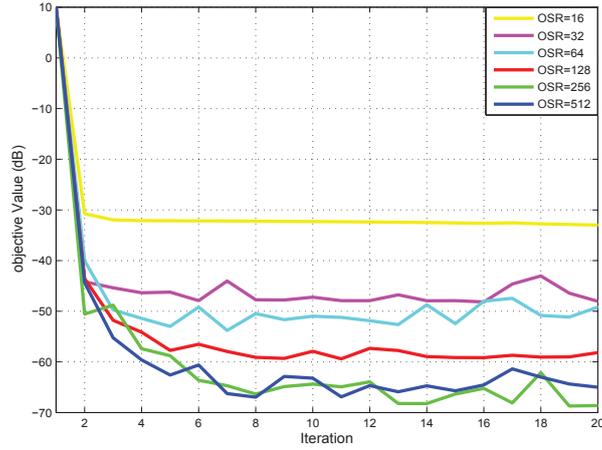


Figure 4.26: Convergence behavior of the iterative algorithm [13] for OSR=[16 32 64 128 256 512]

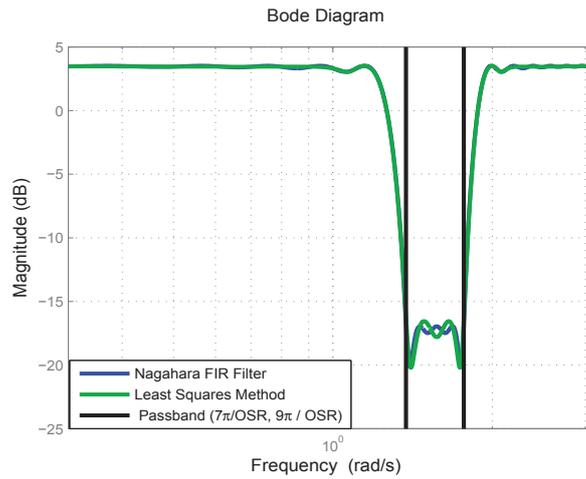


Figure 4.27: Frequency responses of FIR and IIR error feedback filters for the band-pass $\Delta\Sigma$ modulator.

method in [10], we use an approximation method to obtain a low-order bandpass IIR filter. In the previous example, we observed that the LS method outperforms and gives us better approximation than the Yule-Walker method; hence, we only consider the LS method to approximate the bandpass FIR filter.

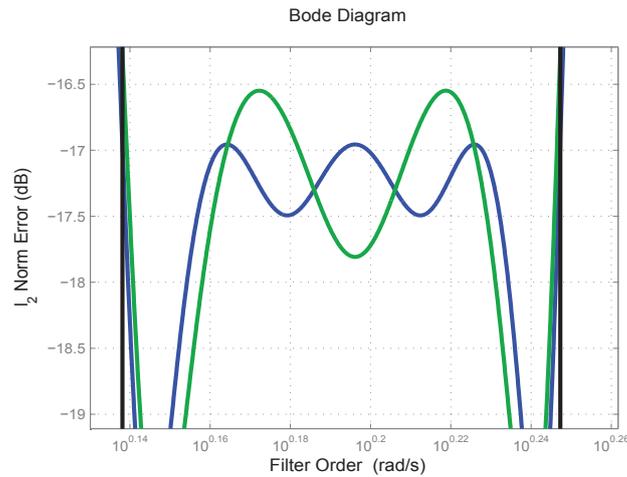


Figure 4.28: Enlarged frequency responses of bandpass filters near cutoff frequencies.

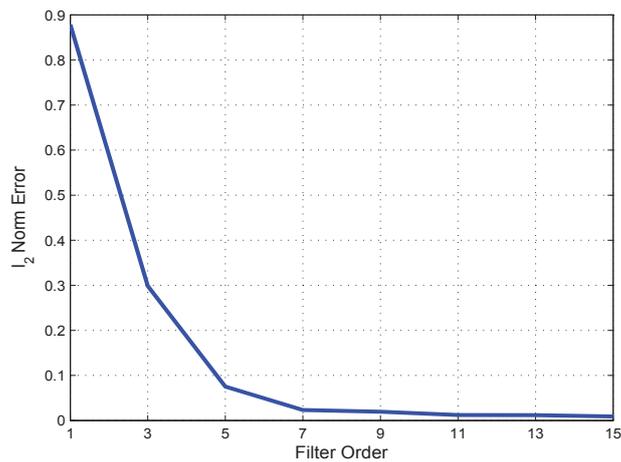


Figure 4.29: l_2 norm error as a function of the order of the bandpass IIR filter obtained by using the LS method.

Fig. 4.27 shows frequency responses of bandpass FIR and IIR filters. We approximate the FIR error feedback filter of order 32 with the IIR filter of order 12 which gives us almost the same performance as that of the FIR filter. The enlarged frequency response in the passband of the $\Delta\Sigma$ modulator is shown in Fig. 4.28. The H_∞ norm of the FIR filter with desired response is -16.96 dB, and the resultant IIR

filter of order 12 gives us -16.26 dB in the passband.

Fig. 4.29 shows us that the l_2 norm of $R_F[z] - R[z]$ for bandpass IIR filter converges to the least approximation error, which can be achieved by using the IIR filter of order 12. The approximation error reduces slowly above the IIR filter of order 12.

4.3 Design Based on the Hybrid Technique

The idea of a hybrid design is proposed for obtaining IIR noise shaping filters in [38]. The hybrid design strategy uses Schreier's method [6] for selecting poles arrangements, while fully optimizing zeros using recent optimal design methods for FIR filter. The zeros of the Schreier's noise shaping filter are constrained to lie on the unit circle, while poles take maximally flat arrangement at low frequencies. The constrained on the zeros can be lifted by using the optimal FIR design method which gives full freedom to choose positions of zeros anywhere in the unit circle. The hybrid design strategy gives us an IIR filter which gets best of both Schreier's and optimal FIR methods which are limited in flexibility and requiring high orders, respectively. However, the method in [38] can only design H_2 norm based noise shaping IIR filter.

We deal with two noise shaping IIR filters based not only on H_2 but also on H_∞ norm by using the hybrid design method. We also show that the hybrid design is superior to the conventional (Schreier's) method which restricts the position of zeros on the unit circle only. Finally, a design example is provided to demonstrate the effectiveness and comparison of noise shaping IIR filters obtained by using the methods proposed in [6, 13, 26, 16, 38].

The transfer function of an IIR filter is defined in (4.1), (4.2), (4.3).

Now, we obtain noise shaping IIR digital filters by minimizing (2.6) under the Lee criterion (2.22) while also considering the non-ideal behavior of the output filter

$H[z]$. We consider the hybrid design to obtain indirect solutions for H_2 and H_∞ norms based noise shaping filters.

The hybrid design assumes that a_1, \dots, a_{N_A} in (4.3) are pre-assigned coefficients which are obtained by using Schreier's method [6]. Then, we are left with the N_B numerator coefficients b_1, \dots, b_{N_B} in (4.2) which are to be found by using the optimal FIR technique in [14]. Hence, the problem can be reduced into LMIs that can be solved by using convex optimization.

4.3.1 Design Example

As a design example, let us consider a filterless audio amplifier [39] which is a widely used electronic component in many portable electronics and mobile phones.

Fig. 4.30 shows a simplified block diagram of a switching audio amplifier, where the lowpass filter is provided by the inherent inductance of the speaker coil and natural filtering of the human ear to recover the information signal. This is known as a filterless solution. The filterless solution is adopted since it greatly reduces the external components that are often used for lowpass filtering; hence, simplifying the circuit design and reducing the system cost. In our method, we utilize the transfer function of the output analog lowpass filter as our weighting function to reduce the in-band quantization noise due to A/D conversion.

We assume that the audio amplifier under consideration utilizes a lowpass $\Delta\Sigma$ modulator with noise shaping IIR filter for A/D conversion. The audio input signal to the lowpass $\Delta\Sigma$ modulator is discretized with an OSR of 256. The analog lowpass filter at the output is of first order with a cutoff frequency $\frac{\pi}{OSR}$. By using the proposed hybrid design and LS approximation methods, we obtain fourth order noise shaping IIR filters which are used in the feedback of the $\Delta\Sigma$ modulator. For the stability of the $\Delta\Sigma$ modulator, the Lee coefficient γ is set to be 1.5 (3.52 dB) to limit the

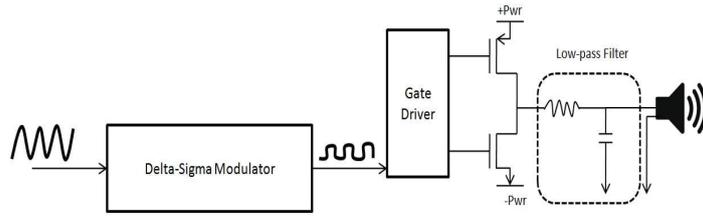


Figure 4.30: Block diagram of an audio amplifier with a $\Delta\Sigma$ modulator as an A/D converter.

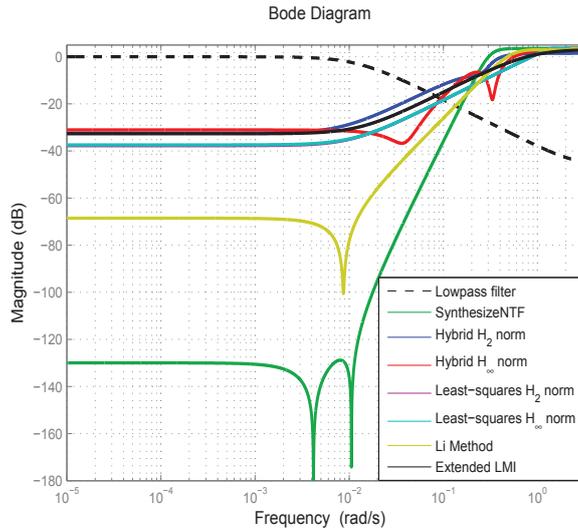


Figure 4.31: Frequency responses of noise shaping IIR filters for a $\Delta\Sigma$ modulator.

maximum out-of-band gain that can overload the quantizer.

Fig. 4.31 depicts frequency responses of noise shaping IIR filters obtained by using proposed and conventional methods. Our proposed noise shaping IIR filters obtained by using hybrid-design and approximation methods match the steepness of the output lowpass filter and gives uniform attenuation throughout the signal band. The `synthesizeNTF` and the method proposed in [13] give better attenuation in the signal passband, but fail to remove the noise near the cutoff frequency since they do not utilize the weighting function to minimize the in-band quantization noise.

Table 4.2: Comparison of $\|H[z]R[z]\|_2^2$ and $\|H[z]R[z]\|_\infty^2$ norm values of different design methods

Design Methods	IIR order	H_2 norm (dB)	H_∞ norm (dB)
Hybrid H_2	4	-36.02	-29.63
Hybrid H_∞	4	-36.76	-31.07
<code>synthesizeNTF</code>	4	-34.53	-26.98
LS H_2	4	-37.99	-37.99
LS H_∞	4	-37.99	-36.68
Li's method [13] H_∞	4	-36.66	-36.66
Extended LMI	1	-37.34	-32.72

Table 4.2 compares H_2 and H_∞ norm values of the in-band quantization noise for each design method. Clearly, both of our proposed H_2 and H_∞ norms hybrid design based noise shaping IIR filters outperform the H_2 norm conventional `synthesizeNTF` based design. On the other hand, the noise shaping IIR filters obtained by using the LS approximation method outperforms the hybrid design and the method proposed in [13]. Since the IIR filter obtained by using the method in [13] does not incorporate the non-ideal output filter, the value of the band quantization noise is slightly higher than that of the LS approximation method. For the extended LMI design [26], the first order H_2 norm based IIR filter outperforms other designs and matches the performance of the fourth order H_2 norm based IIR filter obtained by using the LS method.

Now, let us evaluate the performance of the lowpass $\Delta\Sigma$ modulator with the noise shaping IIR filter obtained by using the LS approximation based on H_2 norm of the in-band quantization noise. The MATLAB function `simulateDSM` in DELSIG toolbox [28] is used to simulate the $\Delta\Sigma$ modulator. The input to the modulator is a sinusoidal signal with a frequency of 100 Hz and an amplitude of 0.5. The uniform quantizer is assumed to have quantization levels $L = 2$ and quantization interval $d = 2$. Fig. 4.32 shows us plots for input, output and quantization error of the $\Delta\Sigma$ modulator. The

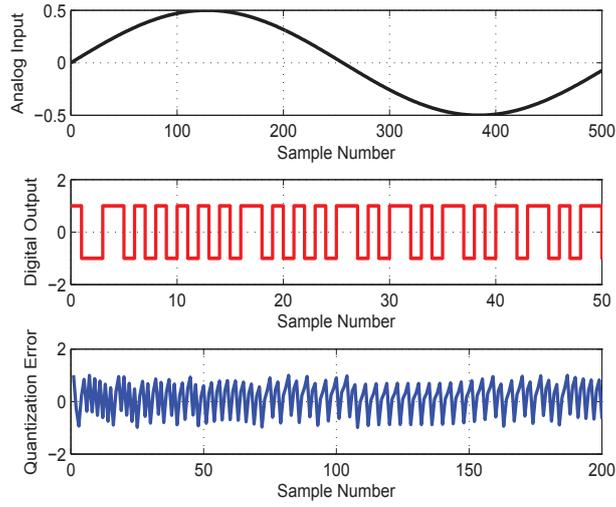


Figure 4.32: Simulation of the lowpass $\Delta\Sigma$ modulator with a noise shaping IIR filter obtained by using the LS approximation method.

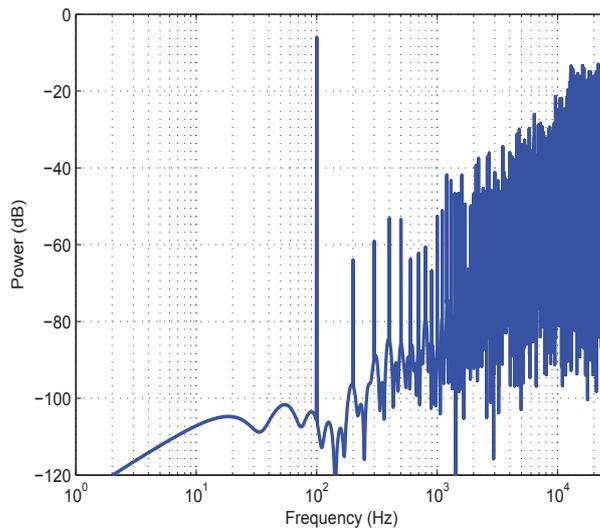


Figure 4.33: Frequency spectrum of the digital output of the lowpass $\Delta\Sigma$ modulator in Fig. 4.32 .

digital output is represented by using +1 and -1 volts for binary 0 and 1 respectively.

The quantization error is a random signal which is bounded by $[-\frac{d}{2}, +\frac{d}{2}]$.

Fig. 4.33 illustrates the frequency spectrum of the digital output of the proposed $\Delta\Sigma$ modulator by using the LS approximation method. The frequency spectrum plot makes it possible to visualize the noise shaping performed by the proposed noise shaping IIR filter. The frequency notch of the input sinusoidal signal appears at 100 Hz.

4.4 Design Based on the Iterative LMI Technique

In this section, we propose an iterative LMI algorithm to solve the non-convex design problem for obtaining the near optimal noise shaping IIR filter. We minimize the variance of the quantization noise at the output of a $\Delta\Sigma$ modulator subject to the stability constraint. Since the non-ideal filter at the output of the $\Delta\Sigma$ modulator is an imperfect filter which may cause quantization noise leakage in the passband of the information signal, the non-ideal behavior of the output filter is also taken into consideration. Moreover, we also design and compare the performances of the noise shaping IIR filters by using the hybrid design and Schreier's method [6]. The design example for the bandpass $\Delta\Sigma$ modulator in a RF transmitter is provided to show the effectiveness of the proposed technique.

The transfer function of an IIR filter is defined in (4.1), (4.2), (4.3). Our design problem for the synthesis of noise shaping IIR is to minimize (2.13) subject to (2.22).

A similar design problem is investigated in [14] for obtaining the noise shaping filter $R[z]$. But the method in [14] only considers the design of a FIR filter, which can be solved to obtain an optimal solution using any convex optimization technique. On the other hand, our objective is to obtain an IIR filter which results in a non-convex optimization design problem.

In [15], an IIR filter is obtained by solving the non-convex optimization problem.

However, the design problem can only be solved if the order of $R[z]$ is identical to the order of $H[z]$. The iterative algorithm proposed in [13, 40, 41] consider minimization problem over a finite frequency range by utilizing GKYP lemma, and it does not consider the non-ideal behavior of the output filter in $\Delta\Sigma$ modulator. On the contrary, we would like to obtain a near optimal solution by overcoming the identical order constraint, considering the output non-ideal filter $H[z]$ and expressing our minimization design problem into matrix inequalities without using the GKYP lemma. The methods [13, 40, 41] consider the min-max design problem, while our design problem is based on the minimization of the variance of the quantization noise.

To minimize the variance of the quantization noise, we can evaluate the H_2 norm of $H[z]R[z]$ numerically. The H_2 norm can be evaluated based on the Lemma 1 in Chapter 3, which gives us a BMI (3.12) and an LMI (3.13) of the function $\|H[z]R[z]\|_2^2 < \mu_2$.

Also, the condition on the stability of $\Delta\Sigma$ modulator can be described by using the BMI (3.24).

Then, we have to minimize μ_2 under the constraints (3.12), (3.13) and (3.24). The constraints (3.12) and (3.24) are BMIs since they contain products of the system variables and Lyapunov matrices. The BMI in (3.12) is bilinear due to the product between the variables P and A matrices. On the other hand, the BMI in (3.24), which results due the stability constraint, is bilinear due to the product between A_R and P_R matrices. The matrices A and A_R are unknown due to the presence of the denominator coefficients a_1, \dots, a_N in A_R . These BMIs are non-convex and NP-hard to solve numerically [42], hence, complicating the design of our noise shaping IIR filter.

When we assign a fix value to one of the unknown variables and optimize with respect to other variable, the resultant design problem becomes a semi-definite pro-

gramming problem consisting of LMIs only. If a_1, \dots, a_N in A_R are assigned constant values, then the BMIs turn into LMIs which can then be solved by using convex optimization techniques. In the following step, we optimize with respect to A_R and C_R by assigning a fix values to P and P_R obtained in the previous step.

Here, we summarize the iterative LMI algorithm to solve the noise shaping IIR design problem as follows:

1. Assign a fix value to the system variable A_R in the composite matrix A . This is done by initializing the denominator coefficients a_1, \dots, a_N with some prior values, which convert the BMIs (3.12), (3.24) into LMIs. Then, the design problem is reduced to convex form, which can then be optimized with respect to P , P_R and C_R variables.
2. Using the values of P and P_R obtained in step 1, optimize with respect to A_R and C_R to obtain N_A denominator and N_B numerator coefficients.
3. Obtain an IIR filter transfer function $R[z]$ by using the N_B numerator and N_A denominator coefficients in step 2. Then, we calculate the H_2 norm of $H[z]R[z]$.
4. Go to step 1 and use the values of a_1, \dots, a_N obtained in step 2. Repeat steps 1-3, until the H_2 norm of $H[z]R[z]$ reaches a given target or decreases less than a given accuracy.

The sequence of $\|H[z]R[z]\|_2^2$ generated by this algorithm is expected to be monotonously non-increasing. The objective function is bounded below by zero which implies that the proposed algorithm converges to some positive value as the iteration increases.

Initialization of the denominator coefficients

The choice of the initial values for the N_A denominator coefficients in the proposed algorithm is an important factor in the optimal design of the IIR filter. The N_A

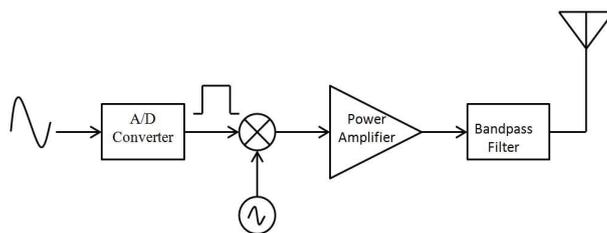


Figure 4.34: A simple block diagram of a $\Delta\Sigma$ modulator based RF transmitter.

denominator coefficients can be chosen as long as the poles of the resultant IIR filter $R[z]$ lie strictly inside the unit circle ($|z| < 1$) in the complex z -plane. However, the random selection of the poles inside the unit circle cannot guarantee us a near optimal solution. To address this issue, we suggest initializing the iterative algorithm with the denominator coefficients of the IIR filter obtained by using either the *LS method* or the *Hybrid design*.

Therefore, the LS method and the hybrid design can be used to design sub-optimal IIR filters whose denominator coefficients are used as initial points in the proposed algorithm. Then, the resultant IIR filter from our iterative LMI algorithm is an improvement over the sub-optimal IIR filters.

4.4.1 Design Examples

Let us consider a design example of a bandpass $\Delta\Sigma$ modulator based RF transmitter. Fig. 4.34 shows a block diagram of a basic wireless transmitter which consists of an A/D converter, a frequency-up converter and a power amplifier (PA) with a bandpass filter (BPF) at its output. The input to the PA is usually a signal with varying envelope, and if the PA is driven to more than its maximum input saturating power, it will cause distortion. The peak power of the input signal with varying envelope happens during very short periods, and most of the time the signal power remains around its average power. Since the average power is much smaller than its peak

power, the PA often works at much lower efficiencies than its maximum efficiency for a varying envelope signal. On the other hand, the constant envelope signal can make the PA work at maximum efficiency [43].

One approach to obtain a constant envelope signal at the input of the PA is to utilize a binary $\Delta\Sigma$ modulator as an A/D in the RF transmitter. By oversampling the information signal and using a binary quantizer, the information signal is encoded to a bi-level constant envelope signal. The main drawback of using a $\Delta\Sigma$ modulator is the quantization noise, which will be amplified alongside the desired signal by the PA. Although the BPF at the output of the PA is used to filter out the quantization noise, but there still remains a significant portion of the quantization noise in the information signal band which affects the overall efficiency of the RF transmitter.

To minimize the quantization noise present in the information signal band, we design noise shaping IIR filters in a $\Delta\Sigma$ modulator by using the proposed algorithm, hybrid design and Schreier's method. We consider a bandpass $\Delta\Sigma$ modulator whose input is an oversampled information signal with an OSR of 256. The bandwidth of the information signal is defined as $[\frac{\pi}{2} - \frac{\pi}{OSR}, \frac{\pi}{2} + \frac{\pi}{OSR}]$, with a center frequency $\frac{\pi}{2}$. The analog BPF at the output is considered to be a first order Butterworth filter with maximally flat response in the passband of the signal. For the stability of the bandpass $\Delta\Sigma$ modulator, the Lee coefficient γ is set to be 1.5 (3.52 dB) to limit the maximum out-of-band gain that can overload the quantizer.

Initialization using the least-squares (LS) method

First, let us utilize the LS method to obtain an IIR filter whose denominator coefficients are then chosen as the initial values a_1, \dots, a_{N_A} for the proposed iterative LMI algorithm.

Fig. 4.35 shows frequency responses of fourth order bandpass noise shaping IIR

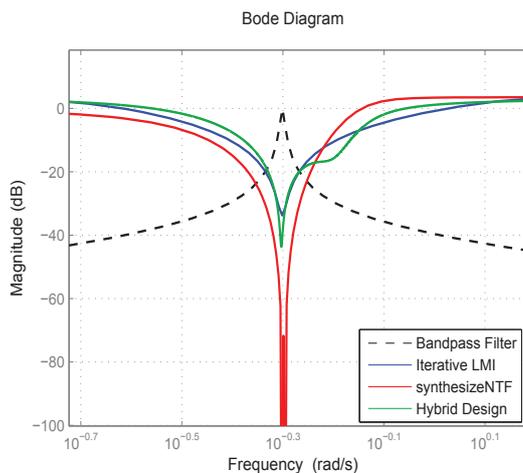


Figure 4.35: Frequency responses of noise shaping IIR bandpass filters obtained by using three different design methods. Here, iterative LMI algorithm is initialized using the LS method.

$\ H[z]R[z]\ _2^2$ norm values of design methods		
Design Methods	IIR Filter Order	H_2 norm (dB)
Iterative LMI	4	-41.79
Hybrid Design	4	-41.35
synthesizeNTF	4	-40.05

Table 4.3: Variance of the quantization noise ϵ , when iterative LMI is initialized using the LS method.

filters $R[z]$ obtained by using the iterative LMI algorithm, the hybrid design and Schreier's method (referred as **synthesizeNTF**). Since we take into account the non-ideal behavior of the bandpass filter $H[z]$, the steepness of the IIR filter $R[z]$ frequency response obtained by using the iterative LMI algorithm follows the steepness of the filter $H[z]$. Also, the iterative algorithm is observed to converge very quickly.

As listed in Table 4.3, the resultant variance of the quantization noise ϵ for the iterative LMI algorithm is -41.79 dB, while the values for the hybrid design and **synthesizeNTF** are -41.35 dB and -40.05 dB, respectively. Our proposed algorithm outperforms the hybrid design and **synthesizeNTF**. We also notice that the Lee

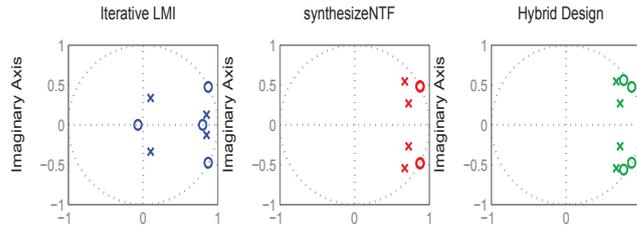


Figure 4.36: Zeros “o” and poles “x” of noise shaping IIR filters obtained by using the iterative LMI algorithm, `synthesizeNTF` and hybrid design methods.

criterion limits the magnitude of the frequency response, with maximum out-of-band gain not exceeding 3.52 dB for each frequency response. For comparison in Table 4.3, we do not list the ϵ of IIR filter obtained using the LS method because it does not satisfy the modulator stability constraint.

Fig. 4.36 shows the zero-pole plots of noise shaping IIR filters obtained by using the iterative LMI algorithm, `synthesizeNTF` and hybrid design methods. The poles “x” for all the designs lie inside the unit circle which verifies the stability of the $\Delta\Sigma$ modulator. Also, we notice that the zeros “o” of hybrid design are optimized to lie anywhere in the unit circle, while the zeros of `synthesizeNTF` are located on the unit circle. Thus, the hybrid design strategy gives us an IIR filter which gets best of both Schreier’s and optimal FIR methods which are limited in flexibility and requiring high orders [38], respectively.

Fig. 4.37 shows the quantization noise spectrum of the bandpass $\Delta\Sigma$ modulator with bi-level quantizer. The fourth order noise shaping filter minimizes the quantization noise present in the information signal band.

Initialization using the hybrid design

Now, let us design an IIR filter using the hybrid design whose denominator coefficients are then used as our initial values a_1, \dots, a_{N_A} . Note that, the hybrid design

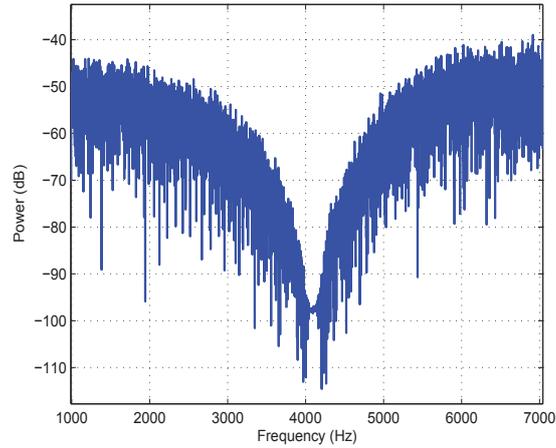


Figure 4.37: Noise Spectrum of the bandpass $\Delta\Sigma$ modulator.

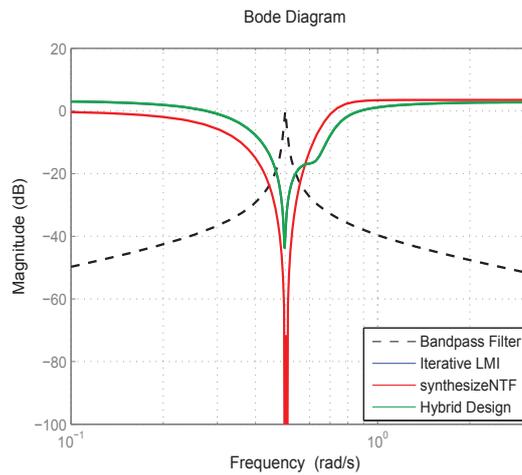


Figure 4.38: Frequency responses of noise shaping IIR bandpass filters obtained by using three different design methods. Here, iterative LMI algorithm is initialized using the hybrid design.

and the `synthesizeNTF` have similar denominator coefficients. Therefore, initialization using the hybrid design is similar to initializing with the `synthesizeNTF`.

Using the same $\Delta\Sigma$ modulator as in the previous section, we design fourth order bandpass noise shaping IIR filters $R[z]$. Fig. 4.38 shows the frequency responses of

$\ H[z]R[z]\ _2^2$ norm values of design methods		
Design Methods	IIR Filter Order	H_2 norm (dB)
Iterative LMI	4	-41.35
Hybrid Design	4	-41.35
<code>synthesizeNTF</code>	4	-40.05

Table 4.4: Variance of the quantization noise ϵ , when iterative LMI is initialized using the hybrid design.

bandpass noise shaping IIR filters obtained by using the iterative LMI algorithm, the hybrid design and the `synthesizeNTF`. The frequency responses of the iterative LMI design and the hybrid design are shown to be very similar to each other.

Table 4.4 lists the resultant variance of the quantization noise ϵ for the iterative LMI algorithm, the hybrid design and the `synthesizeNTF`. Our iterative LMI algorithm outperforms the `synthesizeNTF` design, but gives the same norm as that of the hybrid design. This implies that the iterative LMI algorithm cannot attain the optimum and shows that the initialization with the LS method is better for this example. However, we can say that the proposed algorithm provides the same performance as the hybrid design when initialized with the hybrid design denominator coefficients.

4.5 Conclusions

We have addressed the design of IIR noise shaping filters in the feedback of $\Delta\Sigma$ modulators. At first, we proposed a design based on extended LMIs which provide H_2 and H_∞ norm based filters that outperform the conventional design. We also utilized output weighting function to minimize the weighted quantization noise. However, the order of the IIR noise shaping filters is constrained to be identical to the order of the output weighting function. To overcome this constraint, we introduced approximation techniques like the Yule-Walker method and the LS approximation. Both approximation techniques require us to obtain a high-order optimal FIR filter which

is approximated by using the low-order IIR filter. Through simulations, we observe that the LS approximation technique outperforms the Yule-Walker method and gives us better approximation than the Yule-Walker method. However, the LS approximation does not guarantee $\Delta\Sigma$ modulator stability. Another method known as hybrid design is proposed to obtain sub-optimal IIR noise shaping filters. The hybrid design utilize the optimal FIR design to obtain numerator coefficients of the IIR noise shaping filter. The hybrid design IIR filter can outperform the conventional IIR filter. Then, we propose an iterative algorithm that results in an IIR noise shaping filter and also guarantees the stability of the $\Delta\Sigma$ modulator. The proposed algorithm obtains IIR noise shaping filter which is independent of the FIR filter design unlike approximation techniques. The iterative LMI is initialized with the LS approximation and the hybrid design. Through simulations, we observe that the iterative LMI algorithm gives us a near-optimal solution, while outperforming other techniques discussed in this chapter.

Chapter 5

Conclusions

We have proposed design methods to obtain FIR and IIR noise shaping filters for the minimization of the weighted quantization at the output of a $\Delta\Sigma$ modulator. Our noise shaping filters are based on the H_2 , H_∞ and l_1 norms of the weighted quantization noise. We have also considered the imperfect filter, as a weighting function, attached to the output of a $\Delta\Sigma$ modulator. The stability of the $\Delta\Sigma$ modulator is also ensured by limiting the maximum out-of-band gain of the NTF.

In Chapter 3, we have proposed a design method of the FIR noise shaping filters for $\Delta\Sigma$ modulators based on H_2 , H_∞ , and l_1 norms. The design problem for the minimization of the weighted quantization noise is cast into a convex optimization problem by using LMIs. Our results show that the frequency response of our FIR filters exhibits good performance in the low-frequency region providing uniform attenuation and matching the steepness of the weighting function. Our results show that the proposed method outperforms the existing methods for FIR filter design.

Since the design problem of IIR noise shaping filter cannot be cast into a convex optimization problem, we have proposed several design methods to obtain IIR filters in Chapter 4. The extended LMI technique is used to obtain IIR noise shaping filters by minimizing the variance and the l_2 norm of the quantization noise at the output of the $\Delta\Sigma$ modulator under the constraint on the feedback signal. Then,

the minimization of the error variance under the constraint on the variance of the error feedback signal can be cast into a convex optimization problem. Our proposed design is based on extended LMIs, which provides better filters than the design with conventional LMIs. Design examples are provided to demonstrate the effectiveness of our proposed extended LMIs design.

We have also utilized approximation techniques to obtain IIR noise shaping filters. The Yule-Walker and the LS methods are used to approximate the high-order FIR with low-order IIR filter. First, we have obtained optimal FIR noise shaping filter using the method proposed in Chapter 3, then we use both the Yule-Walker and the LS methods to obtain IIR noise shaping filters which approximate the frequency response of the optimal FIR filter. The LS method gives us better approximation of the FIR filter than the Yule-Walker method. Also, the LS gives us better attenuation of the weighted quantization noise than the Yule-Walker method. However, simulations show that the IIR noise shaping filter obtained by using the LS method does not necessarily satisfy the stability constraint of the $\Delta\Sigma$ modulator.

Moreover, we have proposed hybrid design technique which design IIR noise shaping filter whose numerator coefficients are obtained using the FIR method proposed in Chapter 3. Unlike the LS method, the hybrid design ensures the stability constraint on the $\Delta\Sigma$ modulator by limiting the maximum out-of-band gain of the $R[z]$.

An iterative LMI algorithm has been also proposed which converts the non-convex problem to convex form by converting BMIs into LMIs with alternation of unknown variables at each iteration. The hybrid design and the LS method are used to obtain the IIR filter whose denominator coefficients are used to initialize the proposed algorithm. Our design example utilizes the initial denominator values obtained by using the LS method and the hybrid design. Through simulations and analysis, we have observed that the noise shaping IIR filter obtained by using our the proposed

iterative LMI algorithm can outperform the hybrid design and `synthesizeNTF` if it is initialized with the LS method. If we use the hybrid design for initialization, the iterative LMI design gives same performance as the hybrid design while outperforming the `synthesizeNTF`.

Bibliography

- [1] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
- [2] J. C. Candy and G. C. Temes, *Oversampling Delta-Sigma Data Converters*. IEEE Press New York, 1991.
- [3] K. C. H. Chao, S. Nadeem, W. L. Lee, and C. G. Sodini, “A higher order topology for interpolative modulator for oversampling A/D conversion,” *IEEE Transactions on Circuits and Systems*, vol. 37, pp. 309–318, Mar. 1990.
- [4] E. Janssen, D. Reefman, “Super-audio CD: An introduction,” *IEEE Signal Processing Mag.* , vol. 20, no. 4, pp. 83-90, 2003.
- [5] U. Zoler, *Digital Audio Signal Processing*, 2nd ed. New York: Wiley, 2008.
- [6] R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters*. Wiley-IEEE Press, 2004.
- [7] Tran-Thong and B. Liu, “Error spectrum shaping in narrow-band recursive filters,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 25, no. 2, pp. 200–203, Apr. 1977.
- [8] T. Laakso and I. Hartimo, “Noise reduction in recursive digital filters using high-order error feedback,” *IEEE Transactions on Signal Processing*, vol. 40, no. 5, pp. 1096–1107, May 1992.

- [9] W. Higgins and J. Munson, D.C., “Noise reduction strategies for digital filters: Error spectrum shaping versus the optimal linear state-space formulation,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 30, no. 6, pp. 963–973, Dec. 1982.
- [10] M. Nagahara and Y. Yamamoto, “Frequency domain min-max optimization of noise-shaping delta-sigma modulators,” *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2828–2839, June 2012.
- [11] S. Callegari and F. Bizzarri, “Output filter aware optimization of the noise shaping properties of $\Delta\Sigma$ modulators via semi-definite programming,” *IEEE Transactions on Circuits and Systems I*, vol. 60, no. 9, pp. 2352–2365, Sept. 2013.
- [12] S. Callegari and F. Bizzarri, “Noise weighting in the design of $\Delta\Sigma$ modulators (with a psychoacoustic coder as an example),” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 60, no. 11, pp. 756–760, Nov. 2013.
- [13] X. Li, C. Yu, H. Gao “Design of delta-sigma modulators via generalized Kalman-Yakubovich-Popov lemma,” *Automatica*, vol. 50, pp. 2700–2708, 2014.
- [14] M. R. Tariq and S. Ohno, “ Unified LMI based design of $\Delta\Sigma$ modulators,” *EURASIP Journal of Advances in Signal Processing*, Feb. 2016.
- [15] S. Ohno and M. R. Tariq, “Optimization of noise shaping filter for quantizer with error feedback,” *IEEE Transactions on Circuits and systems I: Regular Papers*, Vol. 64, pp. 918–930, 2017.
- [16] M. R. Tariq and S. Ohno, “Synthesis of IIR error feedback filters for $\Delta\Sigma$ modulators using approximation,” *APSIPA Annual Summit and Conference, Jeju, Korea*, Dec.13-16, 2016.

- [17] M. R. Tariq and S. Ohno, “Approximation of FIR by IIR digital filters in quantizers with error feedback,” *SICE annual conference, Tsukuba, Japan*, 2016.
- [18] M. R. Tariq and S. Ohno, “An indirect approach to synthesis of noise shaping IIR filters in $\Delta\Sigma$ modulators,” *IEEE International Symposium on Circuits and Systems (ISCAS), Baltimore, USA*, 2017.
- [19] M. R. Tariq and S. Ohno, “An iterative LMI algorithm for quantization noise reduction in $\Delta\Sigma$ modulators,” *Signal Processing*, 2018
- [20] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.0 beta,” <http://cvxr.com/cvx>, Sept. 2012.
- [21] S. Park, *Principles of Sigma-Delta Modulation for Analog to Digital converters*. Motorola.
- [22] K.C.H. Chao, S. Nadeem, W.L. Lee and C.G. Sodini, “A higher order topology for interpolative modulators for oversampling A/D converters,” *IEEE Trans. Circuits Syst.*, vol. 37, no. 3, pp. 309-318, Mar. 1997.
- [23] S. Boyd, L. E. Ghaoul, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics, 1997.
- [24] P. Gahinet and P. Apkarian, “A linear matrix inequality approach to H_∞ control,” *International Journal of Robust and Nonlinear Control*, vol. 4, pp. 421-448, 2013.
- [25] K. Sawada and S. Shin, “Synthesis of dynamic quantizers for quantized feedback systems within invariant set analysis framework,” in *American Control Conference (ACC), 2011*, June 2011, pp. 1662–1667.

- [26] S. Ohno, Y. Wakasa and M. Nagata, “Optimal error feedback filters for uniform quantizers at remote sensors,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2015, pp. 3866–3870.
- [27] H. Shingin and Y. Ohta, “Optimal invariant sets for discrete-time systems: Approximation of reachable sets for bounded inputs,” *10th I-FAC/IFORS/IMACS/IFIP Symposium on Large Scale systems: Theory and Applications (LSS)*, pp. 401-406, 2004.
- [28] R. Schreier, Delta Sigma Toolbox,
<https://jp.mathworks.com/matlabcentral/fileexchange/19-delta-sigma-toolbox>
- [29] E. Kaszkurewicz and A. Bhaya, *Matrix diagonal stability in systems and computation*. Springer Science & Business Media, 2012.
- [30] I. Masubuchi, A. Ohara, and N. Suda, “LMI-based controller synthesis: A unified formulation and solution,” *International Journal of Robust and Nonlinear Control*, vol. 8, no. 8, p. 669–686, July 1998.
- [31] C. Scherer, P. Gahinet, and M. Chilali, “Multiobjective output-feedback control via LMI optimization,” *IEEE Transactions on Automatic Control*, vol. 42, no. 7, pp. 896–911, Jul 1997.
- [32] M. C. D. Oliveira, C. Gerome, and J. Bernussou, “Extended H_2 and H_∞ norm characterizations and controller parametrizations for discrete-time systems,” *International Journal of Control*, vol. 75, no. 9, pp. 666–679, 2002.
- [33] A. Gersho and R. M. Gray, *Vector quantization and signal compression*. Springer Science & Business Media, 2012.

- [34] Mathworks, Documentation of Fixed-Point Designer, <https://www.mathworks.com/help/fixedpoint/index.html>, 2016.
- [35] H.K. Kwan and A. Jiang, "Recent advances in FIR approximation by IIR digital filters," *International Conference on Communications, Circuits and Systems*, vol. 1, pp. 185-190, Jun. 2006.
- [36] H. Brandenstein and R. Unbehauen, "Least-squares approximation of FIR by IIR digital filters," *IEEE Trans. Signal Processing*, vol. 46, pp. 21-30, Jan. 1998.
- [37] P. Gahinet, A. Nemirovskii, A.J. Laub and M. Chilali, "LMI control toolbox user's guide," *The Math Works Inc.*, 1995.
- [38] S. Callegari and F. Bizzarri, "Optimal desing of the noise transfer function of $\Delta\Sigma$ modulators: IIR strategies, FIR strategies with preassigned poles," *Signal processing*, vol. 114, pp. 117-130, 2015.
- [39] LM4670 filterless high efficiency 3W switching audio amplifier, Texas Instruments, TI literature number SNAS240B, Jun.2006.
- [40] X. Li, S. Yin, H. Gao, "Passivity-preserving model reduction with finite frequency H_∞ approximation performance," *Automatica*, Vol. 50, pp. 2294-2303, 2014.
- [41] H. Gao, X. Li, J. Qiu, "Finite frequency H_∞ deconvolution with application to approximated bandlimited signal recovery," *IEEE Transactions on Automatic Control*, Vol. 99, 2017.
- [42] A. Hassibi, J. How, S. Boyd, "A path-following method for solving BMI problems in control," *Proceedings of the American Control Conference, USA*, 1999.
- [43] J. S. Walling, D. J. Allstot, "Pulse-width modulated CMOS power amplifiers," *IEEE Microwave Magazine*, Vol. 12, pp.52-60, 2011.

Journal Publications

1. M.R. Tariq, S. Ohno, “An iterative LMI algorithm for quantization noise reduction in $\Delta\Sigma$ modulator”, *Signal Processing*, Vol. 144, pp. 163-168, 2018.
(<https://doi.org/10.1016/j.sigpro.2017.10.018>)
2. S. Ohno, T. Shiraki, M.R. Tariq, M. Nagahara, “Mean squared error analysis of quantizers with error feedback”, *IEEE Transactions on Signal Processing*, Vol. 65, pp. 5970-5981, 2017.
(DOI: 10.1109/TSP.2017.2745450)
3. S. Ohno, M.R. Tariq, “Optimization of noise shaping filter for quantizer with error feedback”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 64, pp. 918-930, 2017.
(DOI: 10.1109/TCSI.2016.2627031)
4. M.R. Tariq, S. Ohno, “Unified LMI-based design of $\Delta\Sigma$ modulators”, *EURASIP Journal of Advances in Signal Processing*, 2016.
(<https://doi.org/10.1186/s13634-016-0326-2>)

Conference Papers

1. S. Ohno, M.R. Tariq, M. Nagahara, “Min-max IIR filter design for feedback quantizers”, *APSIPA Annual Summit and Conference*, Kuala Lumpur, Malaysia, 2017.
2. M.R. Tariq, S. Ohno, “Design of a continuous-time loop filter for $\Delta\Sigma$ modulators with excess loop delay”, *The 49th ISICIE International Symposium on Stochastic Systems Theory and Its Applications*, Hiroshima, Japan, 2017.
3. M.R. Tariq, S. Ohno, “An indirect approach to synthesis of noise shaping IIR filters in $\Delta\Sigma$ modulators”, *IEEE International Symposium on Circuits and Systems (ISCAS)*, Baltimore, USA, 2017.
4. S. Ohno, T. Shiraki, M.R. Tariq, M. Nagahara, “Rate-distortion analysis of delta-sigma modulators”, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, USA, 2017.
5. M.R. Tariq, S. Ohno, “Synthesis of IIR error feedback filters for $\Delta\Sigma$ modulators using approximation”, *APSIPA Annual Summit and Conference*, Jeju, Korea, 2016.
6. M.R. Tariq, S. Ohno, “Approximation of FIR by IIR digital filters in quantizers with error feedback”, *SICE Annual Conference*, Tsukuba, Japan, 2016.