Designs of Orthogonal Filter Banks and Orthogonal Cosine-Modulated Filter Banks

by

Jie Yan
B.Eng., Southeast University, Nanjing, China, 2008

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF APPLIED SCIENCE

in the Department of Electrical and Computer Engineering

Designs of Orthogonal Filter Banks and Orthogonal Cosine-Modulated Filter Banks

by

Jie Yan
B.Eng., Southeast University, Nanjing, China, 2008

Supervisory Committee

Dr. Wu-Sheng Lu, Supervisor
(Department of Electrical and Computer Engineering)

Dr. Pan Agathoklis, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Dale Olesky, Outside Member
(Department of Computer Science)

Supervisory Committee

Dr. Wu-Sheng Lu, Supervisor
(Department of Electrical and Computer Engineering)

Dr. Pan Agathoklis, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Dale Olesky, Outside Member
(Department of Computer Science)

ABSTRACT

This thesis investigates several design problems concerning two-channel conjugate quadrature (CQ) filter banks and orthogonal wavelets, as well as orthogonal cosine-modulated (OCM) filter banks.

It is well known that optimal design of CQ filters and wavelets and optimal design of prototype filters (PFs) of OCM filter banks in the least squares (LS) or minimax sense are nonconvex problems and to date only local solutions can be claimed. In this thesis, we first make some improvements over several direct design techniques for local design problems in terms of convergence and solution accuracy. By virtue of the recent progress in global polynomial optimization and the improved local design methods mentioned above, we describe an attempt at developing several design strategies that may be viewed as our endeavors towards global solutions for LS CQ filter banks, minimax CQ filter banks, and OCM filter banks. In brief terms, the proposed design strategies are based on several observations made among globally optimal impulse responses of low-order filter banks, and are essentially order-recursive algorithms in terms of filter length combined with some techniques in identifying a desirable initial point in each round of iteration.

This main idea is applied to three design scenarios in this thesis, namely, LS design of orthogonal filter banks and wavelets, minimax design of orthogonal filter banks and wavelets, and design of orthogonal cosine-modulated filter banks. Simulation studies are presented to evaluate and compare the performance of the proposed design methods with several well established algorithms in the literature.

Contents

Supervisory Committee **ii**

Abstract **iii**

Table of Contents **iv**

List of Tables **vii**

List of Figures **viii**

Acknowledgements **xii**

Dedication **xiii**

1 Introduction **1**

1.1 Two-Channel Orthogonal Filter Banks 1

1.2 Orthogonal Cosine-Modulated Filter Banks 3

1.3 Overview and Contribution of the Thesis 5

2 Preliminaries **9**

2.1 Two-Channel Orthogonal Filter Banks 9

2.2 Orthogonal Cosine-Modulated Filter Banks 12

2.3 Gauss-Newton Method with Adaptively Controlled Weights 16

2.4 Convex Quadratic Programming and Second-Order Cone Programming 19

2.5 General Nonlinear Optimization Problems 21

2.6 Global Optimization of Small-Size Polynomial Optimization Problems 23

3 Least Squares Design of Orthogonal Filter Banks and Wavelets **25**

3.1 Local LS Design of CQ Filter Banks 26

3.1.1 Sequential Convex-Programming Method 26

3.1.2 Sequential Quadratic-Programming Method 30
3.1.3 Experimenal Results of Local LS Designs 31
3.2 Global LS Design of Low-Order CQ Filter Banks 34
3.3 Potentially Global LS Design of High-Order CQ Filter Banks 36
3.3.1 Pattern of Impulse Responses of Globally Optimal Low-Order Filter Banks 36
3.3.2 A Design Strategy 38
3.4 Design Examples and Performance Evaluation 41
3.4.1 Performance of the Proposed Method for Globally Optimal Low-order Designs 41
3.4.2 Performance of the Proposed Method for Potentially Globally Optimal High-order Designs 42

4 Minimax Design of Orthogonal Filter Banks and Wavelets 47
4.1 Local Minimax Design of CQ Filter Banks 48
4.1.1 Direct Design Method with Feasible Optimization Problems . 48
4.1.2 A Convergence Issue for Local Minimax Designs 49
4.1.3 Experimental Results of Local Minimax Designs 50
4.2 Global Minimax Design of Low-Order CQ Filter Banks 54
4.3 Potentially Global Minimax Design of High-Order CQ Filter Banks . 54
4.4 Design Examples and Performance Evaluations 57
4.4.1 Performance of the Proposed Method for a Low-order Design . 57
4.4.2 Performance of the Proposed Methods for High-order Designs 58
4.5 Comparisons with Other Existing Methods 62
4.5.1 Comparison with a Half-Band Filter Based on the Method in [41] 62
4.5.2 Comparison with the Method of Smith-Barnwell 62

5 Design of Orthogonal Cosine-Modulated Filter Banks 65
5.1 Local Design of OCM Filter Banks 66
5.2 Global Design of Low-Order OCM Filter Banks 67
5.3 Potentially Global Design of High-Order OCM Filter Banks 68
5.3.1 An improvement in initial point when $m = 1$ 72
5.4 Design Examples and Performance Evaluation 73
5.4.1 Performance of the Proposed Method for Low-Order Designs . 73
5.4.2 Performance of the Proposed Method for High-Order Designs 74

6 Conclusions and Future Research **87**

6.1 Conclusions 87

6.2 Future Research 89

A Generating matrices $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}_{l,n}$ for problem (2.19) **90**

Bibliography **92**

List of Tables

Table 3.1 Performance of LS filters designed using different local methods 34
Table 3.2 Performance of globally optimal LS filters of length 96 42
Table 3.3 Performance of a locally optimal LS filter of length 96 44

Table 4.1 Performance of minimax filters designed using different local methods 51
Table 4.2 Performance of the minimax filters from global designs 60
Table 4.3 Performance of the length-96 minimax filter from local design . 60
Table 4.4 Coefficients of $H_0(z)$ of [41] and from global design 63
Table 4.5 Filters performance comparison 63
Table 4.6 Filter length $N = 8$ 64
Table 4.7 Filter length $N = 16$ 64
Table 4.8 Filter length $N = 32$ 64

Table 5.1 Performance comparison for OCM filter banks with $m = 20$, $M = 4$ and $\rho = 1$. 74
Table 5.2 Performance comparison for OCM filter banks with $m = 12$, $M = 16$ and $\rho = 1$. 79
Table 5.3 Performance comparison for OCM filter banks with $m = 7$, $M = 32$ and $\rho = 1$. 86

List of Figures

Figure 1.1 A two-channel CQ filter bank. 2
Figure 1.2 A subband system with a 2-level tree. 2
Figure 1.3 An M-channel maximally decimated filter bank. 4
Figure 1.4 A gradient-based descent algorithm starts at initial point x_a and produces a local solution x_a^*; the same algorithm converges to the global solution if it starts at x_b which falls into a good region (shaded in the figure). 6

Figure 2.1 A 2-channel CQ filter bank. 10
Figure 2.2 An M-channel maximally decimated filter bank. 13
Figure 2.3 Second-order cone in $\mathcal{R}^3$. 20

Figure 3.1 Magnitude response of local LS filter with $N = 30$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the SCP method. 32
Figure 3.2 Impulse response of local LS filter with $N = 30$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the SCP method. 32
Figure 3.3 Magnitude response of local LS filter with $N = 30$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the direct method. 33
Figure 3.4 Impulse response of local LS filter with $N = 30$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the direct method. 33
Figure 3.5 Zero-pole plot of $\mathbf{h}_{\mathrm{LS}}^{(6,2)}$. 35
Figure 3.6 Pattern of LS impulse responses with different length N. 37
Figure 3.7 Pattern of LS impulse responses with various VM L. 37
Figure 3.8 Zero-padded and linearly interpolated initial impulse responses with respect to the globally optimal impulse response of that order. 39
Figure 3.9 Comparison of zero-padded initial point and linearly interpolated initial point in terms of their ℓ_2 distances to globally optimal point for LS designs. 39

Figure 3.10 Magnitude response of globally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 43
Figure 3.11 Impulse response of globally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 43
Figure 3.12 Magnitude response of locally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 45
Figure 3.13 Impulse response of locally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 45
Figure 3.14 Zero-pole plot of globally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 46
Figure 3.15 Zero-pole plot of locally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 46

Figure 4.1 Magnitude response of local minimax filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the SCP-GN method. 52
Figure 4.2 Impulse response of local minimax filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the SCP-GN method. 52
Figure 4.3 Magnitude response of local minimax filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the direct method. 53
Figure 4.4 Impulse response of local minimax filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the direct method. 53
Figure 4.5 Zero-pole plot of $\mathbf{h}_{\text{minimax}}^{(4,1)}$. 55
Figure 4.6 Pattern of minimax impulse responses with different length N. 56
Figure 4.7 Comparison of zero-padded initial point and linearly interpolated initial point in terms of their ℓ_2 distances to globally optimal point for minimax designs. 56
Figure 4.8 Magnitude response of globally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 59
Figure 4.9 Impulse response of globally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 59
Figure 4.10 Magnitude response in passband of globally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 60
Figure 4.11 Zero-pole plot of globally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 61

Figure 4.12 Zero-pole plot of locally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. 61
Figure 4.13 Magnitude response of minimax $H_0(z)$ with $N = 20$ from the global design (solid line) versus that of $H_0(z)$ in [41] (dashed line). 64

Figure 5.1 Pattern of impulse responses of globally optimal PFs. 69
Figure 5.2 Effect of linear interpolation when $m = 1$. 71
Figure 5.3 Effect of zero-padding when $M = 4$. 71
Figure 5.4 Effect of downshifting the linearly interpolated point when $m = 1$. 73
Figure 5.5 Magnitude response of globally optimal PF of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$. 75
Figure 5.6 Impulse response of globally optimal PF of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$. 75
Figure 5.7 Magnitude response of locally optimal PF of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$. 76
Figure 5.8 Impulse response of locally optimal PF of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$. 76
Figure 5.9 Magnitude responses of analysis filters of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$. 77
Figure 5.10 Amplitude distortion of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$. 78
Figure 5.11 Group-delay distortion of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$. 78
Figure 5.12 Worst case aliasing error of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$. 79
Figure 5.13 Magnitude response of globally optimal PF of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$. 80
Figure 5.14 Impulse response of globally optimal PF of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$. 80
Figure 5.15 Magnitude responses of analysis filters of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$. 81
Figure 5.16 Amplitude distortion of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$. 82
Figure 5.17 Group-delay distortion of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$. 82

Figure 5.18 Worst case aliasing error of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$. 83
Figure 5.19 Magnitude response of globally optimal PF of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$. 84
Figure 5.20 Impulse response of globally optimal PF of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$. 84
Figure 5.21 Magnitude response of locally optimal PF of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$. 85
Figure 5.22 Impulse response of locally optimal PF of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$. 85
Figure 5.23 Magnitude responses of analysis filters of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$. 86

ACKNOWLEDGEMENTS

How time flies and it has been one year and a half since I came to the University of Victoria in Canada for graduate studies. At this moment, it is an honor for me to take the opportunity to express my hearty gratitude to my supervisor Dr. Wu-Sheng Lu who has been guiding me throughout this work. Not only for his wisdom and wide range of knowledge, but also his unquenched curiosity and wholehearted dedication to research manifest to me the spirit of a real scientist. From him I discovered the wonderland of the area of digital signal processing, and made up my mind to develop my potential to be a scientific researcher. I am fortunate to have him as my supervisor.

I would like to thank Dr. Pan Agathoklis and Dr. Dale Olesky for serving as my committee members. I learned useful knowledge from their courses and I am influenced by the two wonderful professors in a number of ways. My thanks also go to Dr. Yang Shi for being my external examiner.

It is a pleasure to express my gratitude to my colleagues and friends who made this thesis possible: Chi-Tang Catherine and Ana-Maria for bringing me an agreeable environment in the office where I completed most of the work; Yihai Zhang for his warmhearted assistance in helping me quickly adapted to the life and study in Canada; and Zhong Zhuang for inviting me for a walk many times to show me the amazing nature around campus, in the meantime broadening my perspective in social science and history. I would also like to thank numerous other friends who have helped me in one way or another.

I am also grateful to the Department of Electrical and Computer Engineering along with the faculty and staff of the University of Victoria who have provided guidance and assistance countless times over the years.

My deepest gratitude goes to my parents and my sister who constantly support me when I am in need. Without your love and encouragement, I would not have been where I am today.

DEDICATION

To my parents

Chapter 1

Introduction

Multirate digital signal processing finds applications in speech and image compression, digital audio industry, statistical and adaptive signal processing, and many other fields. Digital filter banks are at the core of many multirate systems. As such, developing effective methods for the design of filter banks with improved performance has been an active research field over the last three decades, and it is also the focus of the present thesis.

In this chapter, the two classes of filter banks to be studied in the thesis, namely the two-channel orthogonal filter banks and orthogonal cosine-modulated filter banks, are introduced in the first two sections. The main contribution of this thesis is our endeavor to develop strategies towards globally optimal designs for these two types of filter banks. A brief overview of the global design idea utilized throughout this thesis is introduced in Sec. 1.3. Finally, Sec. 1.3 lists the contributions of our work reported in the thesis, and conludes this chapter with an outline of the thesis.

1.1 Two-Channel Orthogonal Filter Banks

As illustrated in Fig. 1.1, a two-channel orthogonal filter bank is a filter structure which consists of the analysis filter bank and synthesis filter bank. In the analysis filter bank, an input signal is first split into its lowpass and highpass components, each component is then downsampled by a factor of two. In the synthesis filter bank, the subband signal is upsampled by a factor of two, followed by a filtering process and finally the filtered signals are integrated to reconstruct the original signal. The process of splitting and downsampling (as well as upsampling and integrating) can

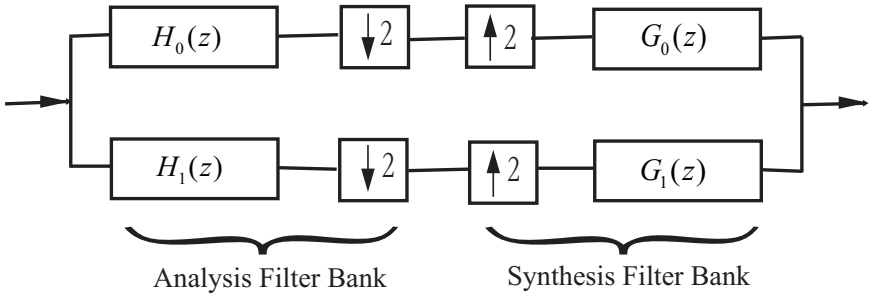

Figure 1.1: A two-channel CQ filter bank.

be repeated for the subband signal and a tree-structured filter bank is formulated in this way, see e.g., Fig. 1.2 representing a 2-level tree structure.

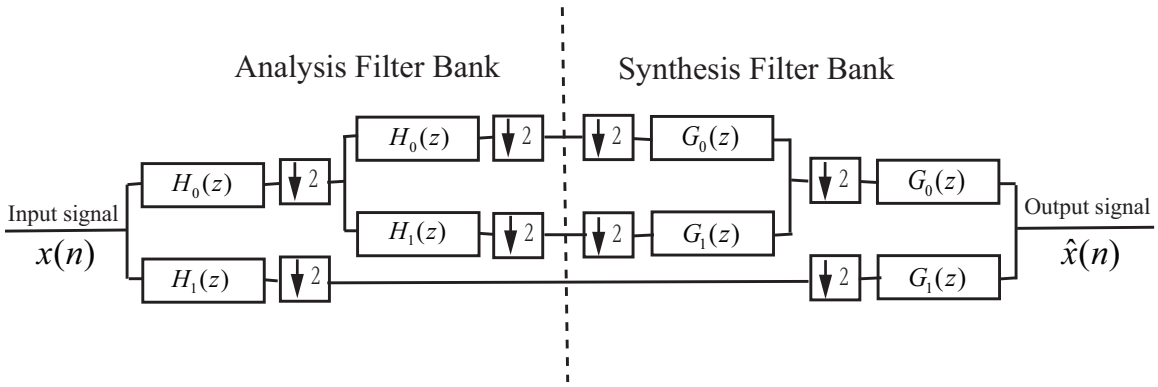

Figure 1.2: A subband system with a 2-level tree.

The class of two-channel conjugate quadrature (CQ) filter banks, also known as power-symmetric filter banks [41], is one of the most well-known building blocks for multirate systems and wavelet-based coding systems as it offers perfect reconstruction (PR) and other desirable properties (such as possessing a certain number of vanishing moments (VMs)), and has wide applications in general areas of digital signal processing, especially in image/speech coding. Many algorithms for the design of CQ filters have been proposed since 1980's, see e.g. references [16, 33, 32, 24, 38, 39, 41] and the work cited therein. In addition, design techniques aimed at FIR compaction filters, which are also applicable to CQ filters, have been investigated in [40] and [6]. The dominating design method in the literature is an indirect methodology in which the design is accomplished in two steps: constructing a halfband filter (subject to certain nonnegativity constraint) followed by spectrum factorization of the halfband filter.

With the astonishing growth of numerical optimization techniques, the task of

optimal design of CQ filters becomes attainable. Generally, two design scenarios are considered, namely,

1. The least squares (LS) design: to design a CQ filter whose stopband energy is minimized [39].

2. The minimax design: to design a CQ filter whose maximum stopband energy is minimized [33, 32, 38, 41].

The reason we address the LS and minimax designs in this thesis is because in digital filters, the magnitude of the largest amplitude-response error is usually required to be as small as possible, thus minimax solutions are preferred [2]. On the other hand, in several applications, especially in telecommunications, digital filters are required to have minimal stopband energy, hence LS solutions are of importance in these applications. A recent progress in this field is the direct design technique proposed in [16], in which halfband filter and spectrum factorization are not required. Instead, a CQ filter is directly optimized subject to the PR and possibly other constraints (such as possessing a certain number of VMs). By implementing the direct design technique, the solution is approached sequentially with each update confined within a small vicinity of the current iterate such that the problem at hand behaves like a convex one. As a result, the update can be obtained as a solution of a convex problem. This proposed method is found useful in designing high-order CQ filters with improved performance.

In this thesis, several local and global methods for the LS and minimax designs of CQ filter banks will be investigated. For local designs, we propose strategies to improve the direct design techniques in [16] so as to make the design methods more robust so that locally optimal filters can be obtained with a high degree of accuracy. More importantly, we will develop methods towards global designs of LS and minimax CQ filter banks.

1.2 Orthogonal Cosine-Modulated Filter Banks

An orthogonal cosine-modulated (OCM) filter bank is an M-channel orthogonal filter bank that consists of the analysis and synthesis filter bank, as illustrated in Fig. 1.3, where each filter is a cosine-modulated version of a common prototype filter (PF).

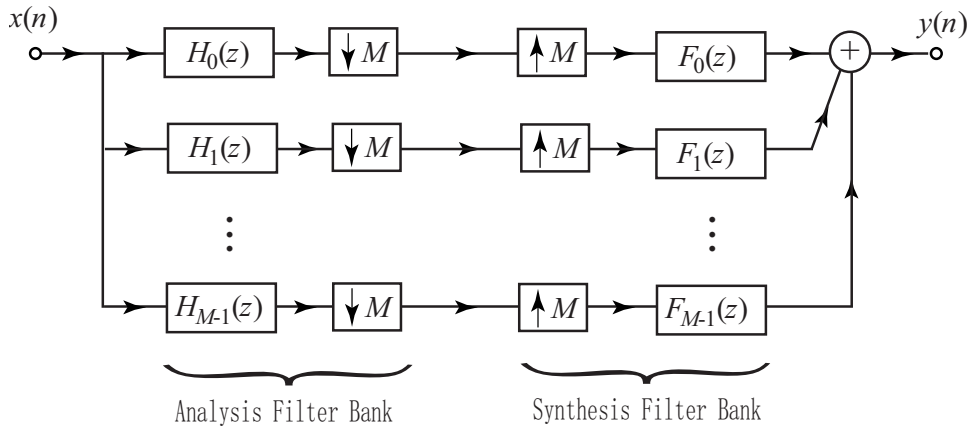

Figure 1.3: An M-channel maximally decimated filter bank.

Orthogonal cosine-modulated (OCM) filter banks are among the most popular filter banks for multirate signal processing as they admit efficient implementation through polyphase decomposition. In addition, optimal synthesis of an M-channel, maximally decimated, OCM multirate system can be carried out with considerably reduced complexity relative to that of a general M-channel system because in the former case one is only focused on a single PF which characterizes the M-channel OCM system. Besides, the PF of an OCM filter bank is a linear-phase FIR filter whose impulse response is symmetrical, thus the design variables are reduced to the first half of the PF's impulse response. The perfect reconstruction (PR) property is maintained and the reconstruction delay of an OCM filter bank is fixed to be the PF order. These benefits, among other things, have rendered the OCM filter banks one of the most useful classes of multirate systems [17, 48]. As a matter of fact, OCM filter banks are widely used in audio, image and video signals coding, as well as applied in communication systems, for instance, in transmultiplexers and multicarrier modulations [31].

A great deal of research on optimal design of OCM filter banks has been made [17, 10, 11, 35]. Available techniques include the quadratic-constrained least-squares (QCLS) method that minimizes the stopband energy of the PF subject to the PR constraints [10, 28, 29]; the factorization-based method [36, 37, 11] in which the stopband energy of the PF is minimized in an unconstrained optimization setting; and the sequential design method [3] that is carried out by gradually increasing the number of channels as well as the filter length employing a technique proposed in [34]. In [18], a convex Lagrangian relaxation method is used to obtain a near PR (NPR) filter

bank, and a technique of alternating null-space projections is further applied as a post-processing step which turns the NPR filter bank into a PR one. In addition, quadratic-constrained-optimization based algorithms [30] and fast designs by optimizing window functions [5, 21] for OCM filter banks have been proposed. In particular, two second-order cone programming (SOCP) based algorithms for designing optimal PR and near PR (NPR) cosine-modulated filter banks have been proposed in [17]. The design technique employed in [17] is in spirit similar to that proposed in [16], which directly minimizes the PF's stopband energy subject to the PR constraints. This algorithm is demonstrated to produce PR OCM filter banks with improved performance compared to several establised design methods.

We in this thesis investigate local and global methods for the design of OCM filter banks. Improvements over the direct technique in [17] are made so that locally optimal OCM filter banks better satisfying the constraints can be designed. Specifically, an order-recursive strategy is proposed for the global design of OCM filter banks.

1.3 Overview and Contribution of the Thesis

In the literature, only locally optimal designs are proposed for two-channel orthogonal filter banks and OCM filter banks. A main contribution of this thesis is the development of several design strategies for potentially globally optimal designs of these two types of filter banks. We provide experimental evidences that are supportive of our speculation.

As the optimization problems for designs of CQ filter banks and OCM filter banks are nonconvex, a primary challenge facing us in finding a global solution is the existene of multiple local solutions, each representing a design of degraded performance, to which iterates generated by a gradient-based algorithm can easily be attracted especially when it starts in a not-so-good region. The scenario is illustrated in Fig. 1.4 where a typical gradient-based algorithm leads initial point x_a to a local and suboptimal solution x_a^*. On the other hand, Fig. 1.4 is also indicative of a solution approach to the global solution, that is to start the algorithm from an initial point like x_b in a good region. The problem is, obviously, how to identify such an initial point.

Fig. 1.4 is only illustrative because it is about a one-variable case without constraints whereas problems related to designs of CQ and OCM filter banks are multivariable problems with nonlinear and nonconvex constraints, yet it does give us some

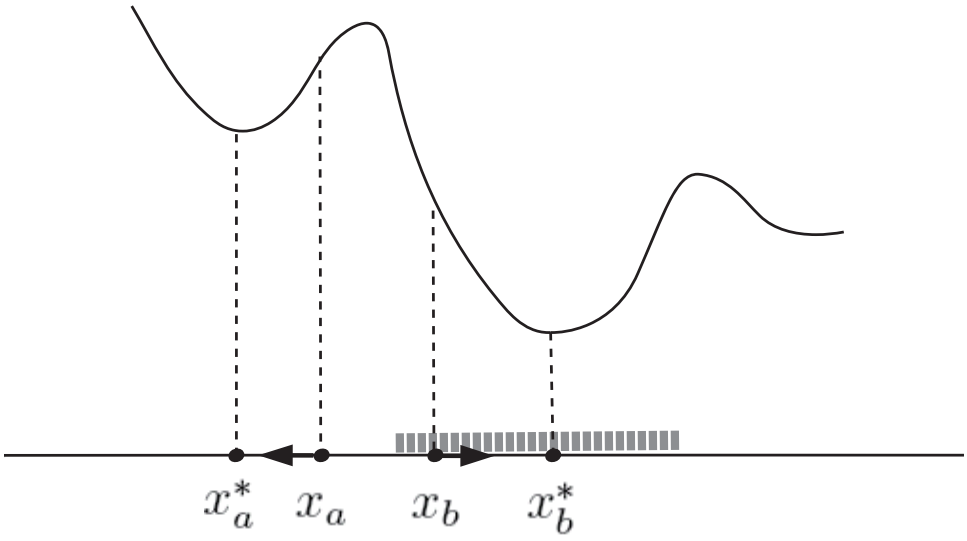

Figure 1.4: A gradient-based descent algorithm starts at initial point x_a and produces a local solution x_a^*; the same algorithm converges to the global solution if it starts at x_b which falls into a good region (shaded in the figure).

insight about the global solution of a general nonconvex problem. A central point in this thesis is to substantiate this notion by developing techniques to secure a good initial point for the design problems concerned.

More specifically, the global design methods proposed in this thesis are made possible by virtue of recent progress in global polynomial optimization [22, 23] and direct design techniques for local designs of CQ and OCM filter banks [16, 17], in conjuction with our observations on a common pattern shared among globally optimal impulse responses of low-order CQ filters (or among low-order PFs of OCM filter banks). Based on a technique proposed for identifying a desirable initial point in each round of the iteration, a progressive design procedure in terms of filter length is proposed for global designs of high-order CQ filter banks. In a similar spirit, an order-recursive strategy is developed for global designs of high-order OCM filter banks.

In addition to the main contributions to global designs of CQ and OCM filter banks, we also make improvements over the direct design techniques in [16, 17] which are known to obtain locally optimal solutions.

Experimental results are presented to demonstrate the performance of the improved local design methods and the proposed global design strategies for LS design of CQ filter banks, minimax design of CQ filter banks, and designs of OCM filter

banks.

The remainder of this thesis is organized as follows.

Chapter 2 presents the preliminaries that provide background for the problems to be studied in this thesis. We introduce basic structures and properties of a two-channel orthogonal filter bank and an orthogonal cosine-modulated filter bank, as well as the associated optimization problems. The Gauss-Newton method with adaptively controlled weights is described. Then we review two classes of convex programming, namely convex quadratic programming and second-order cone programming. We also sketch the sequential quadratic-programming method for general nonlinear optimization problems. At last, concepts for global optimization of polynomial optimization problems (POPs) and software for solving POPs of small size are introduced.

Chapter 3 investigates several methods for local and global LS designs of CQ filter banks. By improving the direct design technique in [16], we study two local methods, i.e., the sequential convex-programming method and the sequential quadratic-programming method. Then, global LS designs for CQ filter banks of low order are carried out using software that can solve POPs of small size. Based on the globally optimal impulse responses of low-order CQ filter banks, we observe a common pattern shared among them. As a result, a design strategy by generating an initial point for a CQ filter bank of higher order, which is believed to be sufficiently close to the global solution point of the CQ filter bank at that order, is proposed. Two techniques for generating an inital point are investigated: the zero-padding technique and the interpolation technique. It is found that the initial point produced by the zero-padding method is much closer to the global solution point than that generated by the interpolation method. Consequently, global LS designs of high-order CQ filter banks are implemented by carrying out the local design methods with a good initial point. Finally, performance of the proposed design algorithm is evaluated.

Chapter 4 is concerned with local and global minimax designs of CQ filter banks. First, we briefly review the direct design technique in [16]. Several improvements are made to the direct method for the local algorithm to achieve convergence at a small specified tolerance. Afterwards, global minimax designs of low-order and high-order CQ filters are investigated. Two methods towards globally optimal minimax filters of high order are proposed. Method 1 is similar to the

strategy utilized in the global LS designs, and it is also found that the zero-padded initial point is a more preferable choice than the linearly interpolated initial point. Method 2 is carried out by passing the globally optimal impulse response obtained from the LS design as the initial point for local minimax design method. In this way, only one round of iteration of the local minimax method is required. We have provided several design examples, and made comparisons with other existing methods well established in the literature.

Chapter 5 is dedicated to local and global designs of OCM filter banks. We first introduce the locally optimal design method which is in spirit similar to that proposed in [16] and in fact an improved version of the direct method reported in [17]. Then, with the observation among impulse responses of PFs of globally optimal low-order OCM filter banks, an order-recursive algorithm in terms of filter length in carrying out potentially global design of high-order OCM filter banks is proposed. The zero-padding method and the interpolation method are both found useful at different stages of the algorithm in generating good initial points for filters of higher order. In addition, the Gauss-Newton method with adaptively controlled weights is further implemented to downshift the linearly interpolated point so that the result becomes extremly close to the global solution point. This chapter concludes with several design examples and performance evaluations.

Chapter 6 summarizes the main ideas of this thesis and suggests several directions for future research.

Chapter 2

Preliminaries

In this chapter, we present preliminaries that provide background for the problems to be studied in the subsequent chapters of the thesis. We begin with introducing basic structures and properties of a two-channel orthogonal filter bank as well as related design problems. Next, the design of orthgonal cosine-modulated (OCM) filter banks is described and the associated optimization problem is formulated. The Gauss-Newton (G-N) method with adaptively controlled weights which will be found useful in Chapters 4 and 5 is introduced in Sec. 2.3. Two classes of convex programming, namely convex quadratic programming (QP) and second-order cone programming (SOCP), that facilitate the development of several design algorithms in Chapters 3, 4 and 5, are reviewed in Sec. 2.4. The sequential quadratic-programming (SQP) method for general nonlinear optimization problems, which will be applied in Chapter 3, is summarized in Sec. 2.5. In Sec. 2.6, concepts for global optimization of polynomial optimization problems (POPs) and software for solving small-size POPs are introduced.

2.1 Two-Channel Orthogonal Filter Banks

Two-channel conjugate quadrature (CQ) filter banks, also known as power-symmetric filter banks, are among the most popular building blocks for multirate systems and wavelet-based coding systems, and have wide applications in general areas of digital signal processing, especially in image/speech coding. A two-channel CQ filter bank offers perfect reconstruction (PR) and other desirable properties. The PR property ensures that the output of the filter bank be an exact reconstruction of the input

signal up to a constant delay, hence an important property in many applications.

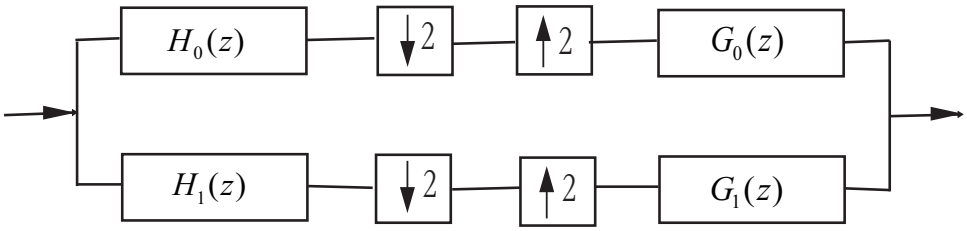

Figure 2.1: A 2-channel CQ filter bank.

A two-channel causal finite-impulser-response (FIR) CQ filter bank consists of a pair of analysis filters H_0, H_1 and a pair of synthesis filters G_0 and G_1 as shown in Fig. 2.1, where the four filters are related by [33]

$$\begin{aligned} H_1(z) &= -z^{-(N-1)} H_0(-z^{-1}) \\ G_0(z) &= H_1(-z) \\ G_1(z) &= -H_0(-z) \end{aligned} \tag{2.1}$$

where $H_0(z) = \sum_{n=0}^{N-1} h_n z^{-n}$ is a lowpass FIR transfer function of length-N with N even. With (2.1), the aliasing is eliminated and the perfect reconstruction (PR) is achieved if $H_0(z)$ satisfies

$$H_0(z)H_0(z^{-1}) + H_0(-z)H_0(-z^{-1}) = 2 \tag{2.2}$$

Eq. (2.2) is equivalent to a set of $N/2$ equality constraints as

$$\sum_{n=0}^{N-1-2m} h_n \cdot h_{n+2m} = \delta_m \quad \text{for } m = 0, 1, ..., (N-2)/2 \tag{2.3}$$

where δ_m is the Dirac sequence with $\delta_0 = 1$ and $\delta_m = 0$ for nonzero m. Eq. (2.3) is known as the double shift orthogonality in the wavelet literature. In addition to the PR condition, CQ filters may be required to meet other constraints such as possessing a certain number of vanishing moments (VMs) for constructing wavelets [32, 24, 38, 39]. The number of VMs is related to the smoothness of the wavelet function associated with the filter. Possessing a certain number of VMs helps reduce the number of nonzero wavelet coefficients in the highpass subbands, hence facilitating many signal compression tasks. It is known that the number of VMs of a CQ filter

bank is equal to the number of zeros of $H_0(z)$ at $\omega = \pi$. Because

$$\frac{d^l H_0(e^{j\omega})}{d\omega^l}\Big|_{\omega=\pi} = (-j)^l \sum_{n=0}^{N-1} (-1)^n \cdot n^l \cdot h_n$$

a CQ filter has L vanishing moments if

$$\sum_{n=0}^{N-1} (-1)^n \cdot n^l \cdot h_n = 0 \quad \text{for } l = 0, 1, ..., L-1 \tag{2.4}$$

In summary, by minimizing the stopband energy subject to the PR and VM constraints, a least squares (LS) design of CQ lowpass filter $H_0(z)$ of length-N having L VMs can be cast as

$$\text{minimize} \qquad \int_{\omega_a}^{\pi} |H_0(e^{j\omega})|^2 d\omega \tag{2.5a}$$

$$\text{subject to:} \qquad \text{constraints (2.3) and (2.4)} \tag{2.5b}$$

where ω_a is the normalized stopband edge of $H_0(z)$. By writing

$$H_0(e^{j\omega}) = \sum_{k=0}^{N-1} h_k e^{-jk\omega} = \mathbf{h}^T \mathbf{p}(\omega) \tag{2.6}$$

with $\mathbf{h} = [h_0 \ h_1 \ ... \ h_{N-1}]^T$ and $\mathbf{p}(\omega) = [1 \ e^{-j\omega} \ \cdots \ e^{-j(N-1)\omega}]^T$, we substitute (2.6) into the stopband energy in (2.5a), and the stopband energy becomes $\mathbf{h}^T\mathbf{Q}\mathbf{h}$ where $\mathbf{Q}$ is a symmetric positive-definite Toeplitz matrix determined by its first row $[\pi - \omega_a, -\sin\omega_a, -\sin 2\omega_a/2, \cdots, -\sin(N-1)\omega_a/(N-1)]$. Thus, the design formulation in (2.5) can be expressed as

$$\text{minimize} \qquad \mathbf{h}^T\mathbf{Q}\mathbf{h} \tag{2.7a}$$

$$\text{subject to:} \qquad \sum_{n=0}^{N-1-2m} h_n \cdot h_{n+2m} = \delta_m \tag{2.7b}$$

$$\sum_{n=0}^{N-1} (-1)^n \cdot n^l \cdot h_n = 0 \tag{2.7c}$$

where $m = 0, \ 1, \ ..., \ (N-2)/2$, $l = 0, \ 1, \ ..., \ L-1$ and δ_m is the Dirac sequence with $\delta_0 = 1$ and $\delta_m = 0$ for nonzero m.

In this thesis, we also consider the minimization of the maximum instantaneous power of a lowpass filter $H_0(z)$ of length-N over its stopband subject to PR and VM constraints. Thus the minimax design can be formulated as

$$\begin{aligned} \text{minimize} \quad & \underset{\omega_a \le \omega \le \pi}{\text{maximize}} \ |H_0(e^{j\omega})| && \text{(2.8a)} \\ \text{subject to:} \quad & \text{constraints (2.3) and (2.4)} && \text{(2.8b)} \end{aligned}$$

where ω_a is the normalized stopband edge of $H_0(z)$. By defining $\mathbf{c}(\omega) = [1 \ \cos\omega \ ... \ \cos(N-1)\omega]^T$ and $\mathbf{s}(\omega) = [0 \ \sin\omega \ ... \ \sin(N-1)\omega]^T$, we can write

$$\begin{aligned} |H_0(e^{j\omega})| &= \sqrt{[\mathbf{h}^T\mathbf{c}(\omega)]^2 + [\mathbf{h}^T\mathbf{s}(\omega)]^2} = \left\| \begin{bmatrix} \mathbf{c}(\omega)^T \\ \mathbf{s}(\omega)^T \end{bmatrix} \cdot \mathbf{h} \right\| \\ &\equiv ||\mathbf{T}(\omega) \cdot \mathbf{h}|| \end{aligned}$$

Then by introducing an upper bound η (as an auxiliary variable) for the objective function in (2.8) over frequency grids $\Omega = \{\omega_a = \omega_1, \ \omega_2, \ ..., \ \omega_K = \pi\}$ in the stopband, problem (2.8) can be formulated as

$$\begin{aligned} \text{minimize} \quad & \eta && \text{(2.9a)} \\ \text{subject to:} \quad & \|\mathbf{T}(\omega) \cdot \mathbf{h}\| \le \eta \ \text{ for } \omega \in \Omega && \text{(2.9b)} \\ & \sum_{n=0}^{N-1-2m} h_n \cdot h_{n+2m} = \delta_m && \text{(2.9c)} \\ & \sum_{n=0}^{N-1} (-1)^n \cdot n^l \cdot h_n = 0 && \text{(2.9d)} \end{aligned}$$

where $m = 0, \ 1, \ ..., \ (N-2)/2$ and $l = 0, \ 1, \ ..., \ L-1$.

2.2 Orthogonal Cosine-Modulated Filter Banks

An M-channel, maximally decimated orthogonal cosine-modulated (OCM) filter bank is illustrated in Fig. 2.2. The coefficients of the analysis and synthesis filters are

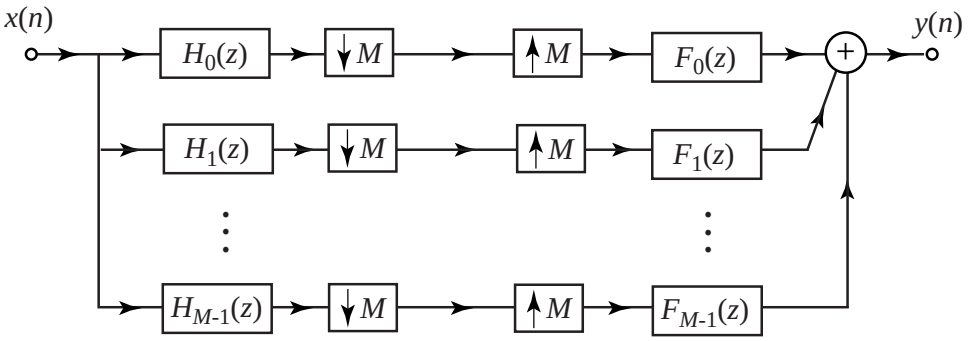

Figure 2.2: An M-channel maximally decimated filter bank.

respectively given by

$$h_k(n) = 2h(n)\cos\left[\frac{\pi}{M}\left(k+\frac{1}{2}\right)\left(n-\frac{D}{2}\right)+(-1)^k\frac{\pi}{4}\right] \tag{2.10a}$$

$$f_k(n) = 2h(n)\cos\left[\frac{\pi}{M}\left(k+\frac{1}{2}\right)\left(n-\frac{D}{2}\right)-(-1)^k\frac{\pi}{4}\right] \tag{2.10b}$$

for $0 \le k \le M-1$ and $0 \le n \le N-1$, where $\{h(n)\}$ is the impulse response of the FIR prototype filter (PF) and D is the system delay. OCM filter banks are one of the most useful classes of multirate systems as they offer the PR property. The PF of an OCM filter bank has a linear-phase response, which is crucial to avoiding phase distortion. In addition, as can be observed from Eq. (2.10), an M-channel OCM filter bank is uniquely characterized by its PF, hence optimal synthesis of an OCM multirate system can be carried out with considerably reduced complexity compared with that of a general M-channel system [17].

There exist discrete cosine transform (DCT) modulations for OCM filter banks with structures other than that of (2.10) [10]. In this thesis, we concentrate on the DCT-IV OCM filter bank as defined in (2.10) and assume that

1) the channel number M is even;

2) the filter length N assumes the form $N = 2mM$ for some positive integer m;

3) the system delay is $D = N-1$ since the PF of an OCM filter bank has a linear phase response.

For the rationale of these assumptions, see [10, 11, 35].

The input-output relation of the system in Fig. 2.2 in the z-domain is given by

$$Y(z) = T_0(z)X(z) + \sum_{l=1}^{M-1} T_l(z)X(ze^{-j2\pi l/M}) \tag{2.11}$$

where

$$T_0(z) = \frac{1}{M}\sum_{k=0}^{M-1} F_k(z)H_k(z)$$

is the distortion transfer function determining the distortion caused by the system for the unaliased component $X(z)$, and

$$T_l(z) = \frac{1}{M}\sum_{k=0}^{M-1} F_k(z)H_k(ze^{-j2\pi l/M}) \quad \text{for } l = 1, 2, \cdots, M-1$$

are the alias transfer functions that determine how the aliased components $X(ze^{-j2\pi l/M})$ are attenuated. The OCM filter bank holds the PR property if and only if

$$T_0(z) = z^{-D}, \qquad T_l(z) = 0 \quad \text{for } l = 1, 2, \cdots, M-1 \tag{2.12}$$

Under this circumstance, (2.11) becomes $Y(z) = z^{-D}X(z)$ and in the time domain the output is a delayed replica of the input as $y(n) = x(n-D)$.

Typically, the "closeness" of an OCM filter bank to the PR property is measured in the frequency domain by means of

1) amplitude distortion

$$e_m(\omega) = 1 - |T_0(e^{j\omega})|, \quad \text{for } \omega \in [0, \pi] \tag{2.13}$$

2) group-delay distortion

$$e_{gd}(\omega) = D - \arg[T_0(e^{j\omega})], \quad \text{for } \omega \in [0, \pi] \tag{2.14}$$

3) worst case aliasing error

$$e_a(\omega) = \max_{1 \le l \le M-1} |T_l(e^{j\omega})|, \quad \text{for } \omega \in [0, \pi] \tag{2.15}$$

For an OCM filter bank which satisfies the PR condition, its $e_m(\omega)$, $e_{gd}(\omega)$ and $e_a(\omega)$ should be uniformly small in magnitude for $\omega \in [0, \pi]$.

Alternatively, the PR condition can be described in the time domain by the quadratic equations [10]

$$a_{l,n}(\mathbf{h}) = \mathbf{h}^T \mathbf{Q}_{l,n} \mathbf{h} - c_n = 0 \tag{2.16a}$$
$$\text{for } 0 \le n \le m-1 \text{ and } 0 \le l \le M/2 - 1$$

where $\mathbf{h} = [h_0 \; h_1 \; \cdots \; h_{N-1}]^T$ denotes the coefficients of the PF, and

$$\mathbf{Q}_{l,n} = \mathbf{V}_{2M-1-l} \mathbf{D}_n \mathbf{V}_l^T + \mathbf{V}_{M-1-l} \mathbf{D}_n \mathbf{V}_{M+l}^T \tag{2.16b}$$

$$\mathbf{D}_n(i,j) = \begin{cases} 1, & \text{if } i+j = n \\ 0, & \text{otherwise} \end{cases} \tag{2.16c}$$

$$\mathbf{V}_l(i,j) = \begin{cases} 1, & \text{if } i = l + 2jM \\ 0, & \text{otherwise} \end{cases} \tag{2.16d}$$

$$c_n = \frac{1}{2M} \delta(n-s) \tag{2.16e}$$

for $i = 0, 1, \cdots, N-1$ and $j = 0, 1, \cdots, N-1$.

In this thesis, we consider designing a PR OCM filter bank with its PF's stopband energy

$$e_2(\mathbf{h}) = \int_{\omega_s}^{\pi} |H(e^{j\omega})|^2 d\omega$$

minimized, where $\omega_s = (1+\rho)\pi/2M$ is the stopband edge of the PF with $\rho > 0$ (ρ is always assumed to be 1 in our designs). As in Sec. 2.1 (see Eq. (2.6)), the stopband energy is given by

$$e_2(\mathbf{h}) = \mathbf{h}^T \mathbf{P} \mathbf{h} \tag{2.17}$$

where $\mathbf{P}$ is a symmetric positive-definite Toeplitz matrix determined by its first row $[\pi - \omega_s, -\sin\omega_s, -\sin 2\omega_s/2, \; \cdots, -\sin(N-1)\omega_s/(N-1)]$. With Eqs. (2.17) and (2.16a), the design of the PF of an OCM filter bank can be expressed as

$$\text{minimize} \quad e_2(\mathbf{h}) = \mathbf{h}^T \mathbf{P} \mathbf{h} \tag{2.18a}$$
$$\text{subject to:} \quad a_{l,n}(\mathbf{h}) = \mathbf{h}^T \mathbf{Q}_{l,n} \mathbf{h} - c_n = 0 \tag{2.18b}$$
$$\text{for } 0 \le n \le m-1 \text{ and } 0 \le l \le M/2 - 1$$

Since the impulse response of the PF for an OCM filter bank is symmetrical such

that the PF has linear phase, the design variables are reduced to only the components in the first half of the PF's impulse response, i.e., $\hat{\mathbf{h}} = [h_0 \ h_1 \ \cdots \ h_{N/2-1}]^T$. In consequence, matrices $\mathbf{P}$ and $\mathbf{Q}_{l,n}$ of size $N \times N$ in (2.18) need to be reduced to matrices $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}_{l,n}$ of size $N/2 \times N/2$. The method to generate the reduced-size matrices is elaborated in the Appendix. Thus, problem (2.18) can be further cast as

$$\begin{aligned} \text{minimize} \quad & e_2(\hat{\mathbf{h}}) = \hat{\mathbf{h}}^T \hat{\mathbf{P}} \hat{\mathbf{h}} && \text{(2.19a)} \\ \text{subject to:} \quad & a_{l,n}(\hat{\mathbf{h}}) = \hat{\mathbf{h}}^T \hat{\mathbf{Q}}_{l,n} \hat{\mathbf{h}} - c_n = 0 && \text{(2.19b)} \\ & \text{for } 0 \le n \le m-1 \text{ and } 0 \le l \le M/2-1 \end{aligned}$$

By solving the above optimization problem, the impulse response $\mathbf{h}$ of the PF can then be obtained as

$$\mathbf{h} = \begin{bmatrix} \hat{\mathbf{h}} \\ \texttt{flipud}(\hat{\mathbf{h}}) \end{bmatrix} \tag{2.20}$$

where $\texttt{flipud}(\hat{\mathbf{h}})$ denotes a vector generated by flipping $\hat{\mathbf{h}}$ upside down.

2.3 Gauss-Newton Method with Adaptively Controlled Weights

In this section, we describe how to solve a system of nonlinear equations using gradient-based optimization method, known as the Gauss-Newton (G-N) method [1]. This method is found useful many places in this thesis.

To begin with, consider a system of nonlinear equations that assumes the following form

$$f_p(\mathbf{x}) = 0 \quad \text{for } p = 1, 2, ..., m \tag{2.21}$$

where $\mathbf{x} = [x_1 \ x_2 \cdots x_n]^T$. In order to find a solution vector $\mathbf{x}$ of (2.21), an unconstrained optimization problem can be formulated as

$$\text{minimize} \quad F(\mathbf{x}) = \sum_{p=1}^{m} f_p(\mathbf{x})^2 \tag{2.22}$$

In this way, the system of nonlinear equations in (2.21) is solved in the least-squares sense.

A problem with the G-N method is that it cannot be utilized to numerically com-

pute a solution satisfying the system of equations (2.21) with equal equation errors. To seek such a solution, the G-N method with adaptively controlled weights can be applied. By implementing this method, a weighted sum-of-squares type objective function is minimized using the G-N algorithm with the weights adjusted according to the absolute equation errors evaluated at the current iterate. In what follows, we present some details for implementing this technique.

As a first step, we define a vector

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \sqrt{w_1} f_1(\mathbf{x}) & \sqrt{w_2} f_2(\mathbf{x}) & \cdots & \sqrt{w_m} f_m(\mathbf{x}) \end{bmatrix}^T$$

where w_p $(1 \leq p \leq m)$ are weight values which need to be adjusted in every round of iteration. As will become transparent shortly, the value assignment for w_p is critical in obtaining the solution that will make the system of equations (2.21) equally satisfied. Accordingly, the problem in (2.22) is modified to

$$\text{minimize} \qquad F(\mathbf{x}) = \mathbf{f}^T(\mathbf{x})\mathbf{f}(\mathbf{x}) = \sum_{p=1}^{m} w_p f_p(\mathbf{x})^2 \tag{2.23}$$

Problem (2.23) is an unconstrained optimization problem for which many gradient-based methods exist. A second-order method known as the Newton method is efficient in dealing with such problem. To solve (2.23) using the Newton method, we define the Jacobian matrix for the system of equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ as

$$\mathbf{J} = \begin{pmatrix} \sqrt{w_1}\frac{\partial f_1}{\partial x_1} & \sqrt{w_1}\frac{\partial f_1}{\partial x_2} & \cdots & \sqrt{w_1}\frac{\partial f_1}{\partial x_n} \\ \sqrt{w_2}\frac{\partial f_2}{\partial x_1} & \sqrt{w_2}\frac{\partial f_2}{\partial x_2} & \cdots & \sqrt{w_2}\frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \sqrt{w_m}\frac{\partial f_m}{\partial x_1} & \sqrt{w_m}\frac{\partial f_m}{\partial x_2} & \cdots & \sqrt{w_m}\frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Then the gradient $\mathbf{g}$ of $F(\mathbf{x})$ in (2.23) can be calculated as

$$\mathbf{g} = 2\mathbf{J}^T\mathbf{f}$$

The Hessian $\mathbf{H}$ of $F(\mathbf{x})$ can be deduced by neglecting the second-derivatives of $f_p(\mathbf{x})$ as

$$\mathbf{H} \approx 2\mathbf{J}^T\mathbf{J} \tag{2.24}$$

If $\mathbf{H}$ is not positive definite, it is forced to be so by letting

$$\mathbf{H} := \frac{\mathbf{H} + \beta \mathbf{I}_n}{1 + \beta}$$

with $\beta > 0$ set to a sufficiently large value and $\mathbf{I}_n$ an $n \times n$ identity matrix. Using the Newton direction, the next iterate $\mathbf{x}_{k+1}$ is calculated as

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{H}^{-1} \mathbf{g}$$

where α_k is the value of α that minimizes $F(\mathbf{x}_k - \alpha \mathbf{H}^{-1}\mathbf{g})$ and can be determined by inexact line search [1]. For the current set of w_p $(1 \leq p \leq m)$, the G-N algorithm converges to $\mathbf{x}_w = \mathbf{x}_{k+1}$ when $\|\mathbf{x}_{k+1} - \mathbf{x}_k\|$ is less than a pre-specified tolerance ε.

In order to identify the solution which makes the errors of the m individual equations in (2.21) practically equal, weights w_p should be adjusted based on equation errors at the most recent solution point $\mathbf{x}_w$. Specifically, we denote the solution of the jth round G-N optimization by $\mathbf{x}_w^{(j)}$ and the current weight values by $w_1^{(j)},\ w_2^{(j)},\ \ldots,\ w_m^{(j)}$. In preparation for the next round of G-N optimization, the weights are adjusted according to

$$w_p^{(j+1)} = w_p^{(j)} + \mu[|f_p(\mathbf{x}_w^{(j)})| - f_{\max}(\mathbf{x}_w^{(j)})] \quad \text{for } p = 1, 2, ..., m$$

where $\mu > 0$ is a parameter controlling the magnitude of the adjustments, and

$$f_{\max}(\mathbf{x}_w^{(j)}) = \text{maximum}\{|f_p(\mathbf{x}_w^{(j)})|\} \quad \text{for } p = 1, 2, ..., m$$

Then the new weights $w_p^{(j+1)}$ are normalized to satisfy $\sum_{p=1}^{m} w_p^{(j+1)} = 1$. A new solution to problem (2.23) is obtained with the new weight values by the G-N algorithm. This procedure is repeated until the weights are stabilized and converge to constant values after a prescribed K times of G-N optimizations. In this way, a solution satisfying the m constraints in (2.21) with practically identical errors is obtained. In the simulations to be presented in this thesis, the initial values $w_p^{(0)}$ for $p = 1,\ 2,\ ...,\ m$ were set to $1/m$, μ was set to 10, and K was set to 20.

2.4 Convex Quadratic Programming and Second-Order Cone Programming

In this section, we briefly review two important classes of convex programming problems, namely convex quadratic programming problems and second-order cone programming problems, which will be used in subsequent chapters of the thesis.

Quadratic programming (QP) is a family of methods, techniques, and algorithms that can be used to minimize quadratic objective functions subject to linear constraints [1]. QP is often used as the basis in constrained nonlinear programming. When the objective function of a QP problem is a convex quadratic function, the problem is called a convex QP problem.

The general form of a QP problem is to minimize a quadratic function subject to a set of linear equality and a set of linear inequality constraints. The linear equality constraints can be eliminated using the method of singular value decomposition (SVD) or QR decomposition [1]. Without loss of generality, a QP problem can be expressed as

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x} + \mathbf{x}^T\mathbf{p} && \text{(2.25a)} \\ \text{subject to:} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} && \text{(2.25b)} \end{aligned}$$

where $\mathbf{H} \in \mathcal{R}^{N\times N}$, $\mathbf{p} \in \mathcal{R}^{N\times 1}$, $\mathbf{A} \in \mathcal{R}^{P\times N}$ and $\mathbf{b} \in \mathcal{R}^{P\times 1}$. The Hessian $\mathbf{H}$ of the objective function in (2.25a) is symmetric and positive semidefinite for a convex QP problem. Many efficient algorithms and reliable software like SeDuMi [51] exist for problem (2.25). The MATLAB optimization toolbox [50] is an efficient solver for convex QP, and the MATLAB command to solve (2.25) with $\mathbf{H}$ positive semidefinite is `quadprog`.

An important branch of convex programming is the second-order cone programming (SOCP) where a linear function is minimized subject to a set of second-order cone constraints and possibly a set of linear inequality constraints

$$\begin{aligned} \text{minimize} \quad & \mathbf{b}^T\mathbf{x} && \text{(2.26a)} \\ \text{subject to:} \quad & \|\mathbf{A}_i^T\mathbf{x} + \mathbf{c}_i\| \leq \mathbf{b}_i^T\mathbf{x} + d_i \ \text{ for } i = 1, 2, ..., q && \text{(2.26b)} \\ & \mathbf{F}\mathbf{x} \geq \mathbf{g} && \text{(2.26c)} \end{aligned}$$

where $\mathbf{b} \in \mathcal{R}^{N\times 1}$, $\mathbf{A}_i \in \mathcal{R}^{N\times(N_i-1)}$, $\mathbf{c}_i \in \mathcal{R}^{(N_i-1)\times 1}$ and $\mathbf{b}_i \in \mathcal{R}^{N\times 1}$. The term "second-

order cone" reflects the fact that each constraint in (2.26b) is equivalent to a conic constraint

$$\begin{bmatrix} \mathbf{b}_i^T \\ \mathbf{A}_i^T \end{bmatrix} \mathbf{x} + \begin{bmatrix} d_i \\ \mathbf{c}_i \end{bmatrix} \in \mathcal{C}_i$$

where $\mathcal{C}_i$ is the second-order cone in $\mathcal{R}^{N_i}$

$$\mathcal{C}_i = \left\{ \begin{bmatrix} t \\ \mathbf{u} \end{bmatrix} : \mathbf{u} \in \mathcal{R}^{(N_i-1)\times 1}, t \geq 0, \|\mathbf{u}\| \leq t \right\}$$

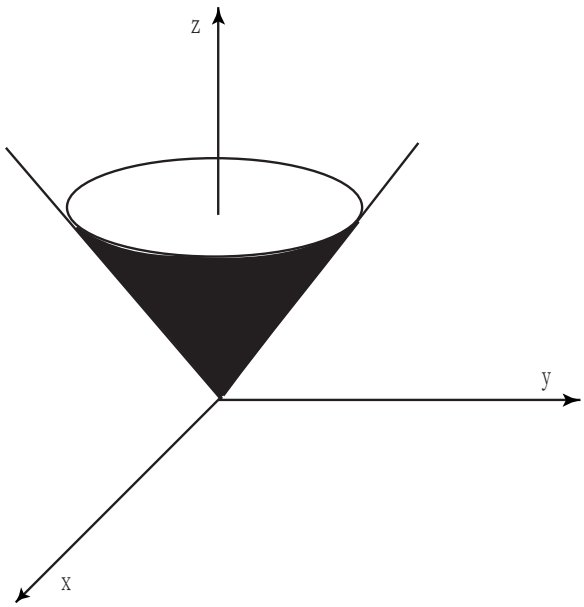

Figure 2.3: Second-order cone in $\mathcal{R}^3$.

The second-order cone in space $\mathcal{R}^3$ is illustrated in Fig. 2.3. It is evident that SOCP includes linear programming and convex QP as special cases. It is also known that SOCP is a subclass of semidefinite programming (SDP) [25, 4, 42] because each constraint in (2.26b) can be expressed as

$$\begin{bmatrix} (\mathbf{b}_i^T\mathbf{x} + d_i)\mathbf{I} & \mathbf{A}_i^T\mathbf{x} + \mathbf{c}_i \\ (\mathbf{A_i}^T\mathbf{x} + \mathbf{c}_i)^T & \mathbf{b}_i^T\mathbf{x} + d_i \end{bmatrix} \succeq \mathbf{0}$$

SOCP problems can be solved by various polynomial-time interior-point optimization algorithms [1, 26] and the MATLAB toolbox SeDuMi [51] is found efficient in solving SOCP problems. It is also important to remark that it is often more efficient to solve problem (2.26) using an SOCP solver rather than treating it as an equivalent

SDP problem.

2.5 General Nonlinear Optimization Problems

The most general class of optimization problems is the class of problems where both the objective function and the constraints are nonlinear and possibly nonconvex [1]. Consider the optimization problem

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{h}) && \text{(2.27a)} \\ \text{subject to:} \quad & a_i(\mathbf{h}) = 0 \ \text{ for } i = 1, 2, ..., p && \text{(2.27b)} \end{aligned}$$

where $f(\mathbf{h})$ and $a_i(\mathbf{h})$ are continuous functions with continuous second partial derivatives. We assume that (2.27b) defines a nonempty feasible region $\mathcal{R}$. Among a variety of methods in dealing with problems like (2.27), the sequential quadratic-programming (SQP) algorithms have proved highly effective. In this section, we briefly review a SQP method for problem (2.27).

The first-order necessary conditions for $\mathbf{h}^*$ to be a local minimizer of (2.27) are that there exists a $\boldsymbol{\lambda}^* \in R^p$ such that

$$\nabla_h \mathcal{L}(\mathbf{h}^*, \boldsymbol{\lambda}^*) = \mathbf{0}, \ \ \nabla_\lambda \mathcal{L}(\mathbf{h}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$$

where the Lagrangian $\mathcal{L}(\mathbf{h}, \boldsymbol{\lambda})$ is defined by

$$\mathcal{L}(\mathbf{h}, \boldsymbol{\lambda}) = f(\mathbf{h}) - \sum_{i=1}^{p} \lambda_i a_i(\mathbf{h})$$

Suppose we are in the kth iteration with iterate $\{\mathbf{h}_k, \boldsymbol{\lambda}_k\}$, which is assumed to be sufficiently close to $\{\mathbf{h}^*, \boldsymbol{\lambda}^*\}$. To find an increment $\{\boldsymbol{\delta}_h, \boldsymbol{\delta}_\lambda\}$ such that $\{\mathbf{h}_{k+1}, \boldsymbol{\lambda}_{k+1}\} = \{\mathbf{h}_k + \boldsymbol{\delta}_h, \boldsymbol{\lambda}_k + \boldsymbol{\delta}_\lambda\}$ is closer to $\{\mathbf{h}^*, \boldsymbol{\lambda}^*\}$, we apply the first-order Taylor approximation of $\boldsymbol{\nabla}\mathcal{L}$[1] for $\{\mathbf{h}_k, \boldsymbol{\lambda}_k\}$, i.e.,

$$\boldsymbol{\nabla}\mathcal{L}(\mathbf{h}_{k+1}, \boldsymbol{\lambda}_{k+1}) \approx \boldsymbol{\nabla}\mathcal{L}(\mathbf{h}_k, \boldsymbol{\lambda}_k) + \boldsymbol{\nabla}^2\mathcal{L}(\mathbf{h}_k, \boldsymbol{\lambda}_k) \begin{bmatrix} \boldsymbol{\delta}_h \\ \boldsymbol{\delta}_\lambda \end{bmatrix}$$

[1]The symbol $\boldsymbol{\nabla}$ is defined as $\boldsymbol{\nabla} = \begin{bmatrix} \nabla_h^T & \nabla_\lambda^T \end{bmatrix}^T$.

We see that $\{\boldsymbol{h}_{k+1},\, \boldsymbol{\lambda}_{k+1}\}$ is an approximation of $\{\boldsymbol{h}^*,\, \boldsymbol{\lambda}^*\}$ if

$$\boldsymbol{\nabla}^2\mathcal{L}(\mathbf{h}_k, \boldsymbol{\lambda}_k)\begin{bmatrix} \boldsymbol{\delta}_h \\ \boldsymbol{\delta}_\lambda \end{bmatrix} = -\boldsymbol{\nabla}\mathcal{L}(\mathbf{h}_k, \boldsymbol{\lambda}_k)$$

which can be reformulated in terms of the Hessian $\mathbf{W}_k$ of the Lagrangian $\mathcal{L}(\boldsymbol{h}_k,\, \boldsymbol{\lambda}_k)$ and the Jacobian $\mathbf{A}_k$ of the equality constraints (2.27b) as

$$\begin{bmatrix} \mathbf{W}_k & -\mathbf{A}_k^T \\ -\mathbf{A}_k & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_h \\ \boldsymbol{\delta}_\lambda \end{bmatrix} = \begin{bmatrix} \mathbf{A}_k^T\boldsymbol{\lambda}_k - \mathbf{g}_k \\ \mathbf{a}_k \end{bmatrix} \tag{2.28}$$

where

$$\mathbf{W}_k = \nabla_h^2 f(\mathbf{h}_k) - \sum_{i=1}^{p}(\boldsymbol{\lambda}_k)_i \nabla_h^2 a_i(\mathbf{h}_k) \tag{2.29a}$$

$$\mathbf{A}_k = \begin{bmatrix} \nabla_h a_1(\mathbf{h}_k) & \nabla_h a_2(\mathbf{h}_k) & \cdots & \nabla_h a_p(\mathbf{h}_k) \end{bmatrix}^T \tag{2.29b}$$

$$\mathbf{g}_k = \nabla_h f(\mathbf{h}_k) \tag{2.29c}$$

$$\mathbf{a}_k = \begin{bmatrix} a_1(\mathbf{h}_k) & a_2(\mathbf{h}_k) & \cdots & a_p(\mathbf{h}_k) \end{bmatrix}^T \tag{2.29d}$$

Eq. (2.28) can also be expressed as

$$\mathbf{W}_k\boldsymbol{\delta}_h + \mathbf{g}_k = \mathbf{A}_k^T\boldsymbol{\lambda}_{k+1} \tag{2.30a}$$

$$\mathbf{A}_k\boldsymbol{\delta}_h = -\mathbf{a}_k \tag{2.30b}$$

and they can be interpreted as the first-order necessary conditions for $\boldsymbol{\delta}_h$ to be a local minimizer of the problem

$$\text{minimize} \quad \frac{1}{2}\boldsymbol{\delta}_h^T\mathbf{W}_k\boldsymbol{\delta}_h + \boldsymbol{\delta}_h^T\mathbf{g}_k \tag{2.31a}$$

$$\text{subject to:} \quad \mathbf{A}_k\boldsymbol{\delta}_h = -\mathbf{a}_k \tag{2.31b}$$

$$||\boldsymbol{\delta}_h|| \text{ is small} \tag{2.31c}$$

Note that constraint (2.31c) is added to ensure the validity of the Taylor approximation. By removing the equality constraint (2.31b) using the singular value decomposition (SVD) or QR decomposition [1], problem (2.31) assumes the form of problem (2.25), which is a QP problem (see Sec. 2.4). Once the minimizer $\boldsymbol{\delta}_h^*$ is found, the

next iterate is set to

$$\mathbf{h}_{k+1} = \mathbf{h}_k + \boldsymbol{\delta}_h^*, \;\; \boldsymbol{\lambda}_{k+1} = (\mathbf{A}_k \mathbf{A}_k^T)^{-1} \mathbf{A}_k (\mathbf{W}_k \boldsymbol{\delta}_h^* + \boldsymbol{g}_k)$$

in which $\boldsymbol{\lambda}_{k+1}$ is calculated using Eq. (2.30a). With $\mathbf{h}_{k+1}$ and $\boldsymbol{\lambda}_{k+1}$ found, $\mathbf{W}_{k+1}$, $\mathbf{A}_{k+1}$, $\mathbf{g}_{k+1}$ and $\mathbf{a}_{k+1}$ can be evaluated for a new round of iteration. The algorithm continues until $||\boldsymbol{\delta}_h^*||$ becomes sufficiently small.

We see that this method consists of solving a series of QP subproblems in a sequential manner, hence the name sequential quadratic-programming (SQP).

2.6 Global Optimization of Small-Size Polynomial Optimization Problems

A real-valued polynomial $f(\mathbf{x})$ in an n-dimensional space R^n can be expressed as

$$f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{F}} c(\boldsymbol{\alpha}) \mathbf{x}^{\boldsymbol{\alpha}} \tag{2.33}$$

where $c(\boldsymbol{\alpha}) \in R$, $\mathbf{x} = [x_1 \; x_2 \; ... \; x_n]$, $\boldsymbol{\alpha} = [\alpha_1 \; \alpha_2 \; ... \; \alpha_n] \in \mathcal{F} \subset \mathcal{Z}_+^n$ — the set of all vectors in R^n whose components are nonnegative integers, and $\mathbf{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \; x_2^{\alpha_2} \; ... \; x_n^{\alpha_n}$. The order (degree) of $f(\mathbf{x})$ is defined as the largest $\sum_i \alpha_i$.

A general polynomial optimization problem (POP) has the form

$$\begin{aligned}
\text{minimize} \quad & f_0(\mathbf{x}) && \text{(2.34a)} \\
\text{subject to:} \quad & f_k(\mathbf{x}) \geq 0 \;\; \text{for } k = 1, ..., L && \text{(2.34b)} \\
& f_k(\mathbf{x}) = 0 \;\; \text{for } k = L+1, ..., K && \text{(2.34c)}
\end{aligned}$$

where $f_k(\mathbf{x})$ for $k = 0, \; 1, \; ..., \; K$ are real-valued polynomials. POPs include linear programming (LP), convex quadratic programming (QP), semidefinite programming (SDP), and second-order cone programming (SOCP) problems as its special cases. More importantly, POPs stand for a substantially broader class that covers many nonconvex optimization problems [22].

A recent breakthrough in the field is made by Lasserre [22] in which it is proved that when the feasible region of (2.34) is compact (not necessarily convex), its global solution can be approximated as closely as desired (and often can be obtained exactly)

by solving a finite sequence of SDP problems. A technical difficulty with the method of [22] is that the size of the SDP problems involved in a POP usually grows very quickly. This may cause numerical difficulties even for POPs of moderate scales.

More recently, sparse SDP relaxation [44] is proposed for global solutions of POPs of relatively larger scales with improved efficiency. The method is supported by the MATLAB toolbox SparsePOP [43][45]. Another MATLAB toolbox for POPs is GloptiPoly [9] which is intended to solve the generalized problem of moments that can be viewed as an extension of the classical problem of moments [23].

While in theory the method in [22] is able to generate global solutions for POPs, the size of an SDP problem it requires to solve grows so quickly with the size of the original POP that the current versions of the software such as GloptiPoly version 3 and SparsePOP V220 can only handle POPs of very limited size. Nevertheless, the globally optimal solutions of small-size POPs provided by the software form one of the key ingredients in the proposed techniques in this thesis towards global designs of several types of filter banks of high order. We shall illustrate this with details in the following corresponding chapters.

Chapter 3

Least Squares Design of Orthogonal Filter Banks and Wavelets

As studied in Sec. 2.1, a least squares (LS) design of a conjugate quadrature (CQ) lowpass filter $H_0(z)$ of length-N with L vanishing moments (VMs) can be cast as

$$(\mathcal{P}_2^N) \qquad \text{minimize} \qquad \mathbf{h}^T\mathbf{Q}\mathbf{h} \tag{3.1a}$$

$$\text{subject to:} \qquad \sum_{n=0}^{N-1-2m} h_n \cdot h_{n+2m} = \delta_m \tag{3.1b}$$

$$\sum_{n=0}^{N-1} (-1)^n \cdot n^l \cdot h_n = 0 \tag{3.1c}$$

where $\mathbf{h} = [h_0 \ h_1 \ ... \ h_{N-1}]^T$, $m = 0,\ 1,\ ...,\ (N-2)/2$, $l = 0,\ 1,\ ...,\ L-1$, and δ_m is the Dirac sequence with $\delta_0 = 1$ and $\delta_m = 0$ for nonzero m. The matrix $\mathbf{Q}$ in (3.1a) is a symmetric positive definite Toeplitz matrix characterized by its first row $\left[\ \pi - \omega_a \quad -\sin\omega_a \quad \cdots \quad -\sin(N-1)\omega_a/(N-1)\ \right]$ with ω_a the normalized stopband edge of $H_0(z)$.

In this chapter, several methods for local and global LS designs of CQ filter banks are investigated. In Sec. 3.1, two local methods, i.e., the sequential convex-programming method and the sequential quadratic-programming method are studied. Sec. 3.2 addresses global LS design for CQ filter banks of low order. Based on observation of a common pattern among global low-order impulse responses, a strategy of

designing potentially global LS filter banks of high order is proposed in Sec. 3.3. Finally, performance of the proposed algorithm for LS low-order and high-order designs is evaluated and illustrated in Sec. 3.4.

3.1 Local LS Design of CQ Filter Banks

The LS design of CQ filters as formulated in (3.1) is a nonconvex problem because of the presence of the quadratic equality constraints in (3.1b). As a result, it possesses multiple local solutions. Several (local) design techniques are available in the literature [16, 33, 32, 24, 38, 39, 41]. In particular, a direct method is recently proposed in [16] and is found to provide improved performance. Based on this direct method, in this section a sequential convex-programming (SCP) method and a sequential quadratic-programming (SQP) method are developed and applied for local designs, and are shown to produce improved results.

3.1.1 Sequential Convex-Programming Method

Lu and Hinamoto proposed in [16] a direct design method for LS designs that deals with problem (3.1) by local convex approximations in a sequential manner. As each update is confined within a small vicinity of the current iterate, the problem at hand behaves like a convex one. However, sometimes inaccurate solutions are obtained and unnecessary computations are involved because the optimization problem formulated in the method can become infeasible. In this section, such a problem is explained in detail and techniques for improvement in this regard are proposed to achieve more efficient designs.

Suppose we are in the kth iteration to compute $\boldsymbol{\delta}_h$ so that $\mathbf{h}_{k+1} = \mathbf{h}_k + \boldsymbol{\delta}_h$ reduces the filter's stopband energy (i.e. the value of the objective function in (3.1a)) and better satisfies the constraints in (3.1b) and (3.1c). Then $\mathbf{h}^T\mathbf{Q}\mathbf{h}$ in (3.1a) becomes

$$\mathbf{h}_{k+1}^T\mathbf{Q}\mathbf{h}_{k+1} = \boldsymbol{\delta}_h^T\mathbf{Q}\boldsymbol{\delta}_h + 2\boldsymbol{\delta}_h^T\mathbf{Q}\mathbf{h}_k + \mathbf{h}_k^T\mathbf{Q}\mathbf{h}_k \tag{3.2}$$

In the same way, (3.1c) can be written as

$$\sum_{n=0}^{N-1}(-1)^n \cdot n^l \cdot (\boldsymbol{\delta}_h)_n = -\sum_{n=0}^{N-1}(-1)^n \cdot n^l \cdot (\mathbf{h}_k)_n \tag{3.3}$$

Now if we write (3.1b) at $\mathbf{h}_{k+1}$ as

$$\sum_{n=0}^{N-1-2m} [(\mathbf{h}_k)_n + (\boldsymbol{\delta}_h)_n] \cdot [(\mathbf{h}_k)_{n+2m} + (\boldsymbol{\delta}_h)_{n+2m}] = \delta_m$$

with $\boldsymbol{\delta}_h$ limited to be small in magnitude so that the second-order term on the left-hand side of the above equation can be neglected, the above equation becomes

$$\begin{aligned} &\sum_{n=0}^{N-1-2m} (\mathbf{h}_k)_n (\boldsymbol{\delta}_h)_{n+2m} + \sum_{n=0}^{N-1-2m} (\mathbf{h}_k)_{n+2m} (\boldsymbol{\delta}_h)_n \\ &\approx \delta_m - \sum_{n=0}^{N-1-2m} (\mathbf{h}_k)_n (\mathbf{h}_k)_{n+2m} \end{aligned} \tag{3.4}$$

By using Eqs. (3.2), (3.4) and (3.3), the objective function turns into a quadratic function of $\boldsymbol{\delta}_h$ and the two constraints become linear equality constraints. With $\boldsymbol{\delta}_h$ bounded to be small, the kth iteration of problem (3.1) now assumes the form

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{\delta}_h^T \mathbf{Q} \boldsymbol{\delta}_h + \boldsymbol{\delta}_h^T \mathbf{g}_k && \text{(3.5a)} \\ \text{subject to:} \quad & \mathbf{A}_k \boldsymbol{\delta}_h = -\mathbf{a}_k && \text{(3.5b)} \\ & \mathbf{C} \boldsymbol{\delta}_h \leq \mathbf{b} && \text{(3.5c)} \end{aligned}$$

with $\mathbf{g}_k = 2\mathbf{Q}\mathbf{h}_k$, $\mathbf{C} = \begin{bmatrix} \mathbf{I}_N \\ -\mathbf{I}_N \end{bmatrix}$ where $\mathbf{I}_N$ is an N by N identity matrix, $\mathbf{b} = \beta \cdot \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$ which is of dimension $2N$ with β a small positive number, and

$$\begin{aligned} \mathbf{A}_k &= \begin{bmatrix} \mathbf{A}_{k_1} & \cdots & \mathbf{A}_{k_{\frac{N}{2}}} & \mathbf{A}_{k_{(\frac{N}{2}+1)}} & \cdots & \mathbf{A}_{k_{(\frac{N}{2}+L)}} \end{bmatrix}^T \\ \mathbf{a}_k &= \begin{bmatrix} a_{k_1} & \cdots & a_{k_{\frac{N}{2}}} & a_{k_{(\frac{N}{2}+1)}} & \cdots & a_{k_{(\frac{N}{2}+L)}} \end{bmatrix}^T \end{aligned}$$

where $\mathbf{A}_{k_i}$ and a_{k_i} are specified below.
(a) For $1 \leq i \leq \frac{N}{2}$, the N by 1 vector $\mathbf{A}_{k_i}$ has its jth element $(\mathbf{A}_{k_i})_j$ defined as

$$(\mathbf{A}_{k_i})_j = h_{j-2i+1} + h_{j+2i-3}, \;\; 1 \leq j \leq N$$

Note here that h_p is the $(p+1)$-th element of $\mathbf{h}_k$ and is defined as 0 if $p < 0$ or

$p > N-1$. For instances when $i = 1$ and $i = \frac{N}{2}$, we have

$$\mathbf{A}_{k_1} = \begin{bmatrix} 2h_0 & 2h_1 & \cdots & 2h_{N-1} \end{bmatrix}^T = 2\mathbf{h}_k$$

$$\mathbf{A}_{k_{\frac{N}{2}}} = \begin{bmatrix} h_{N-2} & h_{N-1} & 0 & \cdots & 0 & h_0 & h_1 \end{bmatrix}^T$$

The scalar a_{k_i} for $1 \leq i \leq \frac{N}{2}$ assumes the form

$$a_{k_i} = \sum_{n=0}^{N-2i+1} h_n \cdot h_{n+2i-2} - \delta_{i-1}$$

(b) For $\frac{N}{2}+1 \leq i \leq \frac{N}{2}+L$, vector $\mathbf{A}_{k_i}$ assumes the form

$$\mathbf{A}_{k_i} = \begin{bmatrix} 0^{l(i)} & -1^{l(i)} & 2^{l(i)} & \cdots & -(N-1)^{l(i)} \end{bmatrix}^T$$

where $l(i) = i - \frac{N}{2} - 1$. And a_{k_i} is simply calculated as

$$a_{k_i} = \mathbf{A}_{k_i}^T \mathbf{h}_k$$

If the VM requirement in (3.1c) is not a part of the design specifications, then we set $L = 0$. In this case, $\mathbf{A}_k$ only has the first $N/2$ row vectors and $\mathbf{a}_k$ only has the first $N/2$ elements.

We note that the equality constraint in (3.5*b*) can be eliminated via the singular value decomposition (SVD) of $\mathbf{A}_k = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^T$ as [1]

$$\boldsymbol{\delta}_h = \mathbf{V}_{k_r}\mathbf{x} + \mathbf{A}_k^{\dagger}(-\mathbf{a}_k) \tag{3.8}$$

where $\mathbf{x}$ is an $(N-r)$-dimensional free vector with r being the rank of $\mathbf{A}_k$ so $r = N/2 + L$ or $N/2$ depending on whether or not the VM requirement is included. The symbol † denotes the pseudo-inverse of a matrix, and $\mathbf{V}_{k_r}$ is the matrix consisting of the last $N-r$ columns of $\mathbf{V}_k$. With (3.8), problem (3.5) is reduced to

$$\text{minimize} \quad \mathbf{x}^T \hat{\mathbf{Q}} \mathbf{x} + \mathbf{x}^T \hat{\mathbf{g}}_k \tag{3.9a}$$

$$\text{subject to:} \quad \hat{\mathbf{C}}\mathbf{x} \leq \hat{\mathbf{b}} \tag{3.9b}$$

where

$$\hat{\mathbf{Q}} = \mathbf{V}_{k_r}^T \mathbf{Q} \mathbf{V}_{k_r}, \;\; \hat{\mathbf{g}_k} = \mathbf{V}_{k_r}^T [\mathbf{g}_k + 2\mathbf{Q}\mathbf{A}_k^\dagger(-\mathbf{a}_k)]$$

$$\hat{\mathbf{C}} = \mathbf{C}\mathbf{V}_{k_r}, \;\; \hat{\mathbf{b}} = \mathbf{b} - \mathbf{C}\mathbf{A}_k^\dagger(-\mathbf{a}_k)$$

Since $\hat{\mathbf{Q}}$ is positive definite, problem (3.9) is a convex QP problem (see Sec. 2.4).

The design steps described above form a basis of the direct design technique in [16]. In practice, however, choosing an appropriate β for the above method is fairly challenging. If β is too small, the solution space of (3.5b) will not intersect with that of (3.5c), hence there is no $\boldsymbol{\delta}_h$ satisfying both constraints. In consequence, there is no $\mathbf{x}$ satisfying (3.9b), making problem (3.9) infeasible. On the other hand, when β is made too large, the convex approximation in a local region for the original nonconvex problem may become invalid. To deal with this problem, we modify constraint (3.9b) so that problem (3.9) becomes

$$\begin{aligned} \text{minimize} \quad & \mathbf{x}^T \hat{\mathbf{Q}} \mathbf{x} + \mathbf{x}^T \hat{\mathbf{g}_k} && \text{(3.10a)} \\ \text{subject to:} \quad & \mathbf{F}\mathbf{x} \le \mathbf{a} && \text{(3.10b)} \end{aligned}$$

where $\mathbf{F} = \begin{bmatrix} \mathbf{I}_M \\ -\mathbf{I}_M \end{bmatrix}$ and $\mathbf{a} = \alpha \cdot \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$ which is of dimension $2M$ with $\alpha > 0$ a small scalar. The matrix $\mathbf{I}_M$ is an M by M identity matrix with $M = N/2 - L$ or $N/2$ depending on whether or not the VM requirement is included. Evidently, (3.10b) requires that each component of $\mathbf{x}$ is less than α in magnitude, which implies that $\|\mathbf{x}\|$ is small. This modification can be justified as follows. With $\mathbf{x}$ small, $\boldsymbol{\delta}_h$ is also limited within a small region since (3.8) implies

$$\|\boldsymbol{\delta}_h\| \le \|\mathbf{x}\| + \|\mathbf{A}_k^\dagger(-\mathbf{a}_k)\|$$

where $\mathbf{A}_k^\dagger(-\mathbf{a}_k)$ is known as the minimum-norm solution of (3.5b), therefore a small $\|\mathbf{x}\|$ implies a small $\|\boldsymbol{\delta}_h\|$. An advantage of adopting problem (3.10) is that it is always feasible and a global optimum can be obtained no matter how small α is (as far as our simulations are concerned, $\alpha = 0.01$ was found to work well).

As will be shown in Sec. 3.1.3, with this modification, the LS design of a CQ lowpass filter converges to a local minimizer satisfying constraints (3.1b) and (3.1c) more accurately than the direct method in [16]. In the rest of the chapter, we call

this modified direct design technique as the SCP method.

3.1.2 Sequential Quadratic-Programming Method

The SQP method for solving general constrained problems has been introduced in Sec. 2.5. In this section, we apply the SQP method to the LS design of CQ filters.

The LS problem (3.1) can be formulated in the form of (2.31) by virtue of the SQP method. By using Eqs. (2.29b), (2.29c) and (2.29d), $\mathbf{A}_k$, $\mathbf{g}_k$ and $\mathbf{a}_k$ in problem (2.31) can be calculated for the LS problem (3.1) and they are found to be the same as those in problem (3.5) of the SCP method. According to Eq. (2.29a), the matrix $\mathbf{W}_k$ for problem (3.1) assumes the form

$$\mathbf{W}_k = 2\mathbf{Q} - \mathbf{P}$$

where $\mathbf{P}$ is an N by N symmetric Toeplitz matrix characterized by its first row $\begin{bmatrix} 2(\boldsymbol{\lambda}_k)_1 & 0 & (\boldsymbol{\lambda}_k)_2 & 0 & \cdots & (\boldsymbol{\lambda}_k)_{\frac{N}{2}} & 0 \end{bmatrix}$. We note that matrix $\mathbf{W}_k$ is not assured to be positive semidefinite.

We also note that $\boldsymbol{\delta}_h$ in problem (2.31) cannot be limited to as small as we want because constraints (2.31b) and (2.31c) may become overly stringent, making the problem infeasible. To deal with this technical difficulty, we use the same technique introduced in the SCP method, i.e., eliminating equality constraints with SVD and bounding the new variables from above. In this way, problem (2.31) is reformulated as

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\mathbf{x}^T\tilde{\mathbf{W}}_k\mathbf{x} + \mathbf{x}^T\tilde{\mathbf{g}_k} && \text{(3.11a)} \\ \text{subject to:} \quad & \mathbf{F}\mathbf{x} \leq \mathbf{a} && \text{(3.11b)} \end{aligned}$$

where

$$\tilde{\mathbf{W}}_k = \mathbf{V}_{k_r}^T\mathbf{W}_k\mathbf{V}_{k_r}, \;\; \tilde{\mathbf{g}_k} = \mathbf{V}_{k_r}^T[\mathbf{g}_k + \mathbf{W}_k\mathbf{A}_k^\dagger(-\mathbf{a}_k)]$$

Here the parameters $\mathbf{V}_{k_r}$, $\mathbf{F}$ and $\mathbf{a}$ are the same as those that appeared in the SCP method in Sec. 3.1.1. With α in $\mathbf{a}$ set sufficiently small (α was set to 0.01 in our experiments), the QP problem in (3.11) is always feasible. It was found in practice that $\tilde{\mathbf{W}}_k$ is positive definite when the iterate is in the neighbourhood of the optimal solution $\{\mathbf{h}^*, \boldsymbol{\lambda}^*\}$. In the case where $\tilde{\mathbf{W}}_k$ becomes indefinite, i.e., some of the

eigenvalues of $\tilde{\mathbf{W}}_k$ are less than zero, we modify $\tilde{\mathbf{W}}_k$ as

$$\tilde{\mathbf{W}}_k := \tilde{\mathbf{W}}_k + (\epsilon - r)\mathbf{I}_N$$

where r is the minimum eigenvalue of $\tilde{\mathbf{W}}_k$, ϵ is a small positive value, and $\mathbf{I}_N$ is an identity matrix of size N by N. In this way, $\tilde{\mathbf{W}}_k$ is positive definite and problem (3.11) becomes a convex QP problem (see Sec. 2.4).

If the SQP algorithm starts with an initial point $\{\mathbf{h}_0, \boldsymbol{\lambda}_0\}$ that is sufficiently close to $\{\mathbf{h}^*, \boldsymbol{\lambda}^*\}$, the performance of the SQP method is found to be very satisfactory. But if the algorithm has to start with an initial point that is rather remote from the solution, it may fail to converge [7]. In Sec. 3.3, we present a technique for the design of potentially global LS CQ filters, where an initial point sufficiently close to the solution point is furnished.

3.1.3 Experimenal Results of Local LS Designs

In this section, we make comparisons between the SCP method and the direct design method in [16] for local LS designs. The performance of the SQP method for local designs is not evaluated because the convergence of the SQP method cannot be ensured with an arbitrary initial point, especially when the initial point is quite distant from the locally optimal solution.

The SCP method and the direct method were implemented to design locally optimal LS CQ filters with $N = 30, L = 2$, and $\omega_a = 0.6\pi$. For all the designs, the initial point used was the same linear-phase lowpass filter obtained by the conventional window-based technique [2]. It was found that the smallest tolerance for the direct method to achieve convergence is approximately 1e-7, whereas the SCP method can achieve convergence with a tolerance as small as 1e-18. To compare the two methods under the same design specifications, we set the convergence tolerance to $\varepsilon =$ 1e-7. Table 3.1 lists the performance of the two filters in terms of the stopband energy, the largest equation error among all constraints in (3.1b) and (3.1c) and the CPU time on a PC laptop with a 1.66 GHz dual-core processor. We observe that the SCP method produced a filter not only satsifying constraints (3.1b) and (3.1c) with a higher precision than the direct method, but also possessing a smaller stopband energy. Magnitude responses for the LS CQ filters designed from these two methods are provided in Figs. 3.1 and 3.3. It can be observed that the shape of magnitude response exhibited in Fig. 3.1 looks better. We also note that the SCP method takes

slightly longer CPU time to finish the design.

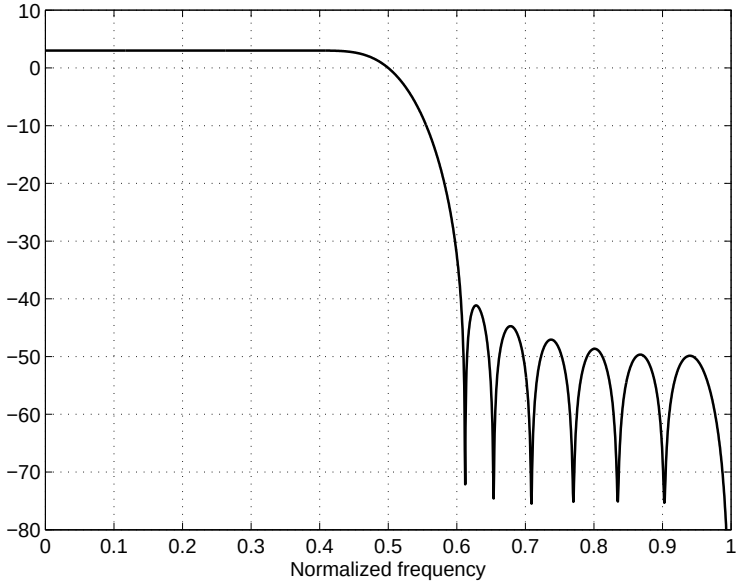

Figure 3.1: Magnitude response of local LS filter with $N = 30$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the SCP method.

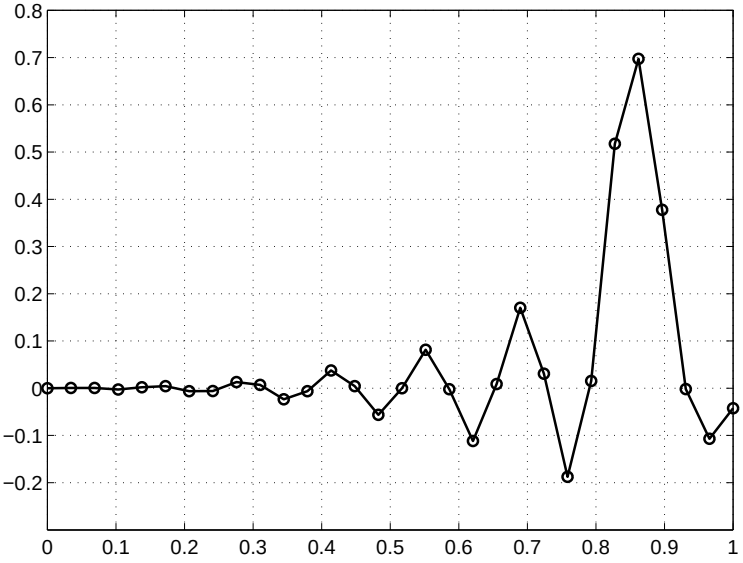

Figure 3.2: Impulse response of local LS filter with $N = 30$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the SCP method.

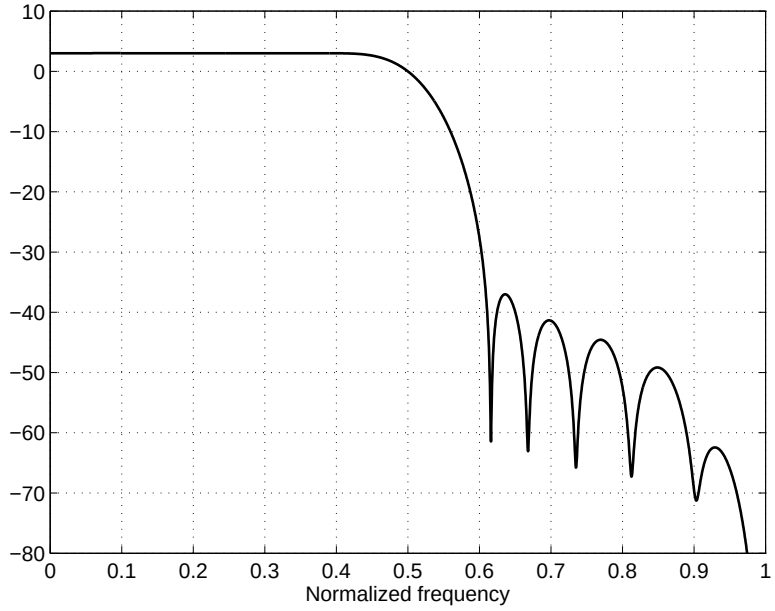

Figure 3.3: Magnitude response of local LS filter with $N = 30$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the direct method.

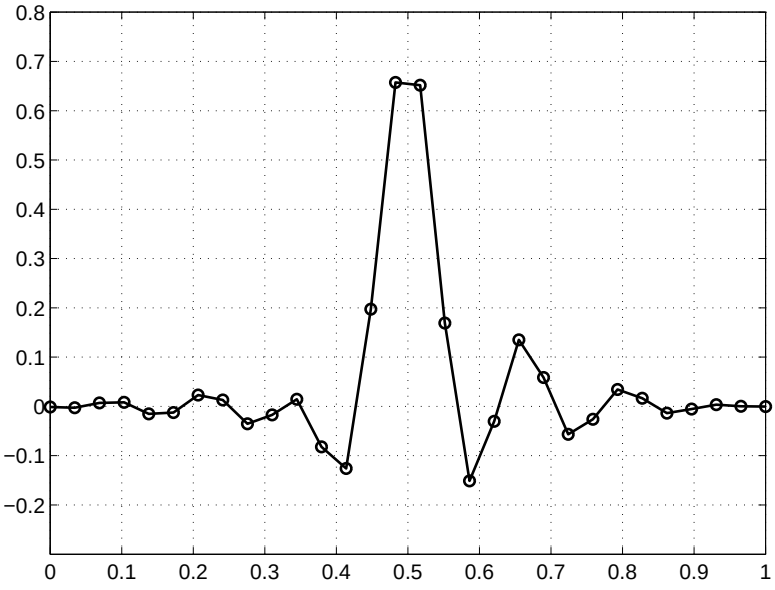

Figure 3.4: Impulse response of local LS filter with $N = 30$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the direct method.

	SCP	Direct method in [16]
Energy in stopband	1.97e-5	5.38e-5
Largest equation error	6.80e-5	9.00e-4
CPU time	4.13s	1.87s

Table 3.1: Performance of LS filters designed using different local methods

3.2 Global LS Design of Low-Order CQ Filter Banks

As argued earlier, the LS design of CQ filters as formulated in (3.1) is a nonconvex problem. As such the problem possesses multiple local solutions. Nevertheless, problem (3.1) is a polynomial optimization problem (POP) (see Sec. 2.6) with $N/2 + L$ or $N/2$ constraints depending on whether or not the VM requirement is included, and the maximum order of all the polynomials involved is two. Therefore, in principle Lasserre's approach [22] is a natural candidate technique for finding the global solution of (3.1).

For low-order filter banks, toolbox GloptiPoly was found to work well. For example, with $N = 6$, $L = 2$ and $\omega_a = 0.56\pi$, the software produces four globally optimal impulse responses as

$$\mathbf{h}_{\mathrm{LS}}^{(6,2)} = \begin{bmatrix} 0.33268098788629 \\ 0.80689591454849 \\ 0.45986215652386 \\ -0.13501431772967 \\ -0.08543638600240 \\ 0.03522516035714 \end{bmatrix}$$

$-\mathbf{h}_{\mathrm{LS}}^{(6,2)}$, $\mathtt{flipud}(\mathbf{h}_{\mathrm{LS}}^{(6,2)})$ and $-\mathtt{flipud}(\mathbf{h}_{\mathrm{LS}}^{(6,2)})$ where $\mathtt{flipud}(\mathbf{h})$ denotes a vector generated by flipping vector $\mathbf{h}$ upside down. We remark that the above four impulse responses satisfy constraints (3.1b) and (3.1c) and they yield the same minimum objective function value as $\mathbf{h}_{\mathrm{LS}}^{(6,2)^T}\mathbf{Q}\mathbf{h}_{\mathrm{LS}}^{(6,2)} = 0.173458$. Also note that $\mathbf{h}_{\mathrm{LS}}^{(6,2)}$ and $-\mathbf{h}_{\mathrm{LS}}^{(6,2)}$ possess minimum phase[1] because no zeros of their corresponding transfer functions are outside the unit circle. The zero-pole plane of $\mathbf{h}_{\mathrm{LS}}^{(6,2)}$ or $(-\mathbf{h}_{\mathrm{LS}}^{(6,2)})$ is shown in Fig. 3.5.

Unfortunately, the software fails to work if the filter length N is greater than or

[1] A linear, time-invariant system is said to be minimum-phase if the system and its inverse are causal and stable [13].

equal to 10. On the other hand, toolbox SparsePOP was found to work for the global design of filter banks up to $N = 16$. A technical problem with SparsePOP is that, unlike GloptiPoly being able to produce multiple global solutions, it requires to set a lower bound and an upper bound for the impulse response and only one global solution that falls within the bounds will be generated. Our design experiences have indicated that the following bounds work well:

$$\mathbf{h_d} - 0.5\mathbf{e} \leq \mathbf{h} \leq \mathbf{h_d} + 0.5\mathbf{e} \tag{3.12}$$

where $\mathbf{h_d}$ is the impulse response of the length-N Daubechies filter [41] and $\mathbf{e}$ is an $N \times 1$ all-one vector. SparsePOP fails to work properly whenever $N > 16$.

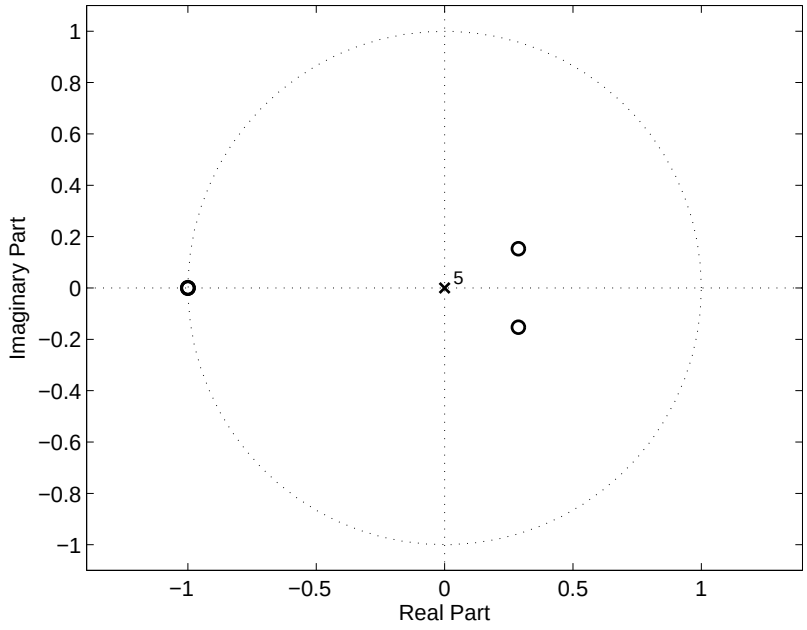

Figure 3.5: Zero-pole plot of $\mathbf{h}_{\mathrm{LS}}^{(6,2)}$.

3.3 Potentially Global LS Design of High-Order CQ Filter Banks

We now address the LS design of CQ filter banks and wavelets with length N too high for the above-mentioned software to handle. A key step in the design strategy is the construction of a good initial point for problem $\mathcal{P}_2^{N_{\rm w}+2}$ (see problem (3.1)) and this is done by appropriately zero-padding the global solution of problem $\mathcal{P}_2^{N_{\rm w}}$. In what follows, we first examine a common pattern that impulse responses of globally optimal CQ filter banks of low-order appear to share.

3.3.1 Pattern of Impulse Responses of Globally Optimal Low-Order Filter Banks

Shown in Fig. 3.6 are the impulse responses of globally optimal minimum-phase lowpass CQ filters of lengths $N = 6,\ 8$ and 10 (all with $L = 2$) obtained using SparsePOP where the impulse responses are plotted over normalized interval $[0,\ 1]$ for better comparison. From this figure, it is clear that these impulse responses are distinctly different from each other. Nevertheless, it is equally clear that they exhibit a similar pattern: each impulse response starts with a short uphill to peak, then goes down to components of small values. In addition, viewing each impulse response as a curve (function), we see that the nearest neighbor to a given curve associated with filter length N is the curve associated with length $N + 2$.

Furthermore, for a fixed filter length N the impulse responses of globally optimal CQ filters with various VMs are clustered and exhibit a pattern similar to that in Fig. 3.6. As an example, Fig. 3.7 shows the impulse responses of lowpass CQ filters with $N = 8$ and $L = 0,\ 1,\ 2,\ 3,\ 4$ obtained using GloptiPoly or SparsePOP.

In what follows, it will become transparent that the observations made on the pattern of the impulse responses of globally optimal low-order designs provide useful clues for tackling the design of potentially globally optimal high-order CQ filters.

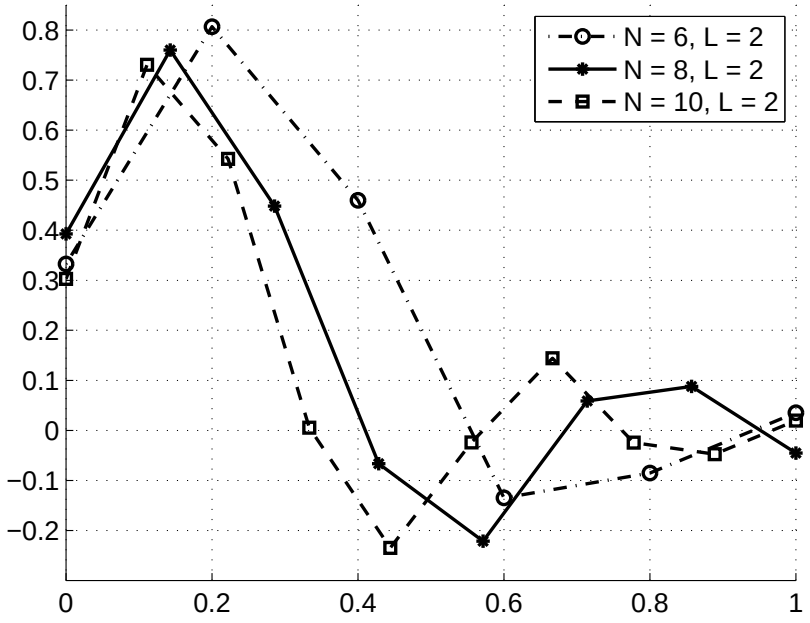

Figure 3.6: Pattern of LS impulse responses with different length N.

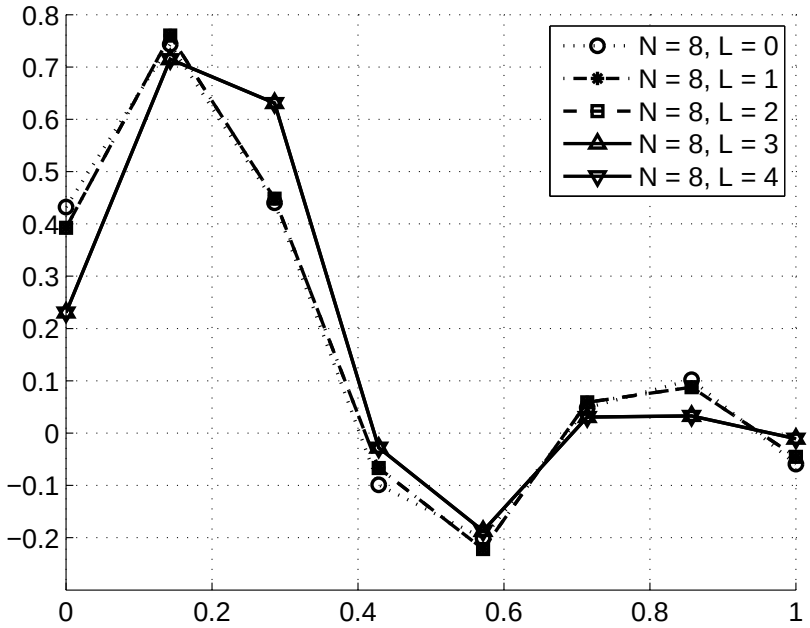

Figure 3.7: Pattern of LS impulse responses with various VM L.

3.3.2 A Design Strategy

The situation facing the designer may be summarized as follows:

(i) Global designs of CQ filters are possible by using the methods of [22, 44], but only for short filter lengths;

(ii) A common pattern exists among the impulse responses of globally optimal low-pass minimum-phase CQ filters of short lengths and, the optimal impulse response of length $N+2$ falls within a small vicinity of the optimal impulse response of length N;

(iii) A local design method that requires a reasonable initial impulse response for producing a locally optimal design is within our reach.

Let $\mathbf{h}_{N_\mathrm{w}}$ be the impulse response of the global solution of problem $\mathcal{P}_2^{N_\mathrm{w}}$ (see Problem (3.1)) and $H_0^{(N_\mathrm{w})}(e^{j\omega})$ be the frequency response of the CQ filter with $\mathbf{h}_{N_\mathrm{w}}$ as impulse response. We construct an impulse response of length $N_\mathrm{w}+2$ by padding two zeros at the end of $\mathbf{h}_{N_\mathrm{w}}$, and denote it by $\mathbf{h}^{\mathrm{zp}}_{N_\mathrm{w}+2}$. We argue that $\mathbf{h}^{\mathrm{zp}}_{N_\mathrm{w}+2}$ is an appropriate initial point for problem $\mathcal{P}_2^{N_\mathrm{w}+2}$. As a matter of fact, from Fig. 3.6 we see that dilating $\mathbf{h}_{N_\mathrm{w}}$ to get an $(N_\mathrm{w}+2)$-dimensional initial point by zero-padding is intuitively reasonable because the tail of every global impulse response in the figure consists of small-value components. A graphical illustration of the effect of zero-padding is given in Fig. 3.8 where $\mathbf{h}_8^{\mathrm{zp}}$ obtained by zero-padding the globally optimal impulse response $\mathbf{h}_6$ (with $L=2$) is shown to be close to the globally optimal impulse response $\mathbf{h}_8$ (with $L=2$). In an early attempt [46], an $(N_\mathrm{w}+2)$-dimensional initial point, denoted by $\mathbf{h}^{\mathrm{interp}}_{N_\mathrm{w}+2}$, was constructed by linearly interpolating $\mathbf{h}_{N_\mathrm{w}}$. As evidenced in Fig. 3.8, $\mathbf{h}_8^{\mathrm{zp}}$ gets closer to $\mathbf{h}_8$ than $\mathbf{h}_8^{\mathrm{interp}}$ does.

Fig. 3.9 depicts the ℓ_2 distance $\|\mathbf{h}_N^{\mathrm{zp}} - \mathbf{h}_N\|$ between the zero-padded initial point and the global solution point, and the ℓ_2 distance $\|\mathbf{h}_N^{\mathrm{interp}} - \mathbf{h}_N\|$ between the linearly interpolated initial point and the global solution point for even N from 12 to 96 (the computation of global solution $\mathbf{h}_N$ is reported in Sec. 3.4). It is observed that $\|\mathbf{h}_N^{\mathrm{zp}} - \mathbf{h}_N\|$ is consistently less than $\|\mathbf{h}_N^{\mathrm{interp}} - \mathbf{h}_N\|$ for the same N.

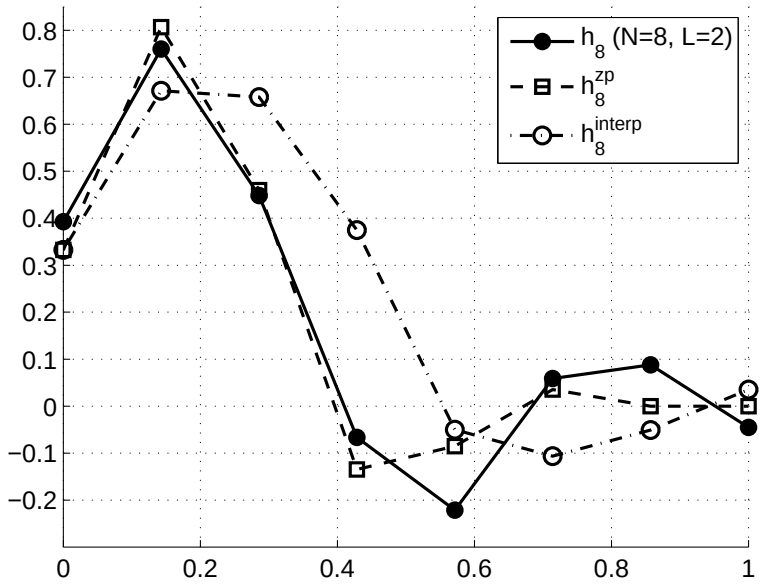

Figure 3.8: Zero-padded and linearly interpolated initial impulse responses with respect to the globally optimal impulse response of that order.

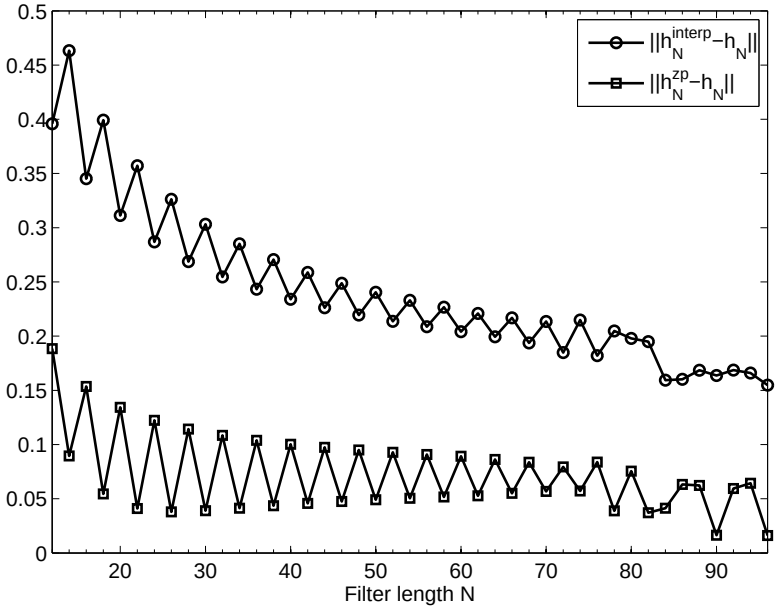

Figure 3.9: Comparison of zero-padded initial point and linearly interpolated initial point in terms of their ℓ_2 distances to globally optimal point for LS designs.

The use of $\mathbf{h}_{N_\mathrm{w}+2}^{\mathrm{zp}}$ can be further justified by several desirable properties it offers:

(i) $\mathbf{h}_{N_\mathrm{w}+2}^{\mathrm{zp}}$ is a feasible point for problem $\mathcal{P}_2^{N_\mathrm{w}+2}$ (i.e. it is in the feasible region characterized by constraints (3.1b) and (3.1c)) because it satisfies both constraints where N is taken to be $N_\mathrm{w}+2$. Verification of this property is straightforward as soon as one realizes that $\mathbf{h}_{N_\mathrm{w}+2}^{\mathrm{zp}}$ is merely a zero-padded $\mathbf{h}_{N_\mathrm{w}}$, and $\mathbf{h}_{N_\mathrm{w}}$ is a solution of $\mathcal{P}_2^{N_\mathrm{w}}$ hence satisfies (3.1b) and (3.1c) (where N is taken to be N_w).

(ii) If we denote the frequency response of the initial CQ filter corresponding to $\mathbf{h}_{N_\mathrm{w}+2}^{\mathrm{zp}}$ by $H_0^{(N_\mathrm{w}+2,\mathrm{zp})}(e^{j\omega})$, it is immediately clear that

$$\int_{\omega_a}^{\pi} |H_0^{(N_\mathrm{w}+2,\mathrm{zp})}(e^{j\omega})|^2 d\omega = \int_{\omega_a}^{\pi} |H_0^{(N_\mathrm{w})}(e^{j\omega})|^2 d\omega$$

That is, the stopband energy of the initial CQ filter (whose length is $N_\mathrm{w}+2$) is the smallest amongst all LS CQ filters of length N_w.

(iii) If $H_0^{(N_\mathrm{w})}(z)$ is a minimum-phase filter, so is $H_0^{(N_\mathrm{w}+2,\mathrm{zp})}(z)$, because the two zero components padding to $\mathbf{h}_{N_\mathrm{w}}$ simply means that the roots of $H_0^{(N_\mathrm{w}+2,\mathrm{zp})}(z)=0$ are the union of the roots of $H_0^{(N_\mathrm{w})}(z)=0$ and two additional roots at the origin.

We remark that these properties do not hold if an interpolated initial impulse response $\mathbf{h}_{N_\mathrm{w}+2}^{\mathrm{interp}}$ is used. Thus, the construction of a suitable initial point described above leads to Algorithm 1 below. In the algorithm, it is implicitly assumed that the filter length N (always an even integer) is too high for the method of [22] to handle.

Algorithm 3.1 (for global LS design of CQ filter banks and wavelets)
Input Data:
Filter length N and stopband edge ω_a.
Step 1
Set an initial working filter length $N_\mathrm{w}=4$. Design a globally optimal, minimum-phase, LS CQ filter of length N_w using the method of [22]. Denote the impulse response obtained by $\mathbf{h}_{N_\mathrm{w}}$.
Step 2
Construct an initial point $\mathbf{h}_{N_\mathrm{w}+2}^{\mathrm{zp}}$ for problem $\mathcal{P}_2^{N_\mathrm{w}+2}$ by padding two zeros at the end of $\mathbf{h}_{N_\mathrm{w}}$.

Step 3
Solve problem $\mathcal{P}_2^{N_w+2}$ using a local design method reported in Sec. 3.1. Denote the impulse response of the CQ filter obtained by $\mathbf{h}_{N_w+2}$.
Step 4
If $N = N_w + 2$, output $\mathbf{h}_{N_w+2}$ as the optimal design and stop; otherwise set $N_w := N_w + 2$ and repeat from Step 2.

Although no theoretical claim about the global optimality of the above design methodology can be made for large N, we speculate that the designs obtained by this approach are quite likely to be globally optimal. This is because in each round of iteration the initial point is sufficiently close to the global minimizer and the local algorithm in Sec. 3.1 is known to converge to a nearby minimizer. In the next section, we provide experimental evidence that supports this speculation.

3.4 Design Examples and Performance Evaluation

3.4.1 Performance of the Proposed Method for Globally Optimal Low-order Designs

The method described in Sec. 3.3.2 was applied to design lowpass minimum-phase LS CQ filters of length $N = 6,\ 8,\ ...,\ 16$. For all designs, the normalized stopband edge was set to $\omega_a = 0.56\pi$ and the number of VMs was set to $L = 1$. In each design, toolbox GloptiPoly was applied only once to (3.1) to generate a globally optimal minimum-phase impulse response with $N = 4$ and $L = 1$, denoted by $\mathbf{h}^{(4,1)}$. In the case of $N = 6$, $\mathbf{h}^{(4,1)}$ was zero-padded to length 6 and then used as the initial point to run the SCP (or SQP) method (note that for the SQP method, the initial $\boldsymbol{\lambda}_0$ is always chosen as an all-zero vector for all our designs), and the impulse response obtained is denoted by $\hat{\mathbf{h}}^{(6,1)}$. In the case of $N = 8$, we first obtained impulse response $\hat{\mathbf{h}}^{(6,1)}$ as above; then $\hat{\mathbf{h}}^{(6,1)}$ was zero-padded to length 8 and used as the initial point to run the SCP (or SQP) method to generate impulse response $\hat{\mathbf{h}}^{(8,1)}$. The designs for $N = 10, 12, 14, 16$ were carried out in a similar manner to produce impulse responses $\hat{\mathbf{h}}^{(N,1)}$.

For comparison purposes, globally optimal impulse responses $\mathbf{h}^{(N,1)}$ for $N = 6,\ 8,\ ...,\ 16$ were obtained using GloptiPoly or SparsePOP. It was found that $\hat{\mathbf{h}}^{(N,1)}$ and $\mathbf{h}^{(N,1)}$ are practically identical for all even N from 6 to 16. We also remark that

with the starting impulse response $\mathbf{h}^{(4,1)}$ having minimum phase, the LS CQ filters so designed all have minimum phase, a desirable property for digital filters to be of practical use.

3.4.2 Performance of the Proposed Method for Potentially Globally Optimal High-order Designs

Supported by the verification of our design concept in Sec. 3.4.1, we now proceed to apply the proposed method to design high-order LS lowpass CQ filters with length N up to 96. As an example, a filter with $N = 96$, $L = 3$, and $\omega_a = 0.56\pi$ was designed. The convergence tolerance was set to $\varepsilon =$ 1e-15. Two local methods, i.e., the SCP and SQP methods with zero-padded initial impulse responses were applied in Step 3 of Algorithm 3.1. The energy in stopband ($\mathbf{h}^T\mathbf{Q}\mathbf{h}$ in (3.1a)), the largest equation error among all constraints in (3.1b) and (3.1c), and the CPU time on a PC laptop with a 1.66 GHz dual-core processor for the two methods are listed in Table 3.2.

We observe that the two methods work quite effectively because globally optimal impulse responses satisfying constraints (3.1b) and (3.1c) with high precision were calculated in a relatively short period of time. We also note that the SQP method achieved a slightly better performance than the SCP method, demonstrating the excellent performance of the SQP method when the initial point is close enough to the optimal solution point. The ℓ_2-norm distance between the two impulse-response vectors was found to be 6.8772e-5, meaning that the global design with SCP method and the global design with SQP method practically converged to the same solution. The magnitude response and the impulse response of the globally optimal LS filter designed by the proposed method are depicted in Figs. 3.10 and 3.11, respectively.

	SCP	SQP
Energy in stopband	1.18103e-9	1.18101e-9
Largest equation error	1e-14	4e-15
CPU time	46s	42s

Table 3.2: Performance of globally optimal LS filters of length 96

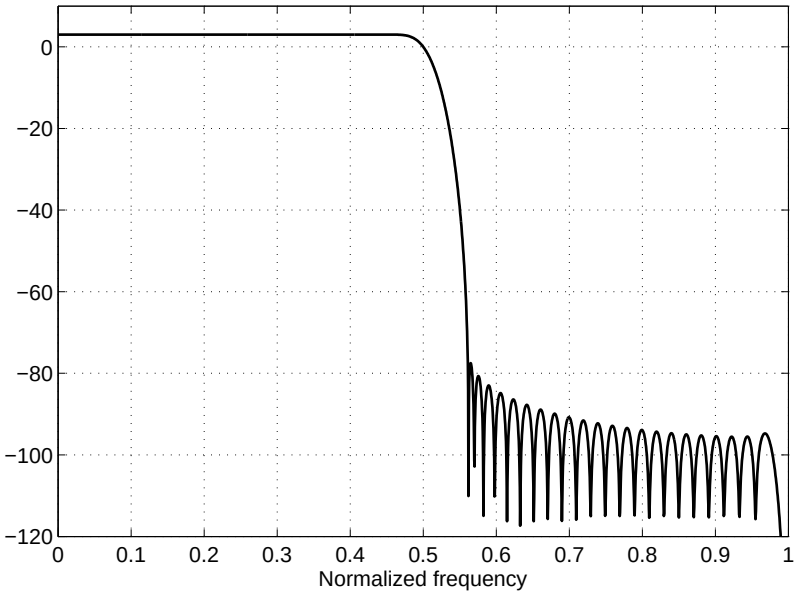

Figure 3.10: Magnitude response of globally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

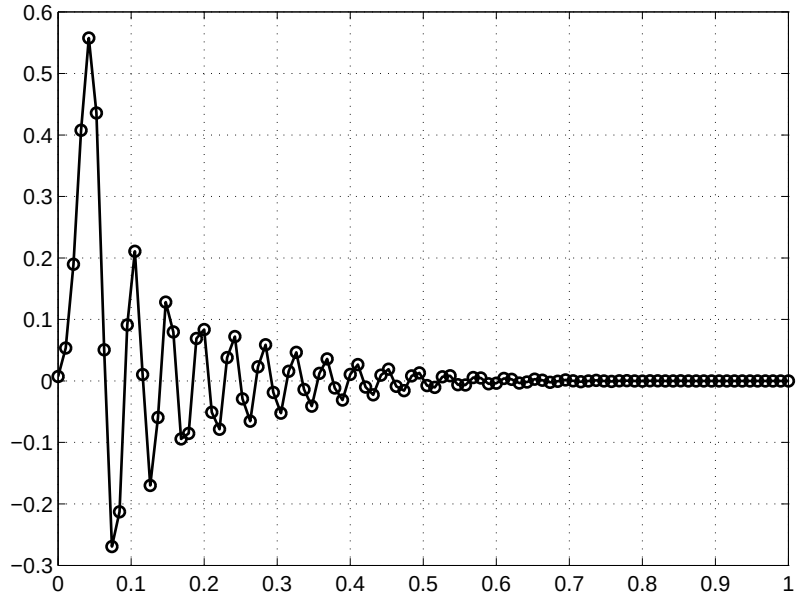

Figure 3.11: Impulse response of globally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

For comparison, a locally optimal LS CQ filter with $N = 96$, $L = 3$, and $\omega_a = 0.56\pi$ was also designed under convergence tolerance $\varepsilon =$ 1e-15 by directly applying the SCP method (see 3.1.1). The initial point used was a linear-phase lowpass filter obtained by the conventional window-based technique [2]. The performance of the local filter designed is listed in Table 3.3. Comparing Table 3.2 with Table 3.3, it is observed that the performance of the LS CQ filters from the global designs indeed excels that from the local design in terms of energy in stopband and largest equation error. The magnitude response and impulse response of the locally optimal filter are shown in Figs. 3.12 and 3.13, respectively. We observe that the stopband attenuation of the globally optimal filter is larger than that of the locally optimal filter. All the observations made above support our belief that the proposed method produces globally optimal CQ filters.

The zero-pole plots of the LS lowpass CQ filters of length 96 obtained using the global and local methods are shown in Figs. 3.14 and 3.15, respectively. It follows that the CQ filter designed from the local method does not possess minimum phase because there are zeros outside the unit circle of the complex plane, whereas the globally optimal CQ filter possesses minimum phase because no zeros are outside the unit circle of the complex plane.

	Local design based on SCP
Energy in stopband	3.15564e-9
Largest equation error	1e-14
CPU time	27s

Table 3.3: Performance of a locally optimal LS filter of length 96

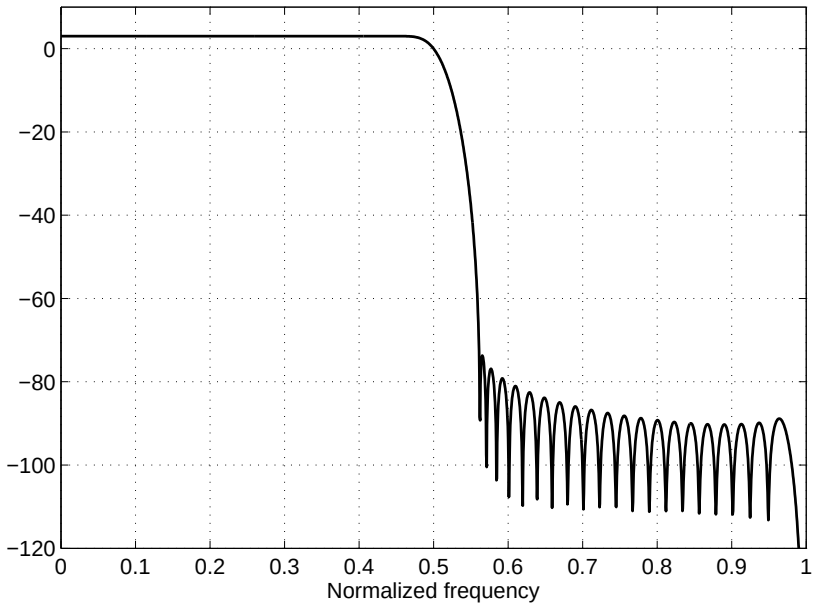

Figure 3.12: Magnitude response of locally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

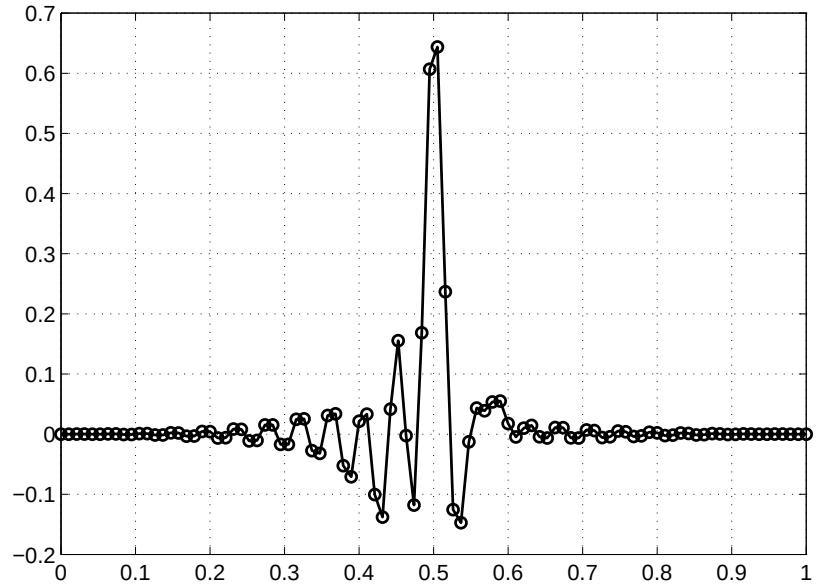

Figure 3.13: Impulse response of locally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

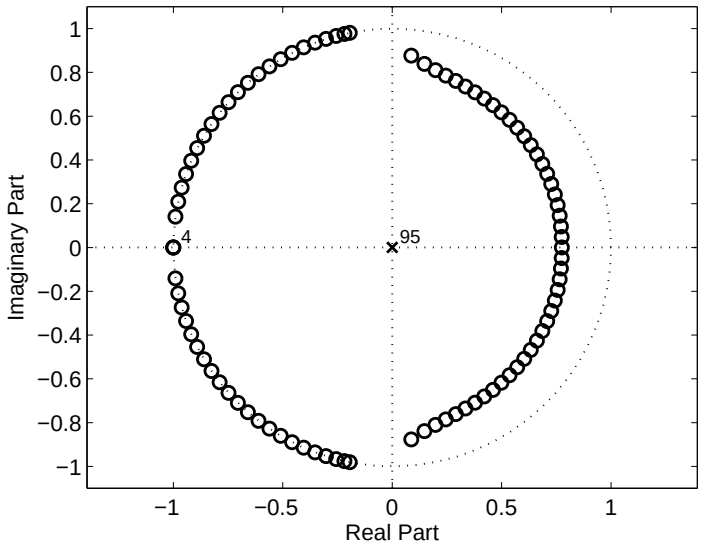

Figure 3.14: Zero-pole plot of globally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

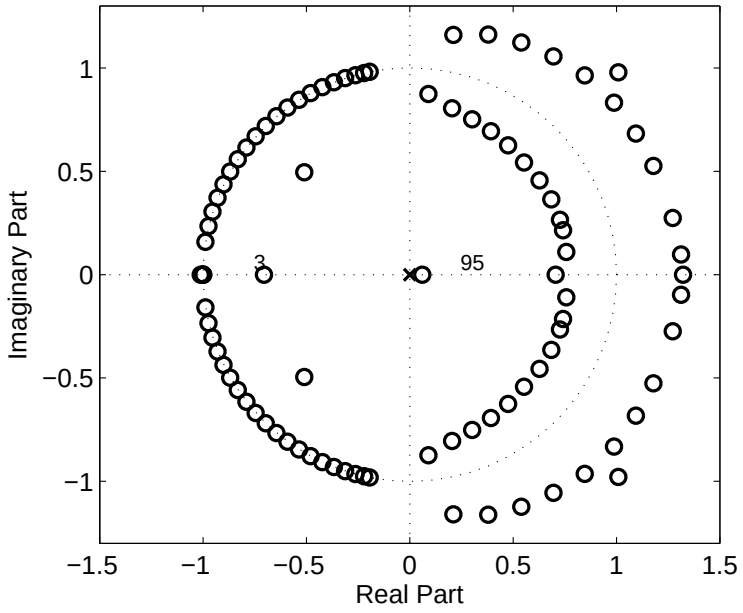

Figure 3.15: Zero-pole plot of locally optimal LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

Chapter 4

Minimax Design of Orthogonal Filter Banks and Wavelets

This chapter is focused on the minimax design of orthogonal filter banks and wavelets. It follows from Sec. 2.1 that the design problem under consideration can be formulated as

$$\begin{aligned} \text{minimize} \quad & \eta && \text{(4.1a)} \\ \text{subject to:} \quad & \|\mathbf{T}(\omega)\cdot\mathbf{h}\| \le \eta \ \text{ for } \omega\in\Omega && \text{(4.1b)} \\ & \sum_{n=0}^{N-1-2m} h_n\cdot h_{n+2m} = \delta_m && \text{(4.1c)} \\ & \sum_{n=0}^{N-1}(-1)^n\cdot n^l\cdot h_n = 0 && \text{(4.1d)} \end{aligned}$$

where $m = 0,\ 1,\ ...,\ (N-2)/2$ and $l = 0,\ 1,\ ...,\ L-1$.

The remainder of this chapter is organized in the following way. Sec. 4.1 studies the local method for the minimax design of conjugate quadrature (CQ) filter banks. The related convergence issue has also been addressed therein. Next, global minimax design of low-order CQ filter banks is studied in Sec. 4.2. In Sec. 4.3, we propose two methods for global minimax design of high-order CQ filter banks. To conclude this chapter, performance of the proposed methods for high-order designs is illustrated in Sec. 4.4, and comparisons with other existing methods well established in the literature are evaluated in Sec. 4.5, separately.

4.1 Local Minimax Design of CQ Filter Banks

Like the least squares (LS) design problem studied in Chapter 3, the minimax design of CQ filters in (4.1) is also a nonconvex problem, hence the problem possesses multiple local solutions. A direct design technique for locally optimal minimax CQ filters is proposed in [16]. In this section we propose several improvements over the design method in [16]. First, the direct method in [16] is modified such that the optimization problem involved is always feasible. This modification is similar to that done for the sequential convex-programming (SCP) method, see Sec. 3.1. Next, a convergence issue encountered in the design is investigated and a technique by locating magnitude-response peaks in conjunction with a Gauss-Newton method with adaptively controlled weights is applied for the minimax design to achieve a solution with a high degree of accuracy. By implementing these improvments, the modified method is found to provide excellent performance in designing locally optimal minimax CQ filter banks. Simulations are presented in Sec. 4.1.3 to demonstrate the effectiveness of the improved method.

4.1.1 Direct Design Method with Feasible Optimization Problems

We begin by a brief review of the method in [16]. Using the analysis described in Sec. 3.1.1, in the kth iteration of the direct design method [16], problem (4.1) assumes the form

$$\begin{aligned} \text{minimize} \quad & \eta && \text{(4.2a)} \\ \text{subject to:} \quad & \|\mathbf{T}(\omega)(\mathbf{h}_k + \boldsymbol{\delta}_h)\| \le \eta \ \text{ for } \omega \in \Omega && \text{(4.2b)} \\ & \mathbf{A}_k \boldsymbol{\delta}_h = -\mathbf{a}_k && \text{(4.2c)} \\ & \mathbf{C}\boldsymbol{\delta}_h \le \mathbf{b} && \text{(4.2d)} \end{aligned}$$

where $\mathbf{A}_k$, $\mathbf{a}_k$, $\mathbf{C}$ and $\mathbf{b}$ are the same as given in (3.5). By utilizing SVD of matrix $\mathbf{A}_k$ as in (3.8) to remove the equality constraints, the problem can be reduced to

$$\begin{aligned} \text{minimize} \quad & \eta && \text{(4.3a)} \\ \text{subject to:} \quad & \|\mathbf{T}_k(\omega)\mathbf{x} + \mathbf{e}_k(\omega)\| \le \eta \ \text{ for } \omega \in \Omega && \text{(4.3b)} \\ & \hat{\mathbf{C}}\mathbf{x} \le \hat{\mathbf{b}} && \text{(4.3c)} \end{aligned}$$

where

$$\hat{\mathbf{C}} = \mathbf{C}\mathbf{V}_{k_r}, \;\; \hat{\mathbf{b}} = \mathbf{b} - \mathbf{C}\mathbf{A}_k^{\dagger}(-\mathbf{a}_k)$$

$$\mathbf{T}_k(\omega) = \mathbf{T}(\omega)\mathbf{V}_{k_r}, \;\; \mathbf{e}_k(\omega) = \mathbf{T}(\omega)[\mathbf{h}_k + \mathbf{A}_k^{\dagger}(-\mathbf{a}_k)]$$

As explained in Sec. 3.1.1, this technique may encounter difficulties as problem (4.3) may become infeasible. As a technical remedy, we propose to directly bound $\mathbf{x}$ so that problem (4.3) becomes

$$\begin{aligned} \text{minimize} \quad & \eta && \text{(4.4a)} \\ \text{subject to:} \quad & \|\mathbf{T}_k(\omega)\mathbf{x} + \mathbf{e}_k(\omega)\| \le \eta \;\; \text{for } \omega \in \Omega && \text{(4.4b)} \\ & \mathbf{F}\mathbf{x} \le \mathbf{a} && \text{(4.4c)} \end{aligned}$$

where $\mathbf{F}$ and $\mathbf{a}$ are defined in the same way as those in Eq. (3.10b). Compared with (4.3), (4.4) is always feasible no matter how small α is. With regard to variable $\{\eta, \;\mathbf{x}\}$, (4.4) is a second-order cone programming (SOCP) problem [1] for which efficient solvers such as SeDuMi [51] exist. After the unique and global solution $\mathbf{x}_k$ of (4.4) is obtained, we set $\mathbf{h}_{k+1} = \mathbf{h}_k + [\mathbf{V}_{k_r}\mathbf{x}_k + \mathbf{A}_k^{\dagger}(-\mathbf{a}_k)]$ for the next iteration. The algorithm continues until $\|\mathbf{h}_{k+1} - \mathbf{h}_k\|$ is less than a prescribed tolerance ε, and $\mathbf{h}_{k+1}$ is claimed to be the locally optimal impulse response.

4.1.2 A Convergence Issue for Local Minimax Designs

In our experimental studies, the choice of the set of frequency grids of Ω in (4.4) is found to be of critical importance for the algorithm to achieve convergence. Occasionally, even if we uniformly partition the interval $[\omega_a, \;\pi]$ with a high density, the local minimax method may remain unconvergent with respect to a certain tolerance. We speculate that this problem is due to the use of an evenly spaced frequency grid set Ω at which a large number of the inequality constraints in (4.4b) remain inactive.

A. Constructing Ω by locating magnitude-response peaks

To deal with the above problem, we first use a set of evenly spaced dense frequency grids as Ω so that the frequencies achieving all magnitude-response peaks in stopband are practically captured. After a few iterations, although the solution is still not accu-

rate, the locations of the peaks of magnitude response in stopband $[\omega_a,\ \pi]$ tend to be stable and it is possible to locate them all. At this point of the algorithm, we construct Ω with only those frequencies corresponding to the peaks and continue the iterations for problem (4.4). As a result, the number of frequency grids in Ω is considerably reduced that helps accelerate the iterations. More importantly, as iteration continues the peaks become increasingly equalized, hence an increasing number of inequality constraints in (4.4b) become active that facilitates the algorithm's convergence to a highly accurate solution.

B. A Gauss-Newton method with adaptively controlled weights

There are instances when the method described above still does not converge. In these cases, it was found that the equality constraint (4.1c) is not satisfied accurately. Under this circumstance, a post-processing method is applied to further improve the solution accuracy in satisfying the equality constraints in (4.1c).

To be specific, let $\mathbf{h}_r$ be the current iterate. It is certain that there exists an accurate solution $\mathbf{h}^*$ in a neighborhood of $\mathbf{h}_r$. Our objective is to identify such a solution $\mathbf{h}^*$, which satisfies the system of nonlinear equations

$$f_m(\mathbf{h}) = 0 \quad \text{for } m = 0, 1, ..., N/2 - 1 \tag{4.5}$$

with practically equal equation errors, where $f_m(\mathbf{h}) = \sum_{n=0}^{N-1-2m} h_n \cdot h_{n+2m} - \delta_m$. The problem here can be resolved by using the Gauss-Newton (G-N) method with adaptively controlled weights introduced in Sec. 2.3. Here the iterate $\mathbf{h}_r$ is utilized as an initial point for the algorithm so that the solution $\mathbf{h}^*$ sufficiently close to $\mathbf{h}_r$ can be obtained.

Together, the two techniques described above facilitate a satisfactory solution for the convergence issue and lead to a locally optimal solution to problem (4.1) with a high degree of accuracy. For convenience, we refer to the local minimax design method investigated in this section as the SCP-GN method.

4.1.3 Experimental Results of Local Minimax Designs

In what follows, we make comparisons between the SCP-GN method and the direct design method [16] for local minimax designs.

The SCP-GN method was applied to design a locally optimal minimax CQ filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$. The initial point was a linear-phase lowpass filter obtained by the conventional window-based technique [2]. The performance of the filter designed in terms of maximum instantaneous stopband energy, largest equation error among all constraints in (4.1c) and (4.1d) and CPU time is listed in Table 4.1.

For comparison, Table 4.1 also lists the performance of the filter of the same parameters (i.e., $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$) obtained using the minimax method in [16]. The initial point was the same as that used in the SCP-GN method. It can be observed that the SCP-GN method produces a filter which has less maximum stopband energy and satisfies the constraints with reduced equation error. The magnitude responses and the impulse responses associated with the local mimimax filter designed by the SCP-GN method and the direct method respectively are plotted in Figs. 4.1-4.4. It is observed from the figures that the two methods converge to two different local solutions, whereas the minimax filter from the SCP-GN method achieves a better magnitude response. It is important to note that magnitude response of a minimax filter is equiripple over stopband.

	SCP-GN	method in [16]
Maximum instantaneous energy in stopband	1.497e-3	1.537e-3
Largest equation error	2.1e-9	2.0e-4
CPU time	19.3s	25.1s

Table 4.1: Performance of minimax filters designed using different local methods

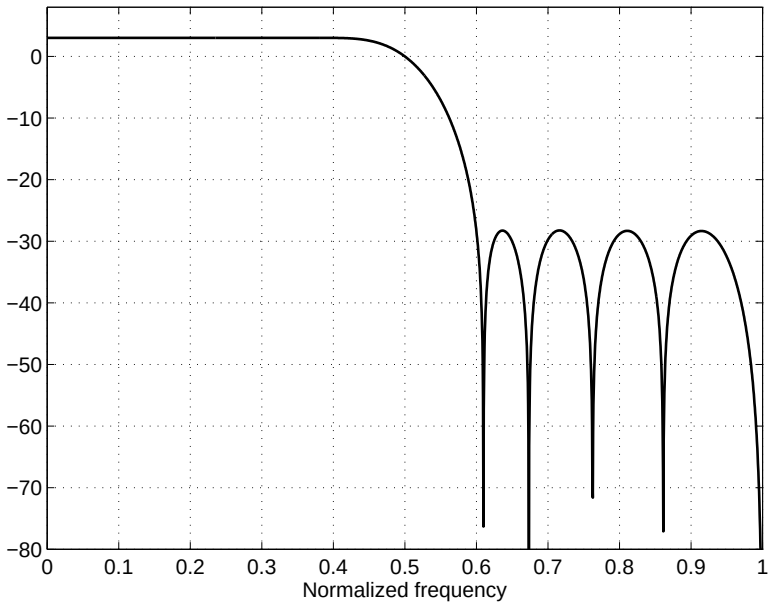

Figure 4.1: Magnitude response of local minimax filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the SCP-GN method.

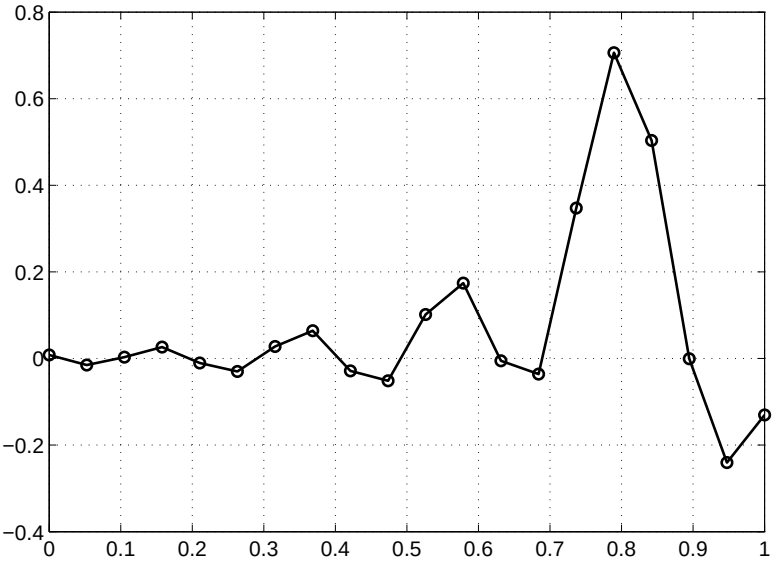

Figure 4.2: Impulse response of local minimax filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the SCP-GN method.

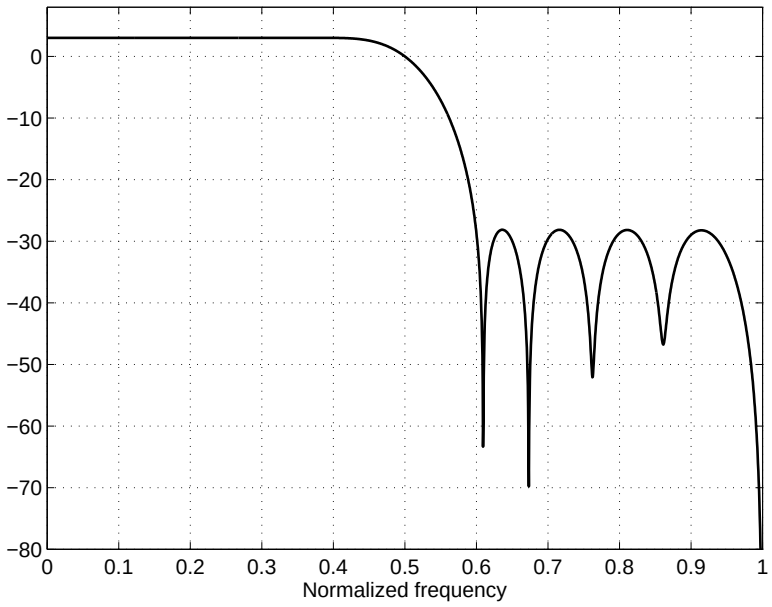

Figure 4.3: Magnitude response of local minimax filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the direct method.

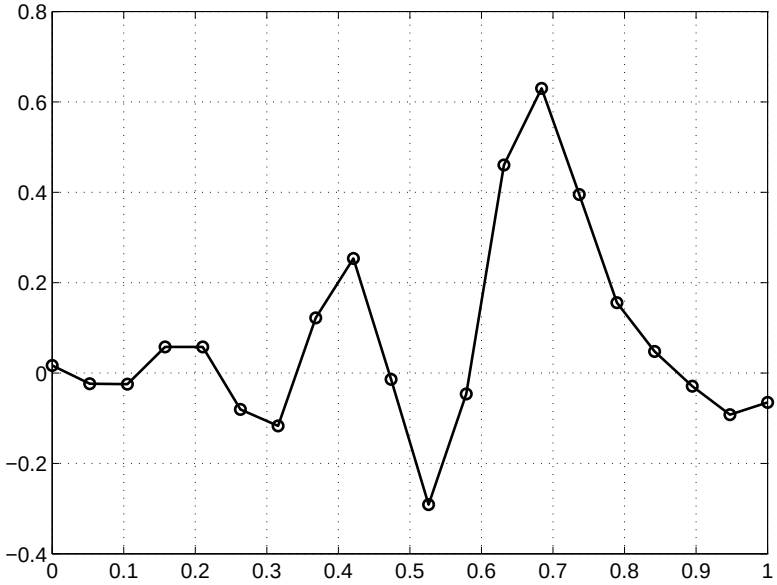

Figure 4.4: Impulse response of local minimax filter with $N = 20$, $L = 2$ and $\omega_a = 0.6\pi$ designed from the direct method.

4.2 Global Minimax Design of Low-Order CQ Filter Banks

Like the LS problem in (3.1), the minimax problem formulated in (4.1) is also a polynomial optimization problem (POP) with $K+N/2+L$ constraints and the maximum order of all the polynomials involved is two. In our experiments, toolbox GloptiPoly was found to work only for filters of order 4. For example, with $N = 4$, $L = 1$, $\omega_a = 0.56\pi$, and $\Omega = \{\omega_a,\ \omega_a + 0.025\pi,\ \omega_a + 0.05\pi,\ ...,\ \pi\}$ (which gives $K = 18$), the toolbox was able to produce four globally optimal impulse responses as

$$\mathbf{h}^{(4,1)}_{\text{minimax}} = \begin{bmatrix} 0.48296282173531 \\ 0.83651623138234 \\ 0.22414405492402 \\ -0.12940935473280 \end{bmatrix}$$

$-\mathbf{h}^{(4,1)}_{\text{minimax}}$, $\texttt{flipud}(\mathbf{h}^{(4,1)}_{\text{minimax}})$ and $-\texttt{flipud}(\mathbf{h}^{(4,1)}_{\text{minimax}})$ where $\texttt{flipud}(\mathbf{h})$ denotes a vector generated by flipping $\mathbf{h}$ upside down. The maximum instantaneous energy in stopband for the above four impulse responses was found to be the same value $\eta^2 = 0.722218$. We also observed that $\mathbf{h}^{(4,1)}_{\text{minimax}}$ and $-\mathbf{h}^{(4,1)}_{\text{minimax}}$ possess minimum phase, as the magnitude of the three zeros of the filter was calculated to be 0.999999953175731, 0.999999953175731 and 0.267948920761278 separately. This can also be observed from the zero-pole plot in Fig. 4.5. However, GloptiPoly failed to work for N even as small as 6 because of the relatively large number of constraints in minimax designs. On the other hand, with the same bounds set for $\mathbf{h}$ as in (3.12), SparsePOP was able to produce global minimax designs for $N = 4$ and 6.

4.3 Potentially Global Minimax Design of High-Order CQ Filter Banks

In this section, we propose two methods for the global minimax design of high-order CQ filter banks.

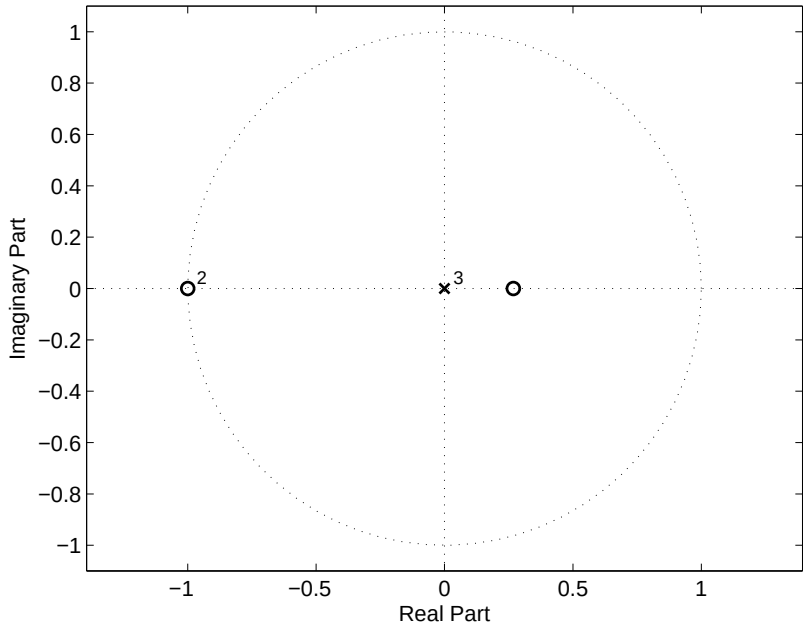

Figure 4.5: Zero-pole plot of $\mathbf{h}_{\text{minimax}}^{(4,1)}$.

Method 1

Similar to the observations made for the LS designs (see Sec. 3.3.1), globally optimal (in minimax sense) impulse responses obtained in Sec. 4.2 appear to exhibit a pattern similar to that in the LS case, as can be seen from Fig. 4.6. With the method of local minimax designs presented in Sec. 4.1 available, it is therefore natural to follow the same strategy as described in Sec. 3.3.2 for minimax designs of high-order filter banks. The zero-padding method is applied to generate an initial point for a higher order design. To justify our choice of the zero-padded initial point instead of an interpolated one, we plot in Fig. 4.7 the ℓ_2 distance $||\mathbf{h}_N^{\text{zp}} - \mathbf{h}_N||$ between the zero-padded initial point and the global solution point, and the ℓ_2 distance $||\mathbf{h}_N^{\text{interp}} - \mathbf{h}_N||$ between the linearly interpolated initial point and the global solution point for filter length N from 6 to 96. It can be observed that $||\mathbf{h}_N^{\text{zp}} - \mathbf{h}_N||$ is consistently less than $||\mathbf{h}_N^{\text{interp}} - \mathbf{h}_N||$. Thus, the zero-padded initial point is deemed as a better choice for producing global minimax designs of high order with satisfactory performance.

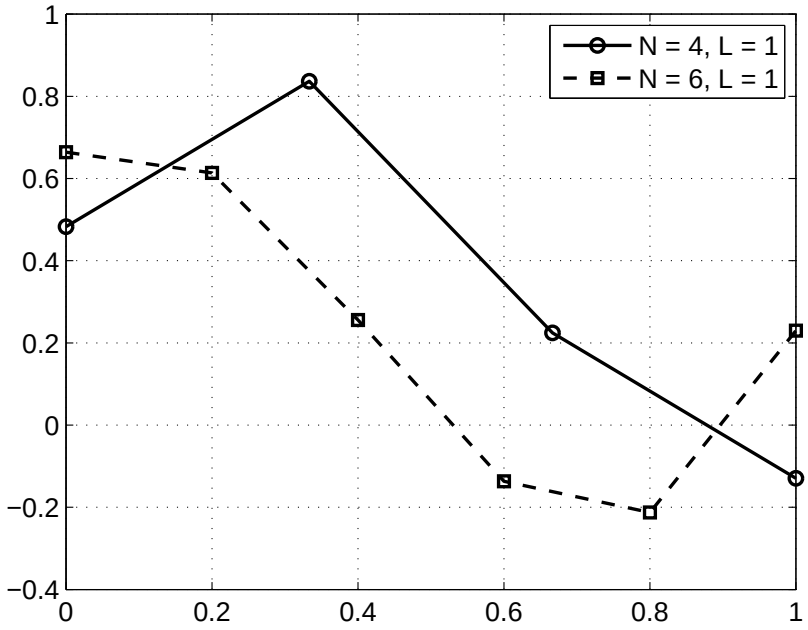

Figure 4.6: Pattern of minimax impulse responses with different length N.

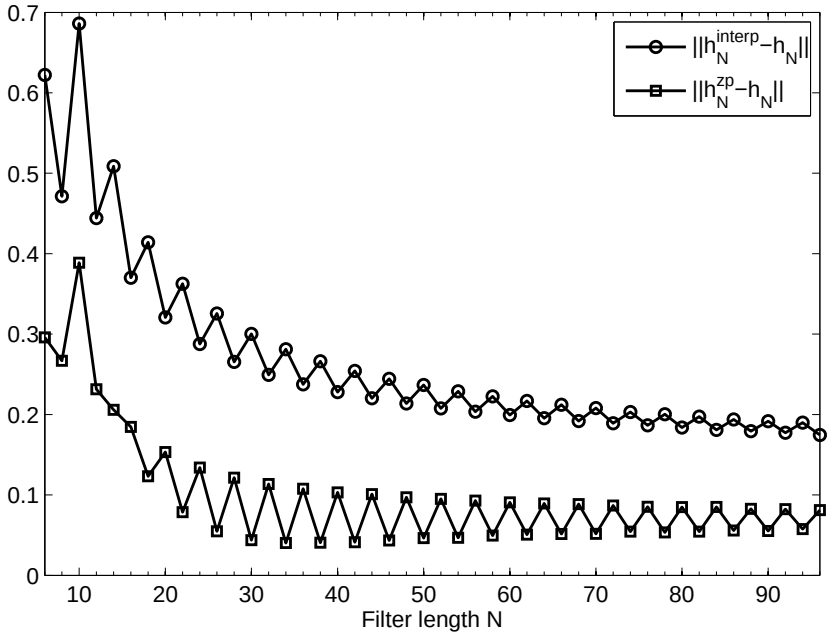

Figure 4.7: Comparison of zero-padded initial point and linearly interpolated initial point in terms of their ℓ_2 distances to globally optimal point for minimax designs.

Method 2

In Chapter 3, we have presented a technique to obtain globally optimal impulse response for an LS lowpass CQ filter. With the global LS filter designed, we propose another design method by simply passing the impulse response of the designed global LS filter as an initial point for the local minimax method to design an optimal minimax filter with the same design specifications N, L and ω_a. The validity of this method is based on the fact that with the same design specifications the globally optimal impulse response of an LS filter is sufficiently close to the globally optimal impulse response of a minimax filter.

An advantage this method offers is an improved design efficiency relative to the first method. By using a good initial point obtained from the global LS design of the same filter length, the order recursion required by the first method is no longer needed. This efficiency improvement is especially pronounced for filter banks of very high order.

4.4 Design Examples and Performance Evaluations

4.4.1 Performance of the Proposed Method for a Low-order Design

The design strategy in Sec. 3.3.2 was applied to design a lowpass minimax CQ filter with $N = 6$, $\omega_a = 0.56\pi$ and $L = 1$. Impulse response $\mathbf{h}_{\text{minimax}}^{(4,1)}$ obtained from GloptiPoly was zero-padded to length 6 and used as the initial point to run the local minimax design algorithm. The choice of Ω follows the technique from Sec. 4.1.2: Initially, we used 80 and 110 frequencies ($K = 190$) to uniformly partition the intervals $[0.56\pi,\ 0.65\pi]$ and $[0.65\pi,\ \pi]$ respectively (the reason we used a set of denser grids on $[0.56\pi,\ 0.65\pi]$ is that the magnitude response tends to have more peaks in this interval as filter length N goes higher). After a number of iterations (30 iterations were performed in all the simulations reported here) and if the problem has not converged yet, we reconstruct Ω with only those frequencies corresponding to the peaks of the magnitude response in stopband and continue the iterations until convergence is achieved. If there is still no convergence after a prescribed number of iterations (e.g., 100), the Gauss-Newton method with adaptively controlled weights is applied. The impulse response obtained is denoted by $\hat{\mathbf{h}}_{\text{minimax}}^{(6,1)}$.

For comparison, SparsePOP was applied to problem (4.1) to generate the globally optimal impulse response $\mathbf{h}_{\text{minimax}}^{(6,1)}$. The two impulse responses, $\hat{\mathbf{h}}_{\text{minimax}}^{(6,1)}$ and $\mathbf{h}_{\text{minimax}}^{(6,1)}$ were found to be practically identical. We also note that with the starting impulse response $\mathbf{h}_{\text{minimax}}^{(4,1)}$ having minimum phase, the CQ filter $\hat{\mathbf{h}}_{\text{minimax}}^{(6,1)}$ possesses minimum phase as well.

4.4.2 Performance of the Proposed Methods for High-order Designs

As an example in this section, we demonstrate the performance of the design methods in Sec. 4.3 by applying them to design a minimax filter bank with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. By implementing Method 1, $\hat{\mathbf{h}}_{\text{minimax}}^{(6,1)}$ was zero-padded to length 8 and used as an initial point for the local minimax algorithm to generate $\hat{\mathbf{h}}_{\text{minimax}}^{(8,1)}$. Such process was repeated until a minimax lowpass CQ filter of length $N = 96$ was designed. In addition, we applied Method 2 to design a minimax filter of length 96 in that a good initial point was supplied to carry out a local minimax design of the same length. This initial point was obtained from the design of a global LS filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$ (see Sec. 3.4.2). The convergence tolerance was set to $\varepsilon =$ 1e-15 for both methods.

Table 4.2 lists the design results in terms of maximum instantaneous energy in stopband, largest magnitude error of equality constraints in (4.1c) and (4.1d), and CPU time on a PC laptop with a 1.66 GHz dual-core processor. The total time required by Method 2 is the sum of the time for the design of a global LS filter and the time for one round of local minimax design. It was observed from Table 4.2 that the minimax filters designed using the two methods have comparable and satisfying performance. The ℓ_2-norm distance between the two impulse responses was 2.8e-3, indicating that the impulse responses obtained using the two methods were practically the same. Whereas for this example, Method 2 was able to produce a filter with less maximum instantaneous stopband energy in a much reduced computational complexity. Figs. 4.8 and 4.9 depict the magnitude response and impulse response of the globally optimal minimax filter designed from Method 2, respectively. It can be observed that the magnitude response is equiripple in stopband. We note that for a minimax filter, the magnitude response is also equiripple in passband, as illustrated in Fig. 4.10.

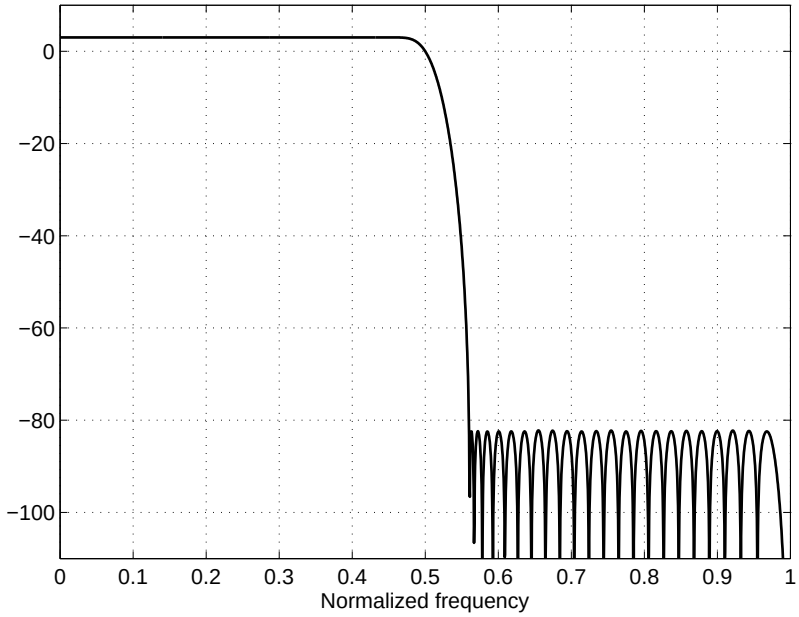

Figure 4.8: Magnitude response of globally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

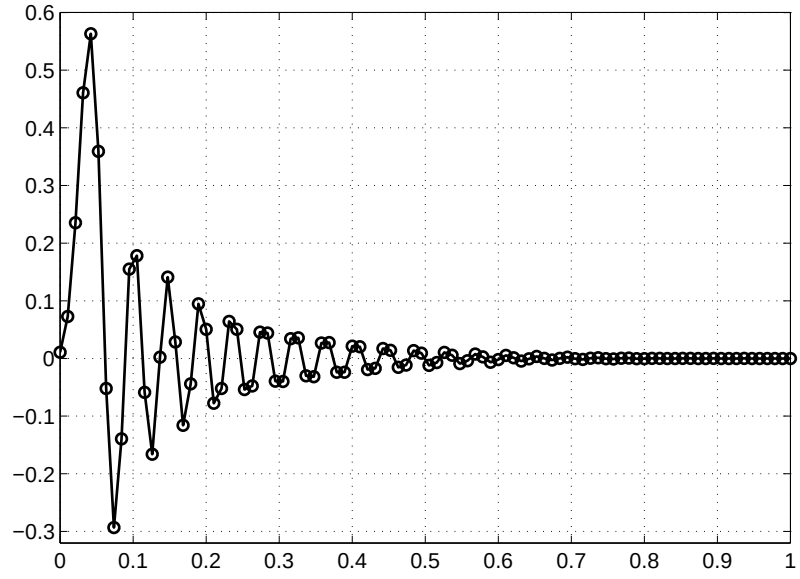

Figure 4.9: Impulse response of globally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

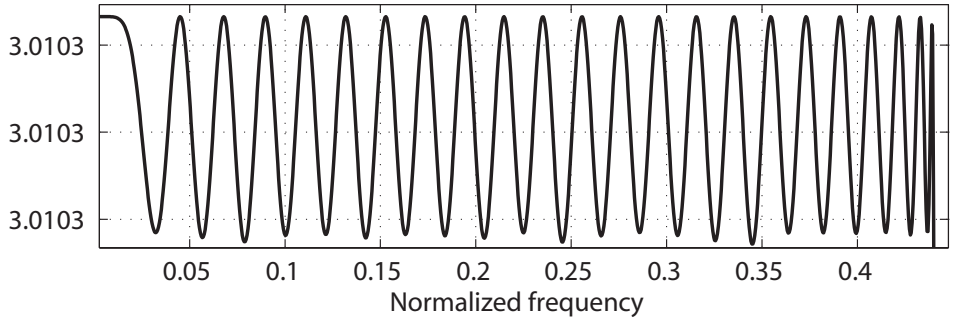

Figure 4.10: Magnitude response in passband of globally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

	Method 1	Method 2
Maximum instantaneous energy in stopband	6.75750e-9	6.02383e-9
Largest equation error	<1e-15	<1e-15
CPU time	1511s	(42+30)s

Table 4.2: Performance of the minimax filters from global designs

For comparison, we applied the SCP-GN method (see Sec. 4.1) to design a locally optimal minimax CQ filter with the same parameters, i.e., $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$. The initial point passed was a linear-phase lowpass filter obtained by the conventional window-based technique [2]. The convergence tolerance was set to $\varepsilon =$ 1e-15. Table 4.3 lists the performance of the designed filter. By comparing Table 4.2 with Table 4.3, it is found that the the global design method indeed produces minimax filters with improved performance. The zero-pole plots of the global and local minimax filters are illustrated in Figs. 4.11 and 4.12, respectively. It was observed that unlike the globally optimal minimax filter, the filter obtained from local method does not possess minimum phase.

	Local design based on SCP-GN
Maximum instantaneous energy in stopband	1.81165e-8
Largest equation error	2.9e-14
CPU time	42s

Table 4.3: Performance of the length-96 minimax filter from local design

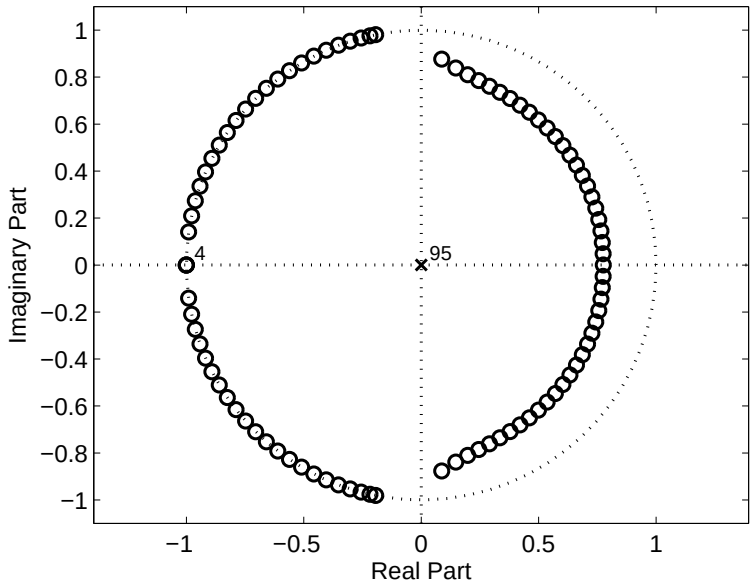

Figure 4.11: Zero-pole plot of globally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

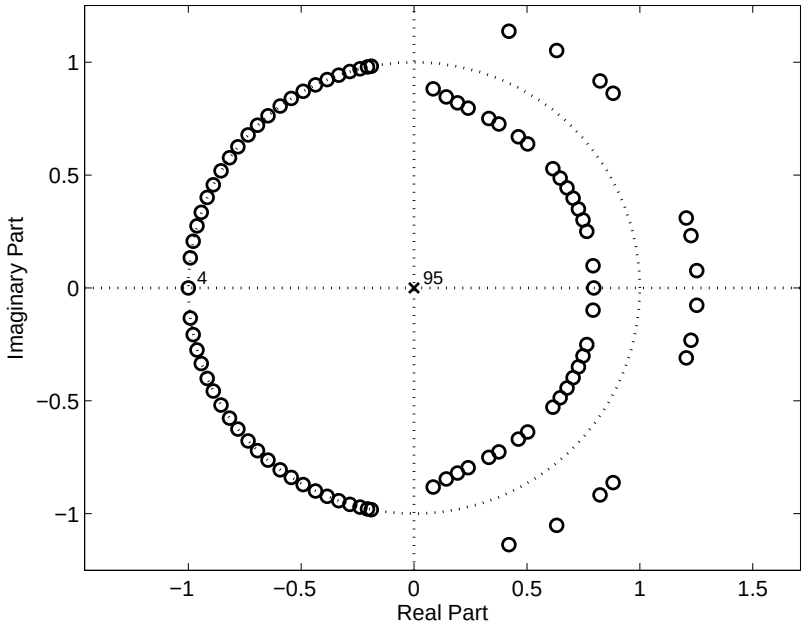

Figure 4.12: Zero-pole plot of locally optimal minimax filter with $N = 96$, $L = 3$ and $\omega_a = 0.56\pi$.

4.5 Comparisons with Other Existing Methods

In this section, we compare the global minimax design of filter banks with two existing methods that are well established in the literature.

4.5.1 Comparison with a Half-Band Filter Based on the Method in [41]

A simple and effective method for two-channel CQ filter banks is described in Sec. 5.3.6 of reference [41]. The method first designs a zero-phase lowpass FIR half-band filter $W(z)$ of order $2(N-1)$ by, e.g., the Parks-McClellan algorithm. One then defines $Y(z) = W(z) + \delta$ with δ the peak stopband ripple of $W(z)$ to ensure $Y(e^{j\omega}) \geq 0$ for all ω. With the same design specifications as in Example 5.3.2 of [41], namely filter length $N = 20$, stopband edge $\omega_a = 0.6\pi$, and number of VMs $L = 0$, Method 1 proposed in Sec. 4.3 was applied. The coefficients of the optimized lowpass CQ filter as well as the CQ filter obtained in Example 5.3.2 of [41] are listed in Table 4.4, and the maximum instantaneous energy η^2 in stopband and the largest magnitude error among all $N/2 + L$ equations in (4.1c) and (4.1d) for the two designs are given in Table 4.5. The magnitude responses of the two filters are depicted in Fig. 4.13. It is apparent that the proposed design technique is able to produce CQ filters with reduced instantaneous stopband energy and more accurate satisfaction of the PR conditions.

4.5.2 Comparison with the Method of Smith-Barnwell

The design technique under consideration is developed in [33] for tree-structured analysis/reconstruction systems and has since been a popular benchmark for performance evaluation and comparison as it also provides accurate numerical results of CQ filters of high quality. For comparison, method 1 proposed in Sec. 4.3 was applied to design minimax lowpass CQ filters of length 8, 16 and 32 to meet the requirement of 40 dB minimum stopband attenuation. The numbers of frequency grids for $N = 8,\ 16$ and 32 were set to $K = 25,\ 28$ and 50 respectively, and the number of vanishing moments was set to zero in all three designs. Minimax designs of CQ filters with the same filter lengths and stopband attenuation are also reported in [33]. The design results in terms of maximum instantaneous stopband energy η^2 and largest equation error in magnitude are shown in Tables 4.6, 4.7 and 4.8 for the CQ filters of length $N = 8,\ 16$

and 32 respectively. It is observed that in all three instances the proposed method produces designs with improved performance.

$H_0(z)$ of [41]	$H_0(z)$ from global design
0.1605476	0.151132584528507
0.4156381	0.406751138104326
0.4591917	0.465073716955923
0.1487153	0.164100745264147
−0.1642893	−0.159230874305372
−0.1245206	−0.132446162371893
0.08252419	0.077632712518187
0.08875733	0.092962310257929
−0.05080163	−0.047219604152222
−0.06084593	−0.062990546331313
0.03518087	0.032739512761500
0.03989182	0.040781157734971
−0.02561513	−0.023979520144301
−0.02440664	−0.024517125714218
0.01860065	0.017475746917452
0.01354778	0.013173547986633
−0.01308061	−0.012191078582176
−0.007449561	−0.006651596447548
0.01293440	0.011254662805676
−0.004995356	−0.004181786154910

Table 4.4: Coefficients of $H_0(z)$ of [41] and from global design

	$H_0(z)$ of [41]	$H_0(z)$ from global design
η^2	0.954568e-3	0.709881e-3
Largest equation error	5.591345e-7	<1e-15

Table 4.5: Filters performance comparison

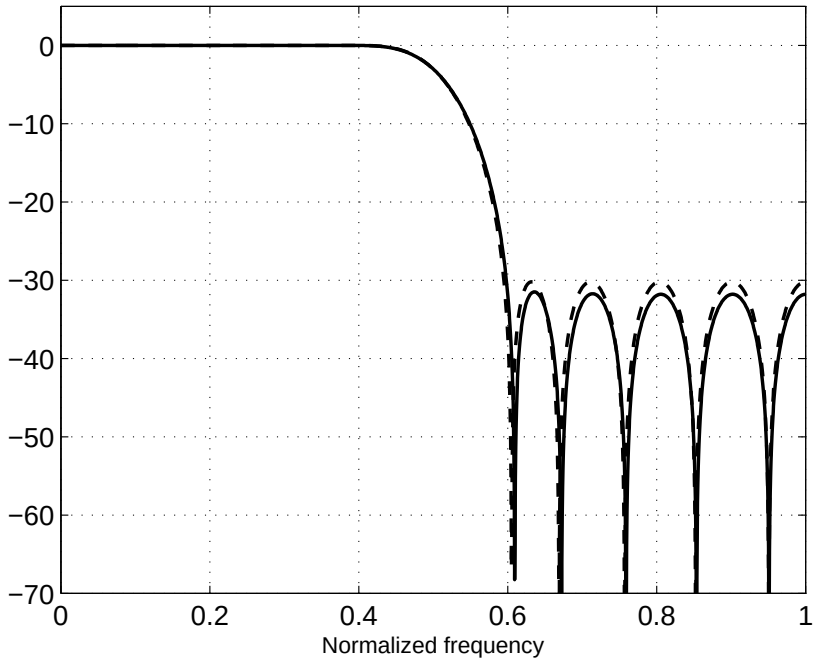

Figure 4.13: Magnitude response of minimax $H_0(z)$ with $N = 20$ from the global design (solid line) versus that of $H_0(z)$ in [41] (dashed line).

	$H_0(z)$ of [33]	$H_0(z)$ from global design
η^2	0.101105e-3	0.100221e-3
Largest equation error	8.316794e-8	<1e-15

Table 4.6: Filter length $N = 8$

	$H_0(z)$ of [33]	$H_0(z)$ from global design
η^2	0.101276e-3	0.099122e-3
Largest equation error	2.635550e-6	<1e-15

Table 4.7: Filter length $N = 16$

	$H_0(z)$ of [33]	$H_0(z)$ from global design
η^2	0.101805-3	0.098748-3
Largest equation error	2.162285-6	<1e-15

Table 4.8: Filter length $N = 32$

Chapter 5

Design of Orthogonal Cosine-Modulated Filter Banks

In Chapter 2, we have formulated the design of the prototype filter (PF) of an orthogonal cosine-modulated (OCM) filter bank as a constrained optimization problem

$$\begin{aligned} \text{minimize} \quad & e_2(\hat{\mathbf{h}}) = \hat{\mathbf{h}}^T \hat{\mathbf{P}} \hat{\mathbf{h}} && \text{(5.1a)} \\ \text{subject to:} \quad & a_{l,n}(\hat{\mathbf{h}}) = \hat{\mathbf{h}}^T \hat{\mathbf{Q}}_{l,n} \hat{\mathbf{h}} - c_n = 0 && \text{(5.1b)} \\ & \text{for } 0 \le n \le m-1 \text{ and } 0 \le l \le M/2-1 \end{aligned}$$

where $\hat{\mathbf{P}}, \hat{\mathbf{Q}}_{l,n}$ and c_n are calculated in Sec. 2.2. The impulse response $\mathbf{h}$ of the PF can then be obtained as

$$\mathbf{h} = \begin{bmatrix} \hat{\mathbf{h}} \\ \texttt{flipud}(\hat{\mathbf{h}}) \end{bmatrix} \tag{5.2}$$

This chapter is structured as follows. Sec. 5.1 introduces the method in finding a locally optimal solution to problem (5.1). Then, global design of low-order OCM filter banks using software for polynomial optimization problems (POPs) is studied in Sec. 5.2. In Sec. 5.3, we have made some observations among globally optimal low-order impulse responses of PFs, and have proposed an algorithm in carrying out potentially global design of high-order OCM filter banks. We conclude this chapter with several design examples and performance evaluation in Sec. 5.4.

5.1 Local Design of OCM Filter Banks

The minimization problem in (5.1) is a nonconvex problem with a quadratic objective function and a set of quadratic equality constraints. In this chapter, we describe a design technique similar in spirit to that proposed in [16], however here it is for PFs in OCM filter banks while [16] concerns two-channel orthogonal filter banks and wavelets.

Suppose that we are in the kth iteration to compute $\boldsymbol{\delta}$ so that $\hat{\mathbf{h}}_{k+1} = \hat{\mathbf{h}}_k + \boldsymbol{\delta}$ reduces the PF's stopband energy in (5.1a) and better satisfies the PR conditions in (5.1b). The objective function in (5.1a) then becomes

$$\hat{\mathbf{h}}_{k+1}^T \hat{\mathbf{P}} \hat{\mathbf{h}}_{k+1} = \boldsymbol{\delta}^T \hat{\mathbf{P}} \boldsymbol{\delta} + 2\boldsymbol{\delta}^T \hat{\mathbf{P}} \hat{\mathbf{h}}_k + \hat{\mathbf{h}}_k^T \hat{\mathbf{P}} \hat{\mathbf{h}}_k \tag{5.3}$$

For a $\boldsymbol{\delta}$ with $\|\boldsymbol{\delta}\|$ small, (5.1b) can be linearly approximated as

$$\begin{gathered} a_{l,n}(\hat{\mathbf{h}}_k + \boldsymbol{\delta}) \approx a_{l,n}(\hat{\mathbf{h}}_k) + \mathbf{g}_{l,n}^T(\hat{\mathbf{h}}_k)\boldsymbol{\delta} = 0 \\ \text{for } 0 \le n \le m-1 \text{ and } 0 \le l \le M/2-1 \end{gathered} \tag{5.4}$$

where $\mathbf{g}_{l,n}(\hat{\mathbf{h}}_k) = 2\hat{\mathbf{Q}}_{l,n}\hat{\mathbf{h}}_k$ is the gradient of $a_{l,n}(\hat{\mathbf{h}})$ at $\hat{\mathbf{h}}_k$. Using (5.3) and (5.4), the kth iteration of (5.1) assumes the form

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{\delta}^T \hat{\mathbf{P}} \boldsymbol{\delta} + \boldsymbol{\delta}^T \mathbf{b}_k && \text{(5.5a)} \\ \text{subject to:} \quad & \mathbf{G}_k \boldsymbol{\delta} = -\mathbf{a}_k && \text{(5.5b)} \\ & \|\boldsymbol{\delta}\| \text{ is small} && \text{(5.5c)} \end{aligned}$$

where $\mathbf{b}_k = 2\hat{\mathbf{P}}\hat{\mathbf{h}}_k$, $\mathbf{G}_k \in R^{N/4 \times N/2}$ collects the $N/4$ rows $\mathbf{g}_{l,n}^T(\hat{\mathbf{h}}_k)$ and $\mathbf{a}_k \in R^{N/4}$ consists of $N/4$ components $a_{l,n}(\hat{\mathbf{h}}_k)$. The equality constraint (5.5b) can be eliminated via the singular value decomposition of $\mathbf{G}_k = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ as [1]

$$\boldsymbol{\delta} = \mathbf{V}_e \boldsymbol{\phi} + \boldsymbol{\delta}_s \tag{5.6}$$

If $\mathbf{G}_k$ has full row-rank that is $N/4$, then $\boldsymbol{\phi}$ is an $N/4$-dimensional free vector, $\mathbf{V}_e$ is a matrix consisting of the last $N/4$ columns of $\mathbf{V}$, and $\boldsymbol{\delta}_s = -\mathbf{G}_k^\dagger \mathbf{a}_k$ where $\dagger$ denotes

the pseudo-inverse of a matrix. Thus, (5.5) can be cast as

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{\phi}^T \tilde{\mathbf{P}}_k \boldsymbol{\phi} + \boldsymbol{\phi}^T \tilde{\mathbf{b}}_k && \text{(5.7a)} \\ \text{subject to:} \quad & ||\boldsymbol{\phi}|| \text{ is small} && \text{(5.7b)} \end{aligned}$$

where $\tilde{\mathbf{P}}_k = \mathbf{V}_e^T \hat{\mathbf{P}} \mathbf{V}_e$, $\tilde{\mathbf{b}}_k = 2\mathbf{V}_e^T \hat{\mathbf{P}}(\hat{\mathbf{h}}_k + \boldsymbol{\delta}_s)$. Note that the constraint on the smallness of $\|\boldsymbol{\delta}\|$ is replaced by a constraint on the smallness of $\|\boldsymbol{\phi}\|$. This change is justified by noticing that (5.6) implies that $\|\boldsymbol{\delta}\| \leq \|\boldsymbol{\phi}\| + \|\boldsymbol{\delta}_s\|$ where $\boldsymbol{\delta}_s$ is a minimum-norm solution of (5.5b), hence a sufficiently small $\|\boldsymbol{\phi}\|$ implies a small $\|\boldsymbol{\delta}\|$.

Since (5.7) is a convex QP problem (see Sec. 2.4), its unique and global minimizer can be calculated efficiently using e.g. MATLAB command `quadprog`. Using the solution of (5.7) in Eq. (5.6), an optimal $\boldsymbol{\delta}_k$ is obtained and point $\hat{\mathbf{h}}_k$ is updated to point $\hat{\mathbf{h}}_{k+1}$. This iterative procedure continues until $\|\boldsymbol{\delta}_k\|$ is less than a prescribed tolerance and the local design is then complete.

It was found that occasionally the algorithm does not converge to a highly accurate solution because of numerical difficulties. In such a case, the filter coefficients obtained need to be slightly adjusted using a Gauss-Newton technique [1] with adaptively controlled weights in order for the final solution to satisfy the PR conditions in (5.1b) with a high degree of accuracy. The Gauss-Newton method with adaptively controlled weights was reviewed in Sec. 2.3.

5.2 Global Design of Low-Order OCM Filter Banks

It is not hard to see that problem (5.1) is a polynomial optimization problem (POP) since the objective function and constraints are all polynomials of degree two. Earlier, we have observed that (5.1) is a nonconvex problem. As a result, it admits local minimizers whose performance may be less than satisfactory. As introduced in previous chapters for the global design of CQ filter banks, GloptiPoly [9] and SparsePOP [43] are two popular solvers in finding global solutions of POPs of small sizes. For problem (5.1), GloptiPoly was found to work well. As an example, with $M = 2$, $m = 1$ and

$\rho = 1$, the toolbox produces two globally optimal impulse responses as

$$\mathbf{h}^{(2,1)} = \begin{bmatrix} \hat{\mathbf{h}}^{(2,1)} \\ \texttt{flipud}(\hat{\mathbf{h}}^{(2,1)}) \end{bmatrix} = \begin{bmatrix} 0.235923416966353 \\ 0.440840267366581 \\ 0.440840267366581 \\ 0.235923416966353 \end{bmatrix}$$

and $-\mathbf{h}^{(2,1)}$. Unfortunately, the usefulness of the software is limited only to OCM filter banks of low order. In our simulations, the software was found to work only for the following cases: a) $M = 2$, $1 \leq m \leq 5$; b) $M = 4$, $1 \leq m \leq 3$; c) $M = 6$, $m = 1$; d) $M = 8$, $m = 1$.

5.3 Potentially Global Design of High-Order OCM Filter Banks

Similar to the global design of high-order CQ filter banks, the principal idea we follow in this paper is to conduct a local search in a region that is sufficiently close to the global solution. To secure such a good starting (i.e., initial) point, we propose an order-recursive strategy based on the observation that the profile of the global solution of order $N = 2Mm$ is quite close to that of the global solution of order $N' = 2M'm'$ as long as the Euclidean distance between (m, M) and (m', M') remains small. The design is accomplished by

(i) obtaining a low-order global design;

(ii) using linear interpolation/zero-padding of the impulse response obtained to produce a desirable initial point for PF of slightly increased order and carrying out the design by a locally optimal method;

(iii) repeating step (ii) until the filter order reaches the targeted value.

In what follows, we describe the technical details that implement this design strategy.

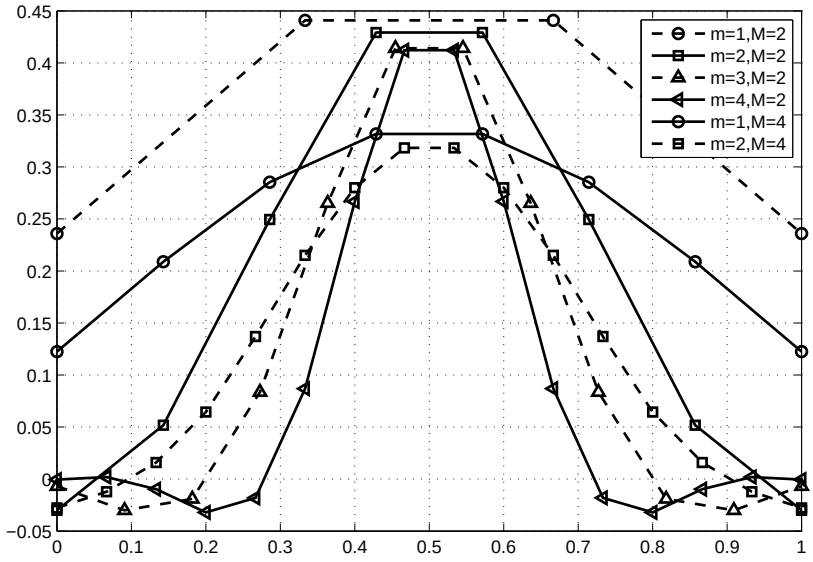

Figure 5.1: Pattern of impulse responses of globally optimal PFs.

First, we observe a common pattern among globally optimal impulse responses of low-order PFs obtained using GloptiPoly, as shown in Fig. 5.1, where the impulse responses are plotted over normalized interval $[0, 1]$. We note that

1. For a fixed M, the impulse responses with different m exhibit a similar pattern and are close to each other.

2. For $m = 1$, the impulse responses with different M also exhibit a similar shape.

These observations allow us to follow the design strategy described earlier. The main design steps can now be summarized as an algorithm below.

Algorithm 5.1
Input Data: Target filter order $\tilde{N} = 2\tilde{m}\tilde{M}$.
Step 1
Design a globally optimal PF with $m = 1$ and $M = 2$ using GloptiPoly. Denote the solution by $\mathbf{h} = \left[\begin{array}{cc} \hat{\mathbf{h}}^T & \texttt{flipud}(\hat{\mathbf{h}})^T \end{array} \right]^T$.
Step 2
Design a globally optimal PF with $m = 1$ and $M = \tilde{M}$ by the following steps:

Step 2.1

Set $M := M + 2$. Generate a vector $\hat{\mathbf{h}}_0^{\text{int}}$ of length-(mM) by linearly interpolating

$\hat{\mathbf{h}}$.

Step 2.2

Obtain a locally optimal design $\breve{\mathbf{h}}$ with $\hat{\mathbf{h}}_0^{\text{int}}$ as initial point, using the local design method described in Sec. 5.1. Set $\hat{\mathbf{h}} := \breve{\mathbf{h}}$. Denote the solution by $\mathbf{h} = \left[\begin{array}{cc} \hat{\mathbf{h}}^T & \texttt{flipud}(\hat{\mathbf{h}})^T \end{array} \right]^T$.

Step 2.3

If $M = \tilde{M}$, go to Step 3; otherwise, repeat from Step 2.1.

Step 3

Design a globally optimal PF with $m = \tilde{m}$ and $M = \tilde{M}$ by the following steps:

Step 3.1

Set $m := m + 1$. Generate a vector $\hat{\mathbf{h}}_0^{\text{zp}}$ of length $m\tilde{M}$ by padding zeros at the front of $\hat{\mathbf{h}}$.

Step 3.2

Obtain a locally optimal design $\breve{\mathbf{h}}$ with $\hat{\mathbf{h}}_0^{\text{zp}}$ as initial point (see Sec. 5.1). Set $\hat{\mathbf{h}} := \breve{\mathbf{h}}$. Denote the solution by $\mathbf{h} = \left[\begin{array}{cc} \hat{\mathbf{h}}^T & \texttt{flipud}(\hat{\mathbf{h}})^T \end{array} \right]^T$.

Step 3.3

If $m = \tilde{m}$, output $\mathbf{h}$ as the optimal design and terminate; otherwise, repeat from Step 3.1.

As an example, Fig. 5.2 illustrates Step 2.1 of the algorithm, where the impulse response of the globally optimal PF with $m = 1$ and $M = 2$ is linearly interpolated to yield a vector $\mathbf{h}_0^{\text{int}}$ (obtained by symmetrical extension of $\hat{\mathbf{h}}_0^{\text{int}}$) whose profile is similar to that of the globally optimal impulse response with $m = 1$ and $M = 4$. Another instance is shown in Fig. 5.3 to illustrate Step 3.1 of the algorithm where vector $\mathbf{h}_0^{\text{zp}}$ (obtained by symmetrical extension of $\hat{\mathbf{h}}_0^{\text{zp}}$) is produced by padding zeros to the globally optimal impulse response with $m = 1$ and $M = 4$. We see that $\mathbf{h}_0^{\text{zp}}$ is quite close to the globally optimal impulse response with $m = 2$ and $M = 4$.

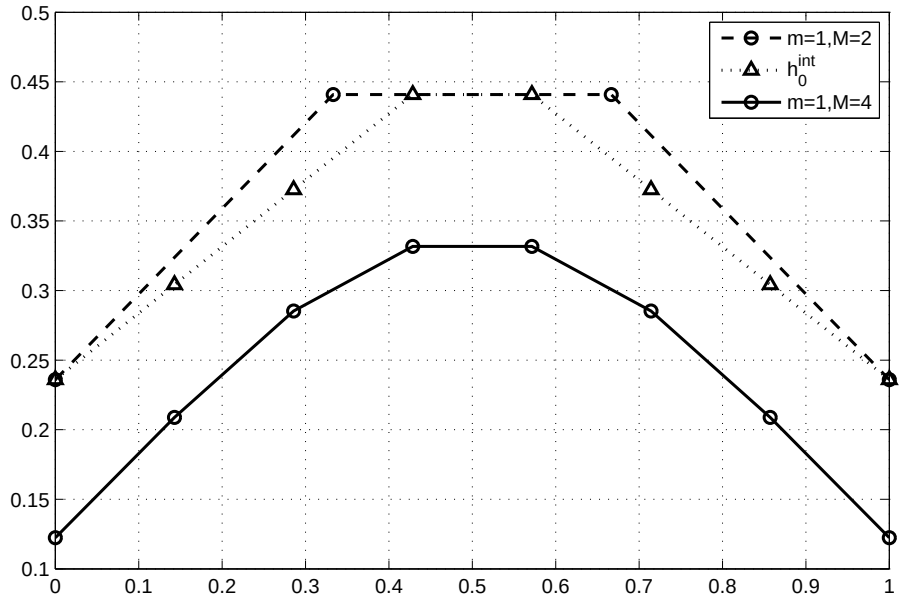

Figure 5.2: Effect of linear interpolation when $m = 1$.

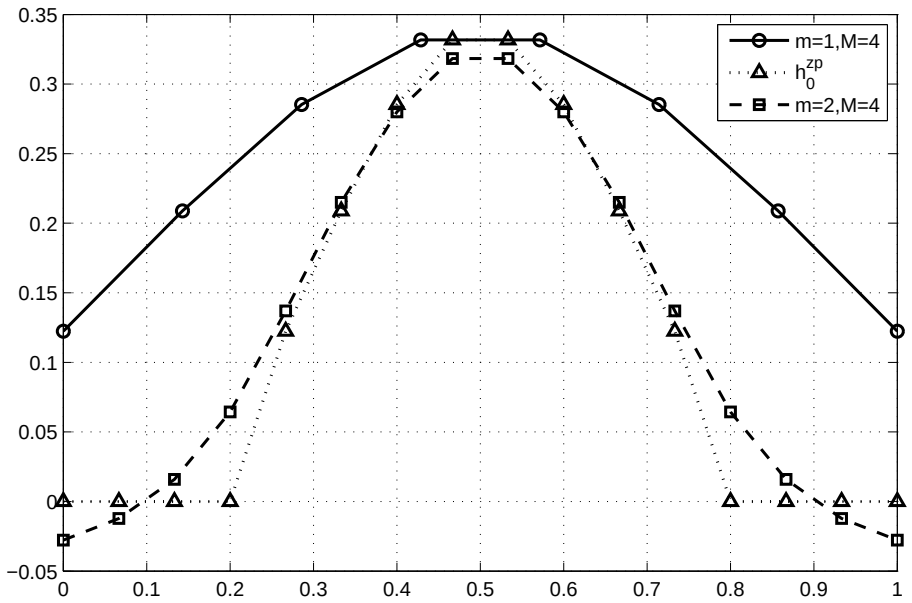

Figure 5.3: Effect of zero-padding when $M = 4$.

5.3.1 An improvement in initial point when $m = 1$

In Step 2.1 of Algorithm 5.1, an initial point is generated by linear interpolation for the design of an OCM filter bank with $m = 1$ and M increased by 2. Fig 5.2 illustrates the effect of linear interpolation as an example. Although the interpolated initial point appears close enough to the global solution point, we propose an improvement so that the initial point generated can be even closer to the global impulse response.

We observe from Fig. 5.2 that $\mathbf{h}_0^{\text{int}}$ exhibits a similar shape to the global impulse response with $m = 1$ and $M = 4$. Moreover, it appears that if all the values in $\mathbf{h}_0^{\text{int}}$ are downshifted by an appropriate constant value, $\mathbf{h}_0^{\text{int}}$ would become much closer to the global solution. Specifically, assume we have already obtained $\mathbf{h}_0^{\text{int}}$ from Step 2.1 of Algorithm 5.1, and we want to calculate a value d so that the vector $\mathbf{h}_0 = \mathbf{h}_0^{\text{int}} + d\mathbf{e}$ is sufficiently close to the global impulse resposne $\mathbf{h}$ of the same m and M ($\mathbf{e}$ is an all-one vector).

Apparently, the global solution $\mathbf{h}$ perfectly satisfies the set of constraints (5.1b). If an appropriate d is calculated so that $\mathbf{h}_0$ is sufficiently close to $\mathbf{h}$, then $\mathbf{h}_0$ is supposed to satisfy (5.1) at least approximately. Based on this, the following unconstrained optimization problem is formualted in order to obtain a good value of d:

$$\begin{aligned} \text{minimize} \quad & \sum_{l,n} a_{l,n}^2 (\hat{\mathbf{h}}_0^{\text{int}} + d\hat{\mathbf{e}}) \\ & \text{for } 0 \le n \le m-1 \text{ and } 0 \le l \le M/2 - 1 \end{aligned} \tag{5.8}$$

where $\hat{\mathbf{h}}_0^{\text{int}}$ and $\hat{\mathbf{e}}$ are related to $\mathbf{h}_0^{\text{int}}$ and $\mathbf{e}$ similarly as (5.2). In the case when $\hat{\mathbf{h}}_0^{\text{int}} + d\hat{\mathbf{e}}$ is sufficiently close to $\hat{\mathbf{h}}$, the minimum objective function value of (5.8) should be approximatly equal to zero. By comparing (5.8) with (2.22), it is not hard to see that the scalar variable d in problem (5.8) can be computed using the Gauss-Newton method with adaptively controlled weights.

To demonstrate the effectiveness of the above technique, we present an example for global design of a PF with $m = 1$ and $M = 4$. Fig. 5.4 illustrates the effect of downshifting the linearly interpolated impulse response $\mathbf{h}_0^{\text{int}}$ by a value d obtained by solving problem (5.8), where the downshifted impulse response is denoted as $\mathbf{h}_0$ and the global impulse response is denoted as $\mathbf{h}$. It is observed that after downshifting $\mathbf{h}_0^{\text{int}}$, the new impulse response $\mathbf{h}_0$ nearly overlaps with the global impulse response $\mathbf{h}$. In the following, as far as our simulations are concerned, the above technique was employed to calculate $\mathbf{h}_0$, and $\hat{\mathbf{h}}_0$ was used as the initial point instead of $\hat{\mathbf{h}}_0^{\text{int}}$ in Step

2.2 of Algorithm 5.1 for global design of PFs of OCM filter banks ($\hat{\mathbf{h}}_0$ is related to $\mathbf{h}_0$ similarly as (5.2)).

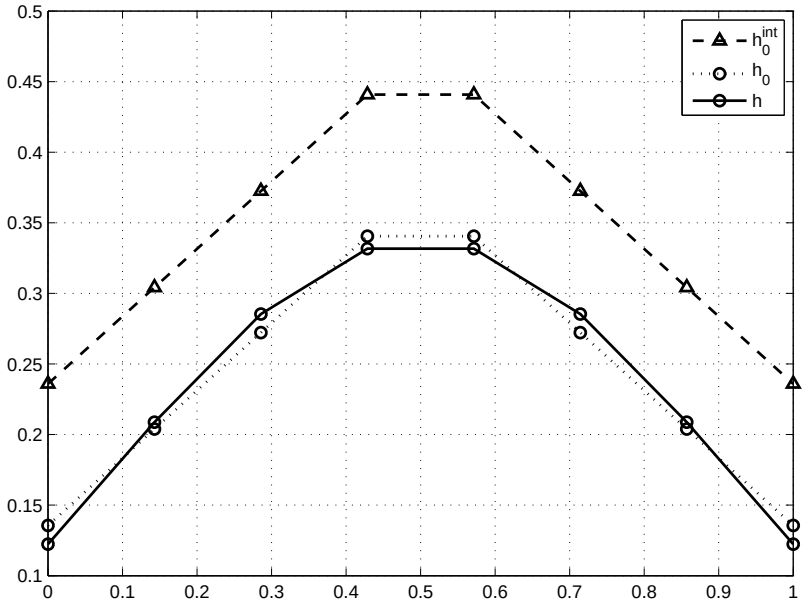

Figure 5.4: Effect of downshifting the linearly interpolated point when $m = 1$.

5.4 Design Examples and Performance Evaluation

5.4.1 Performance of the Proposed Method for Low-Order Designs

Algorithm 5.1 was applied to design low-order PFs for the following cases:

a) $M = 2$, $2 \leq m \leq 5$;

b) $M = 4$, $1 \leq m \leq 3$;

c) $M = 6$, $m = 1$;

d) $M = 8$, $m = 1$.

The impulse responses obtained were found to be practically identical to those generated from GloptiPoly in Sec. 5.2, which gives a convincing support to our speculation that Algorithm 5.1 produces globally optimal solution for problem (5.1).

5.4.2 Performance of the Proposed Method for High-Order Designs

Example 1:

The proposed method was also applied to design PFs of OCM filter banks with high order. As an example, an OCM filter bank with $m = 20$ and $M = 4$ was designed using Algorithm 5.1. The design was performed on a PC laptop with a 1.66 GHz dual-core processor. The magnitude response of the optimized PF is shown in Fig. 5.5 and the impulse response is shown in Fig. 5.6 where the coefficients are plotted over interval [0,1]. For comparison, an OCM filter bank with the same design specifications was designed by the local method described in Sec. 5.1 (the initial point for local design was calculated according to Eq. (35) of [17]). The convergence tolerance for both methods were set to $\varepsilon =$ 1e-12. The magnitude response and the impulse resposne of the locally optimal PF are shown in Figs. 5.7 and 5.8, respectively. It can be observed that the magnitude response of the locally optimal PF is not as good as that of the globally optimal PF because the oscillation in passband and stopband degrades the performance.

Table 5.1 lists the performance of the two filter banks in terms of stopband energy (see (5.1a)) of the optimized PFs and the largest equation error amongst all quadratic equality constraints in (5.1b). From the table, the superior performance of the filter bank produced by the proposed design technique is observed. We remark however that the performance gain was achieved at the cost of increased computational complexity: the proposed design required 35.6 seconds of CPU time versus 11.6 seconds required by the local method.

Table 5.1: Performance comparison for OCM filter banks with $m = 20$, $M = 4$ and $\rho = 1$.

	Global design	Local design
Energy in stopband	8.226e-13	6.585e-10
Largest equation error	1.839e-15	2.297e-10
CPU time	35.6s	11.6s

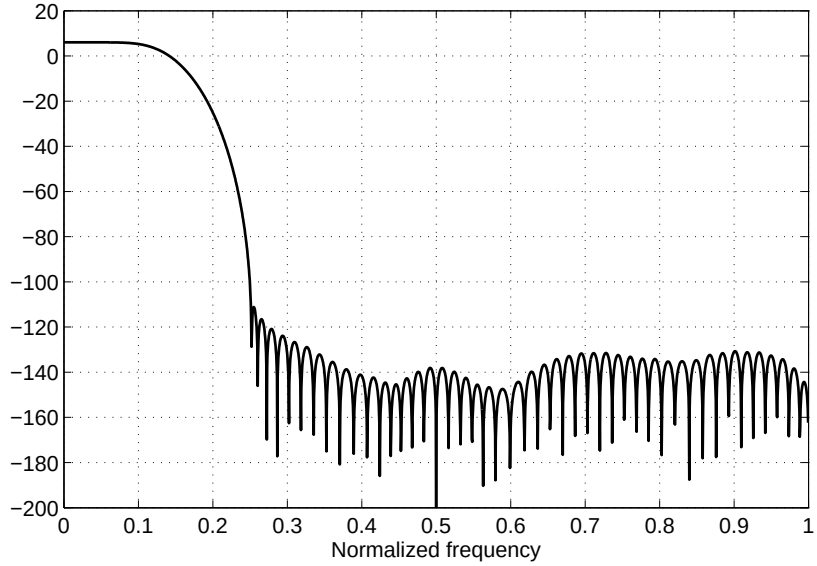

Figure 5.5: Magnitude response of globally optimal PF of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$.

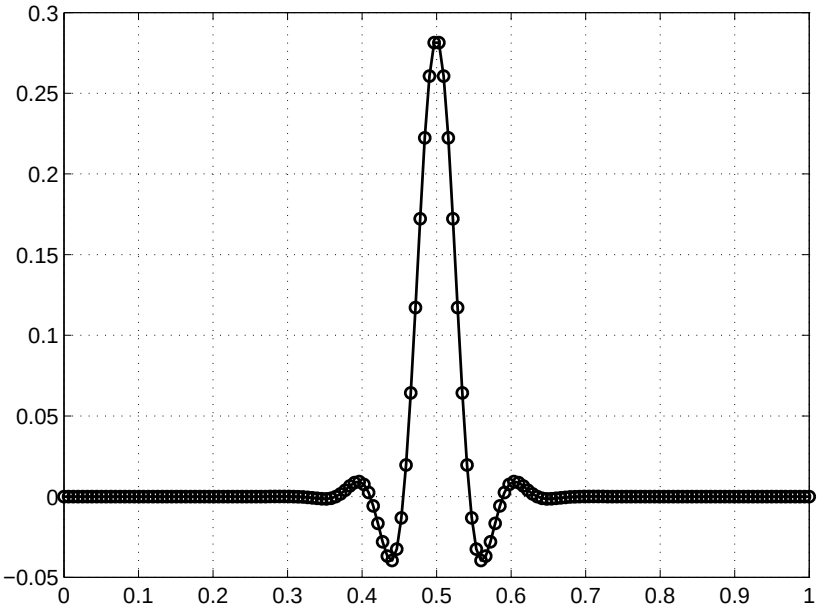

Figure 5.6: Impulse response of globally optimal PF of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$.

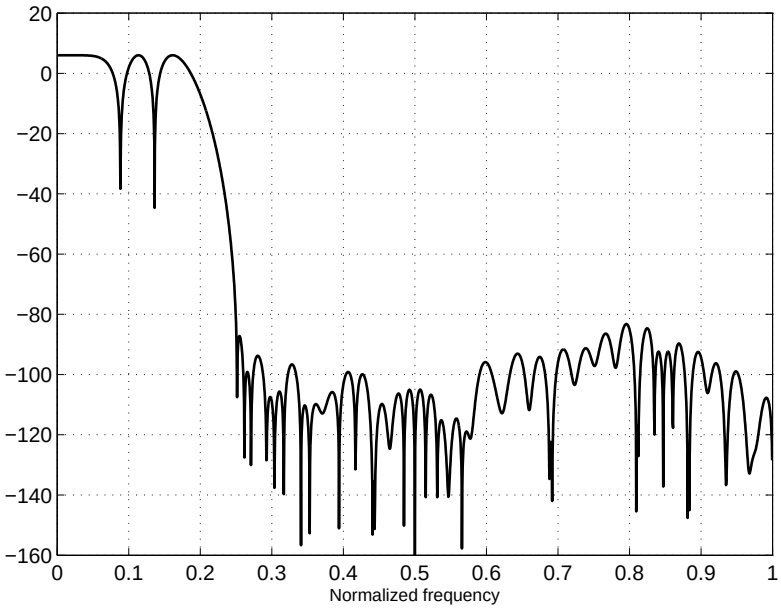

Figure 5.7: Magnitude response of locally optimal PF of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$.

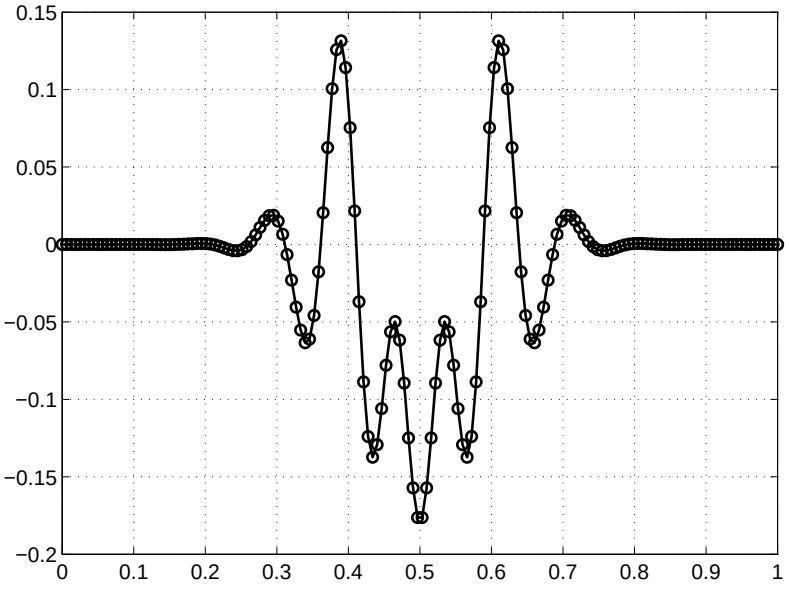

Figure 5.8: Impulse response of locally optimal PF of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$.

In addition, Fig. 5.9 illustrates magnitude responses of the $M = 4$ analysis filters (see Eq. (2.10a)) designed by the proposed method. The amplitude distortion, group-delay distortion and worst case aliasing error for the globally optimal OCM filter bank are calculated according to Eqs. (2.13)-(2.15) and plotted over the normalized frequency in Figs. 5.10-5.12. It is observed that the distortion and aliasing errors are sufficiently small, demonstrating that the OCM filter bank designed by the proposed method holds the PR property with a high degree of accuracy.

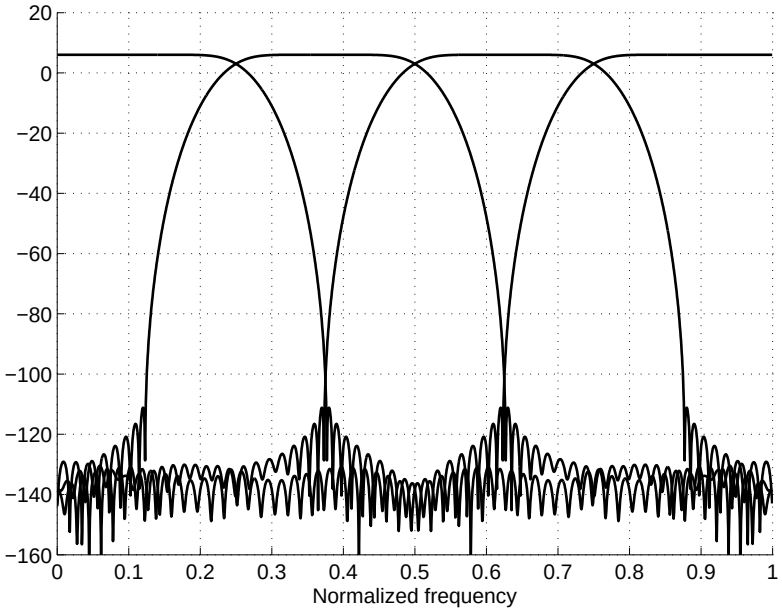

Figure 5.9: Magnitude responses of analysis filters of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$.

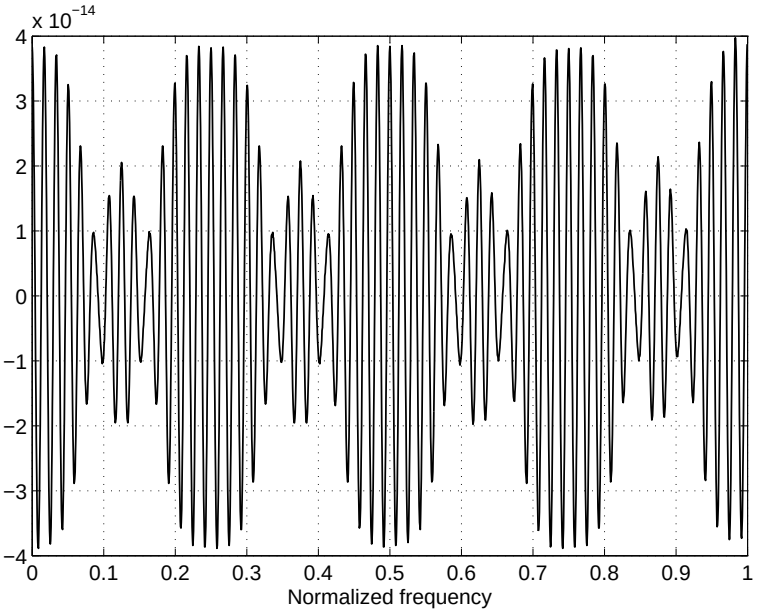

Figure 5.10: Amplitude distortion of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$.

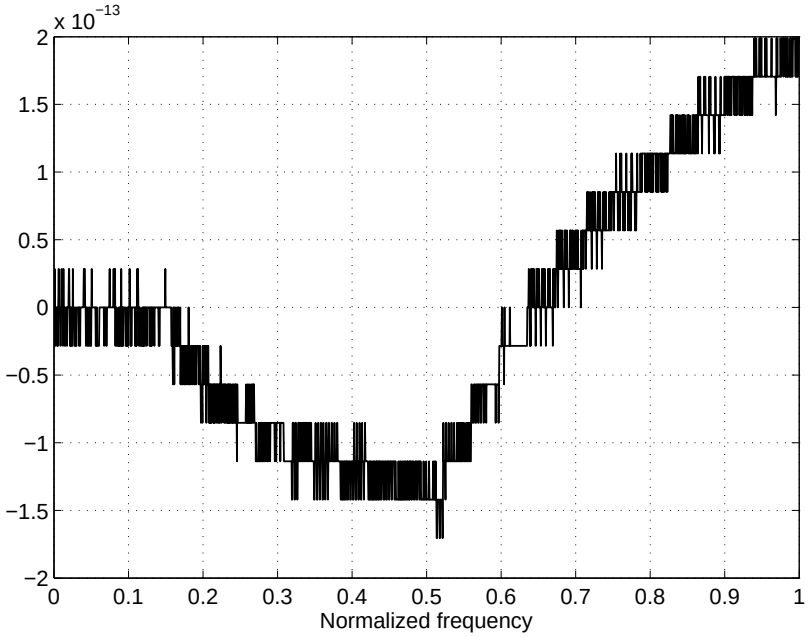

Figure 5.11: Group-delay distortion of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$.

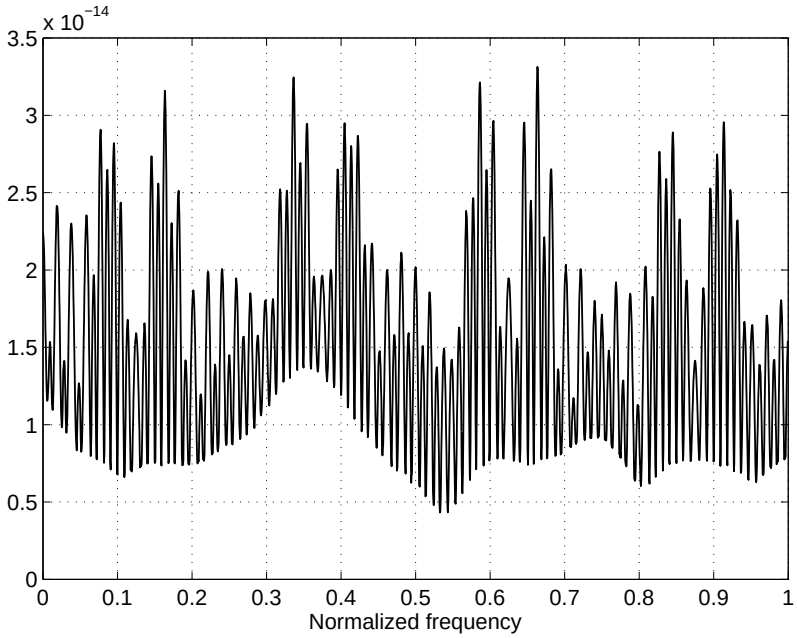

Figure 5.12: Worst case aliasing error of an OCM filter bank with $m = 20$, $M = 4$ and $\rho = 1$.

Example 2:

This example is concerned with the global design of an OCM filter bank with $m = 12$ and $M = 16$. Algorithm 5.1 was applied to design the PF, whose magnitude and impulse responses are depicted in Figs. 5.13 and 5.14, respectively. The convergence tolerance was set to $\varepsilon =$ 1e-12. For comparison, another OCM filter bank with the same specifications was designed using the local method described in Sec. 5.1. Table 5.2 compares the PF obtained from global design with that obtained by the local design in terms of stopband energy and largest PR equation error. It is observed that the proposed method yields a PF with smaller energy in stopband, at the cost of increased CPU time.

Table 5.2: Performance comparison for OCM filter banks with $m = 12$, $M = 16$ and $\rho = 1$.

	Global design	Local design
Energy in stopband	5.538e-10	8.269e-9
Largest equation error	2.806e-13	7.000e-15
CPU time	206.2s	68.7s

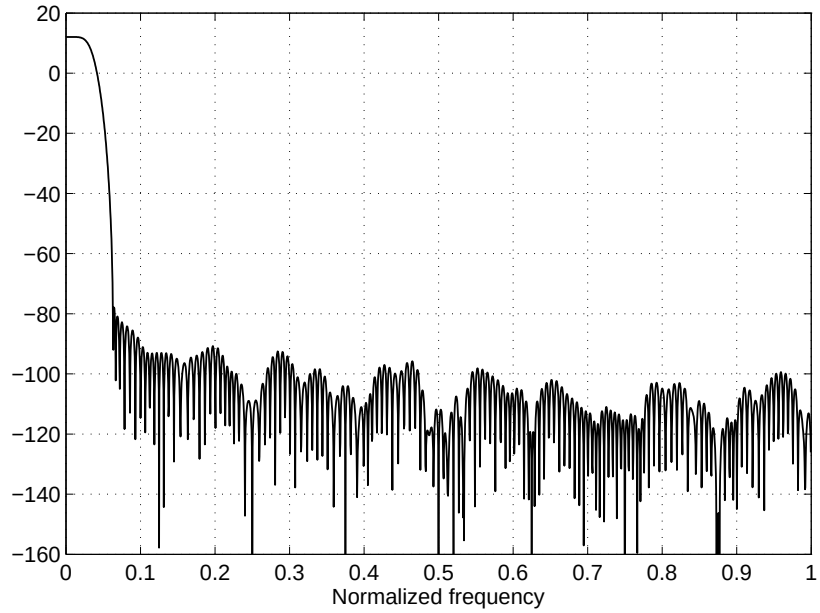

Figure 5.13: Magnitude response of globally optimal PF of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$.

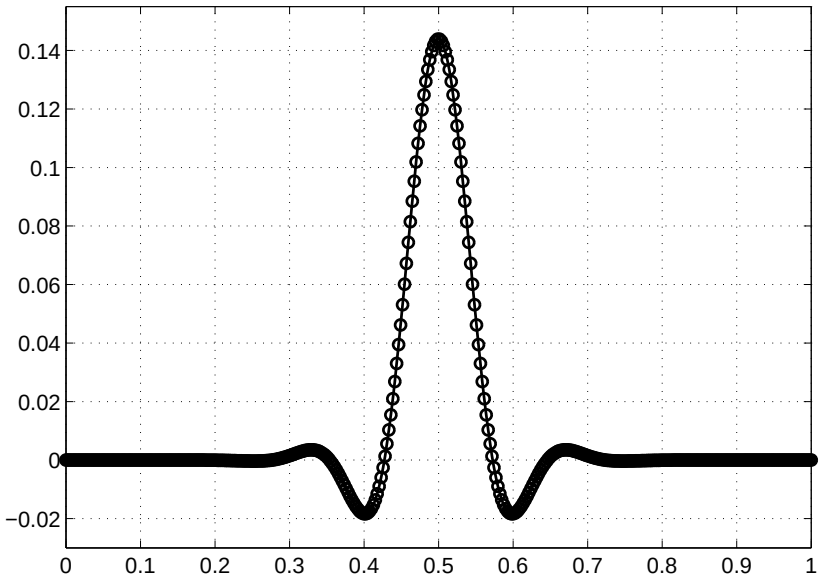

Figure 5.14: Impulse response of globally optimal PF of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$.

The magnitude responses of the $M = 16$ analysis filters are illustrated in Fig. 5.15. The performance of the globally optimal OCM filter bank designed from the proposed method was also evaluated in terms of the amplitude distortion (see Fig. 5.16), group-delay distortion (see Fig. 5.17), and worst case aliasing error (see Fig. 5.18). It can be seen from Fig. 5.18 that all amplitude responses of $T_l(e^{j\omega})$ are below 4.876e-12 (i.e., -226.24 dB), which is compared favorably with -144.61 dB achieved by the locally optimal OCM filter bank with the same design specifications as reported in [17].

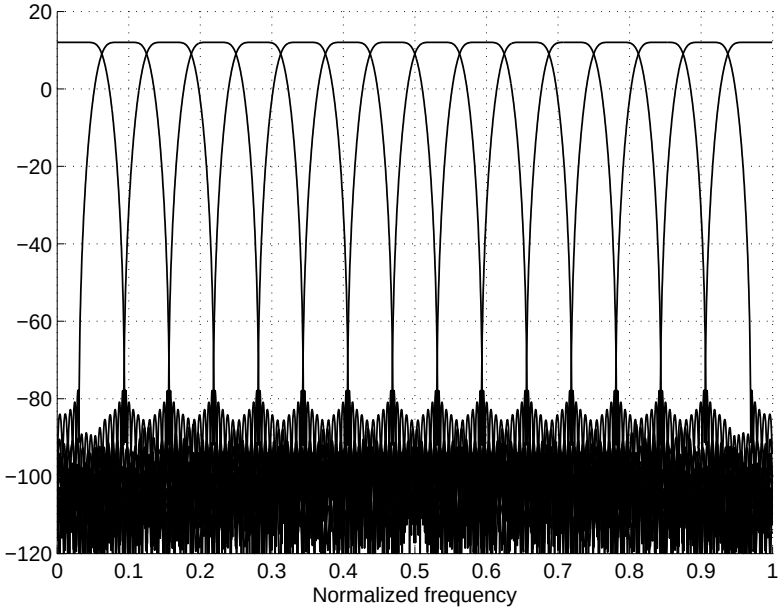

Figure 5.15: Magnitude responses of analysis filters of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$.

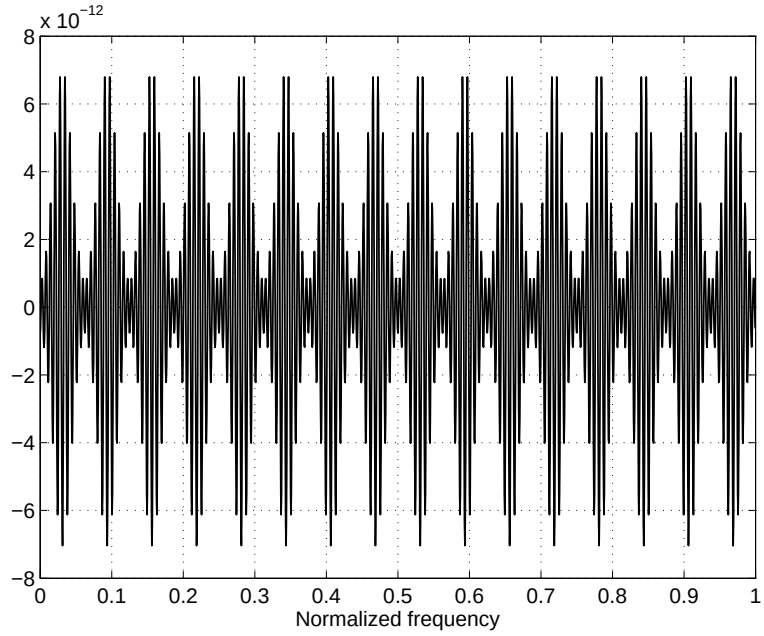

Figure 5.16: Amplitude distortion of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$.

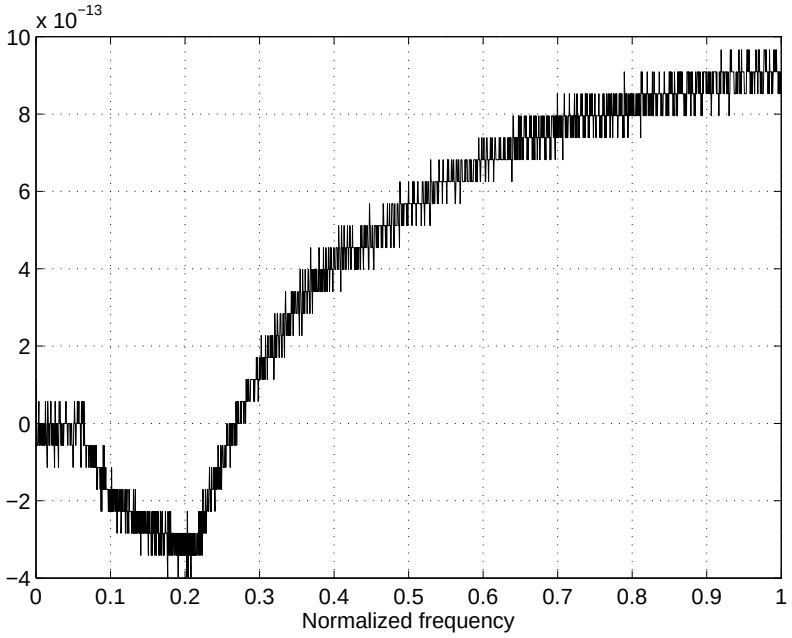

Figure 5.17: Group-delay distortion of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$.

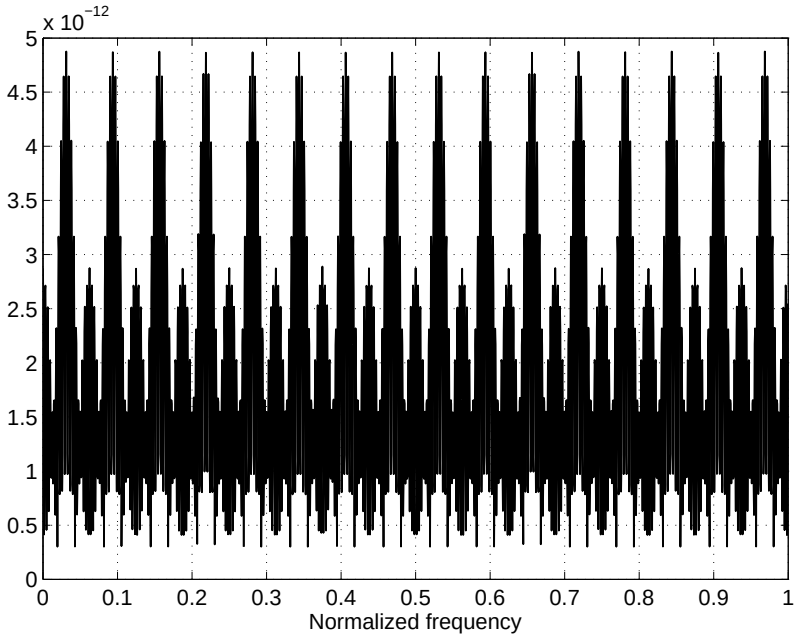

Figure 5.18: Worst case aliasing error of an OCM filter bank with $m = 12$, $M = 16$ and $\rho = 1$.

Example 3:
As the last set of example, global and local designs of OCM filter bank with $m = 7$ and $M = 32$ were implemented using Algorithm 5.1 and the local method reported in Sec. 5.1, respectively. Figs. 5.19-5.22 plot the magnitude and impulse responses for the globally and locally optimal PFs. Similarly, Table 5.3 compares the performance between the two designs. It was found that all these simulation results support the belief that the proposed method produces globally optimal OCM filter banks.

Besides, the magnitude responses of all the analysis filters are illustrated in Fig. 5.23. The amplitude distortion $e_m(\omega)$, group-delay distortion $e_{gd}(\omega)$ and worst case aliasing error $e_a(\omega)$ are also computed (the plots are omitted for this example) and were found sufficiently small. Specifically, the maximum values of $|e_m(\omega)|$ and $|e_a(\omega)|$ were calculated to be 8.59e-8 and 5.06e-8 respectively. By comparing the corresponding values of the near PR (NPR) locally optimal OCM filter bank of the same parameters, i.e., $m = 7$, $M = 32$ and $\rho = 1$ designed in [17], which are 1.09e-3 and 1.40e-7 respectively, it is apparent that the proposed global method offers designs with improved performance.

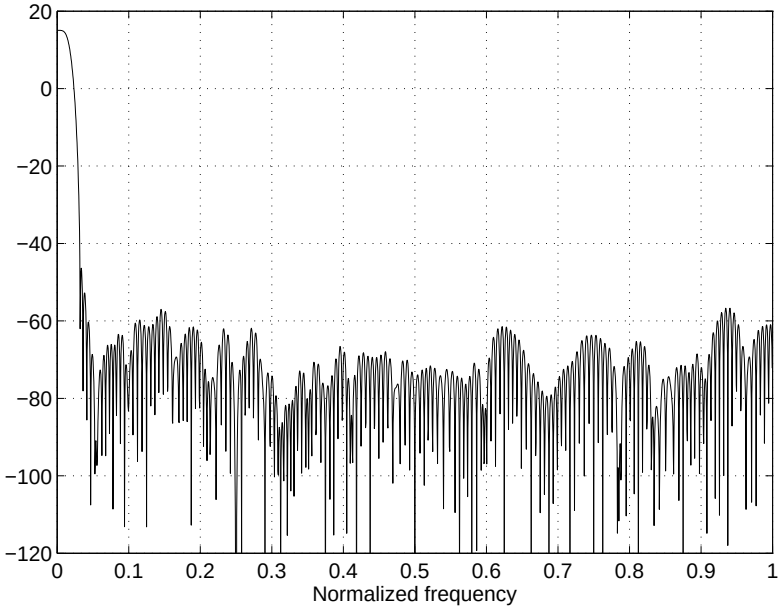

Figure 5.19: Magnitude response of globally optimal PF of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$.

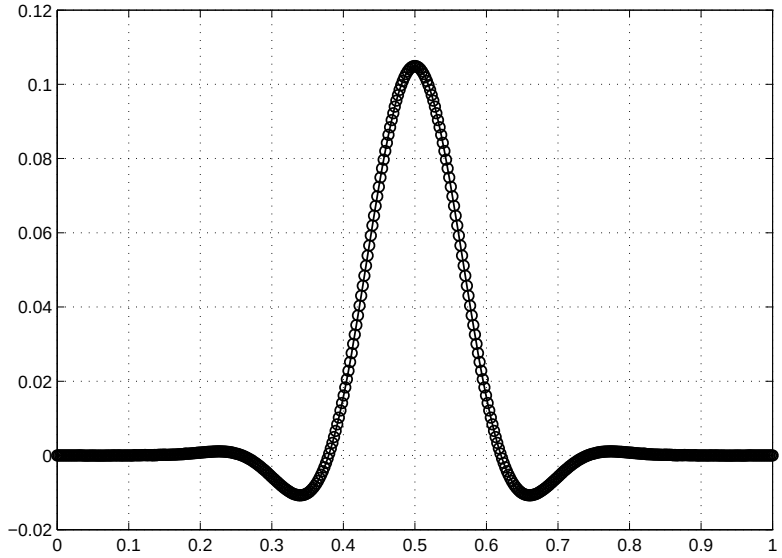

Figure 5.20: Impulse response of globally optimal PF of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$.

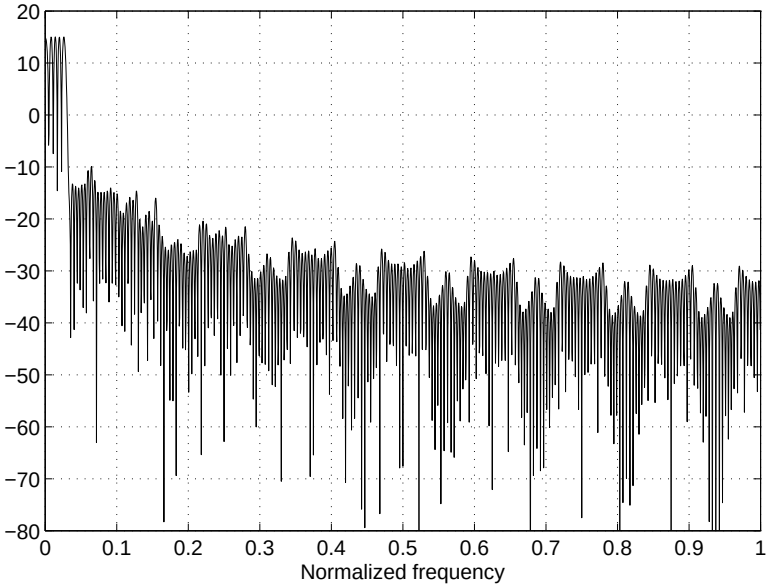

Figure 5.21: Magnitude response of locally optimal PF of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$.

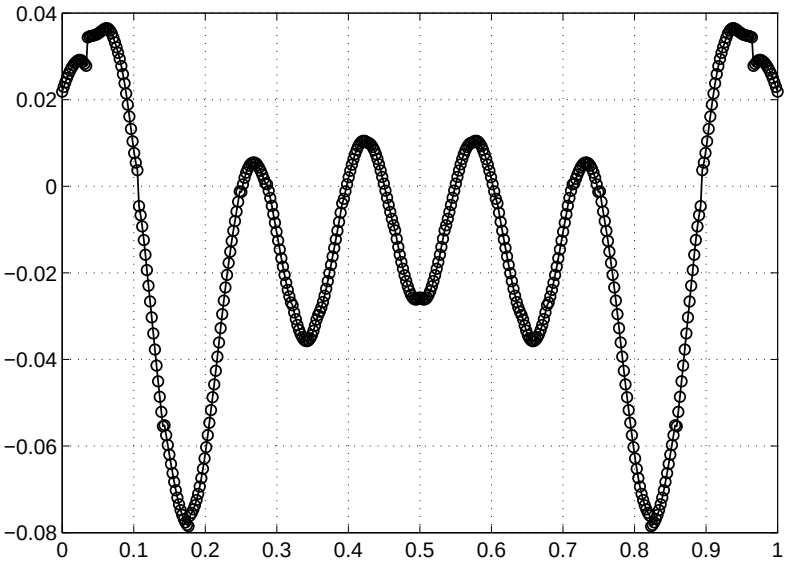

Figure 5.22: Impulse response of locally optimal PF of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$.

Table 5.3: Performance comparison for OCM filter banks with $m = 7$, $M = 32$ and $\rho = 1$.

	Global design	Local design
Energy in stopband	7.911e-7	9.676e-3
Largest equation error	2.232e-9	<1e-15
CPU time	180.1s	64.1s

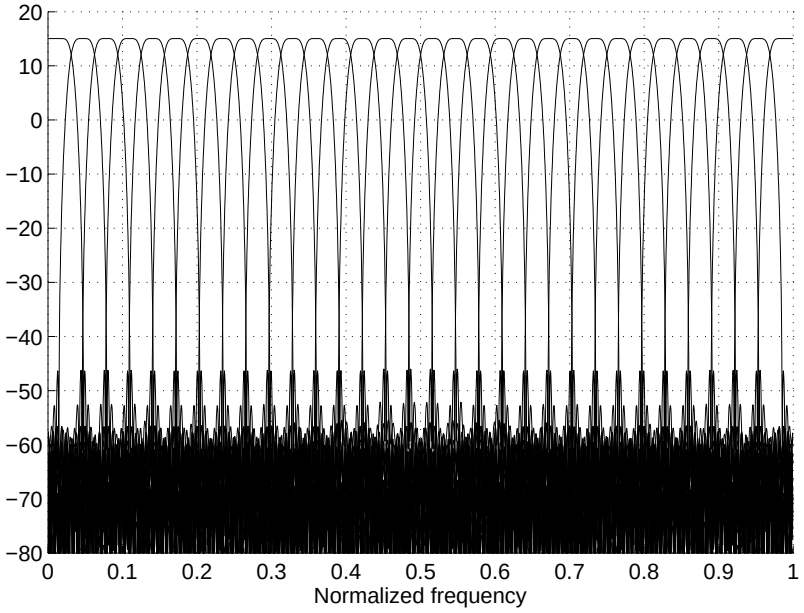

Figure 5.23: Magnitude responses of analysis filters of an OCM filter bank with $m = 7$, $M = 32$ and $\rho = 1$.

Chapter 6

Conclusions and Future Research

6.1 Conclusions

In this thesis, we have studied several design problems, namely, least squares (LS) design of orthogonal filter banks and wavelets, minimax design of orthogonal filter banks and wavelets, and design of orthogonal cosine-modulated (OCM) filter banks. In particular, we have proposed new methods for global designs of LS conjugate quadrature (CQ) filter banks, minimax CQ filter banks and OCM filter banks. The proposed methods are built on some recent progress in global optimization for polynommial optimization problems, the improved locally optimal design techniques for CQ filters and OCM filter banks, in conjunction with several critical observations on the globally optimal impulse responses of low-order CQ filter banks (as well as on the globally optimal impulse responses of low-order prototype filters of OCM filter banks). By using a zero-padding technique to generate the initial point, a progressive design procedure in terms of filter length has been implemented towards global designs of LS and minimax CQ filters. In addition, a similar order-recursive strategy combined with some techniques for identifying a desirable initial point in each round of the recursion has been proposed as an attempt towards global design of OCM filter banks. The work reported in this thesis can be summarized as follows.

For local design of LS CQ filter banks, the sequential convex-programming method and the sequential quadratic-programming method have been studied as an improvement of the direct design technique [16]. Experimental results have demonstrated that the improved methods produce a filter not only satisfying constraints with a higher precision, but also possessing a smaller stopband energy than the direct method.

Then, a method of global design of LS CQ filter banks has been investigated in detail. The zero-padding technique in generating an initial point to carry out a new round of local design has been shown to be superior relative to the interpolation technique reported in our earlier work [46]. Performance of the global LS CQ filters has been evaluated and compared with the local designs. All simulation studies conducted so far have supported our speculation that the proposed method does produce globally optimal CQ filters. Furthermore, it is found that the filters obtained from global designs possess minimum phase, which is a desirable property for filters to be of practical use.

In addition, we have made several improvements over the direct method for local minimax design of CQ filter banks for the algorithm to achieve convergence to solutions with high degree of accuracy. Simulation results are provided to demonstrate and compare the performance of the improved method with the original direct method. Subsequently, we have presented global minimax design of low-order and high-order CQ filter banks. In particular, two methods have been proposed to design globally optimal high-order minimax CQ filter banks. Performance of the proposed methods for global low-order and high-order designs of minimax CQ filters has been evaluated. We have also compared our methods with several existing methods that are well established in the literature. It is found that the proposed method produces minimax CQ filters with satisfactory performance. Like the LS designs, all minimax filters produced by the global method possess minimum phase.

We have also applied our design concept to the design of OCM filter banks. The local design method studied in this thesis is in spirit similar to that proposed in [16], and is in fact an improved version of the direct method presented in [17]. The global design method based on an order-recursive strategy has been made available with some techniques for generating a desirable initial point, wherein the zero-padding method and the interpolation method in conjuction with a downshifting technique based on the Gauss-Newton method with adaptively controlled weights have been utilized. The initial point produced in this way is found to be sufficiently close to the global solution point. Comparisons between performance of the OCM filter banks from global and local designs have been made. It is observed that the proposed global method offers designs with improved performance than the designs reported in [17].

6.2 Future Research

There are several problems that are pertinent to the work described in this thesis and appear to be worthwhile to pursue. As a matter of fact, our design idea for local and global designs of LS CQ filter banks, minimax CQ filter banks and OCM filter banks is quite general, and is also applicable to several other design scenarios. For instance, we can also minimize the maximum stopband energy of the prototype filter of an OCM filter bank such that local and global minimax designs of OCM filter banks can be implemented. We have also attempted to extend our design concept towards globally optimal biorthogonal cosine-modulated (BCM) filter banks. Unfortunately, a common pattern seems to not exist among impulse responses of globally optimal PFs of low-order BCM filter banks, which impedes us from following a similar strategy described in Algorithm 5.1. Consequently, global design of BCM filter banks remains to be an ongoing research problem. Besides, it is important to note that although our proposed methods towards global designs of serval types of filter banks considered in this thesis yield satisfactory filter banks that are compared favorably with several existing methods and are believed to be globally optimal, at the moment no rigorous analytical arguments are available to theoretically prove the global optimality of the designs obtained. The theoretical proof in question remains to be an interesting and challenging problem for future research.

Appendix A

Generating matrices $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}_{l,n}$ for problem (2.19)

As was mentioned in Sec. 2, the matrices $\mathbf{P}$ and $\mathbf{Q}_{l,n}$ need to be reduced from size $N \times N$ to $N/2 \times N/2$ so that problem (2.19) is formulated where the number of variables is reduced in half. In what follows, we present the derivations of $\hat{\mathbf{P}}$. Matrix $\hat{\mathbf{Q}}_{l,n}$ can be obtained in a similar way.

Matrix $\hat{\mathbf{P}}$ is required to satisfy the following equation

$$\mathbf{h}^T \mathbf{P} \mathbf{h} = \hat{\mathbf{h}}^T \hat{\mathbf{P}} \hat{\mathbf{h}} \tag{A.1}$$

We partition $\mathbf{P}$ into 4 blocks, each with size $N/2 \times N/2$, as

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_A & \mathbf{P}_B \\ \mathbf{P}_C & \mathbf{P}_D \end{bmatrix} \tag{A.2}$$

and use Eq. (2.20) to write

$$\begin{aligned} \mathbf{h}^T \mathbf{P} \mathbf{h} &= \begin{bmatrix} \hat{\mathbf{h}}^T & \texttt{flipud}(\hat{\mathbf{h}})^T \end{bmatrix} \begin{bmatrix} \mathbf{P}_A & \mathbf{P}_B \\ \mathbf{P}_C & \mathbf{P}_D \end{bmatrix} \begin{bmatrix} \hat{\mathbf{h}} \\ \texttt{flipud}(\hat{\mathbf{h}}) \end{bmatrix} \\ &= \hat{\mathbf{h}}^T \mathbf{P}_A \hat{\mathbf{h}} + \hat{\mathbf{h}}^T \mathbf{P}_B \texttt{flipud}(\hat{\mathbf{h}}) + \texttt{flipud}(\hat{\mathbf{h}})^T \mathbf{P}_C \hat{\mathbf{h}} + \texttt{flipud}(\hat{\mathbf{h}})^T \mathbf{P}_D \texttt{flipud}(\hat{\mathbf{h}}) \\ &= \hat{\mathbf{h}}^T [\mathbf{P}_A + \texttt{fliplr}(\mathbf{P}_B) + \texttt{flipud}(\mathbf{P}_C) + \texttt{fliplr}(\texttt{flipud}(\mathbf{P}_D))] \hat{\mathbf{h}} \end{aligned}$$

where `fliplr` denotes flipping a vector or a matrix left to right, and `flipud` denotes

flipping a vector or a matrix up to down. Thus, matrix $\hat{\mathbf{P}}$ can be obtained as

$$\hat{\mathbf{P}} = \mathbf{P}_A + \texttt{fliplr}(\mathbf{P}_B) + \texttt{flipud}(\mathbf{P}_C) + \texttt{fliplr}(\texttt{flipud}(\mathbf{P}_D))$$

Finally, $\hat{\mathbf{P}}$ can be made symmetric by

$$\hat{\mathbf{P}} := \frac{1}{2}(\hat{\mathbf{P}} + \hat{\mathbf{P}}^T)$$

It can easily be verified that matrix $\hat{\mathbf{P}}$ obtained in this way is a positive-definite matrix.

In a similar manner, matrix $\hat{\mathbf{Q}}_{l,n}$ can be generated and made symmetric.

Bibliography

[1] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*, Springer, 2007.

[2] A. Antoniou, *Digital Signal Processing: Signals, Systems, and Filters*, McGraw-Hill, 2005.

[3] R. Bregović and T. Saramäki, "An efficient approach for designing nearly perfect-reconstruction low-delay cosine-modulated filter banks," *IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 825-828, Scottsdale AZ., May 2002.

[4] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*, Philadelphia, PA: SIAM, 2001.

[5] C. D. Creusere and S. K. Mitra, "A simple method for designing high-quality prototype filters for M-band pseudo QMF banks," *IEEE Trans. Signal Processing*, vol. 43, pp. 10051007, Apr 1995.

[6] B. Dumitrescu and C. Popeea, "A low complexity SDP method for designing optimum compaction filters," *Proc. ICASSP*, pp. 516-519, 2000.

[7] R. Fletcher, *Pratical Methods of Optimization*, 2nd ed., Wiley, New York, 1987.

[8] P. E. Gill, W. Murray, and M. A. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization," *SIAM Review*, vol. 47, no. 1, pp. 99-131, 2005.

[9] D. Henrion, J. B. Lasserre, and J. Lofberg, "GloptiPoly 3: Moments, optimization and semidefinite programming," *LAAS-CNRS Report*, Toulouse, France Sep 2008.

[10] P. N. Heller, T. Karp, and T. Q. Nguyen, "A general formulation of modulated filter banks," *IEEE Trans. Signal Processing*, vol. 47, pp. 986-1002, Apr. 1999.

[11] T. Karp, A. Mertins, and G. Schuller, "Efficient biorthogonal cosine-modulated filter banks," *Signal Processing*, vol. 81, pp. 997-1016, May. 2001.

[12] R. D. Koilpillai and P. P. Vaidyanathan, "Cosine-modulated FIR filter banks satisfying perfect reconstruction," *IEEE Trans. Signal Processing*, vol. 40, pp. 770-783, April 1992.

[13] T. Kailath, A. Sayed, B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ: Prentice Hall, 2000.

[14] A. Karmakar, A. Kumar and R. K. Patney, "Design of an optimal two-channel orthogonal filterbank using semidefinite programming," *IEEE Signal Processing Letters*, vol.14, no.10, pp.692-694, Oct 2007.

[15] W.-S. Lu, *Digital Signal Processing III*, Lecture Notes, University of Victoria, 2007.

[16] W.-S. Lu and T. Hinamoto, "Direct design of orthogonal filter banks and wavelets," *ISCAS* 2009, pp. 81-84, Taipei, Taiwan, May 24-27, 2009.

[17] W.-S. Lu, T. Saramäki and R. Bregović, "Design of practically perfect-reconstruction cosine-modulated filter banks: a second-order cone programming approach," *IEEE Trans. Signal Processing*, vol. 51, pp. 552-563, Mar. 2004.

[18] W.-S. Lu, R. Bregović, and T. Saramäki, "Efficient design of perfect-reconstruction biorthogonal cosine-modulated filter banks using convex Lagrangian relaxation and alternating null-space projections," *ISCAS* 2003, vol.4, pp. IV-141- IV-144, 2003.

[19] W.-S. Lu, "Digital filter design: Global solutions via polynomial optimization," *APCCAS* 2006, pp. 49-52, Singapore, Dec 2006.

[20] Y.-P. Lin and P. P. Vaidyanathan, "Linear phase cosine modulated maximally decimated filter banks with perfect reconstruction", *IEEE Trans. Signal Processing*, vol. 43, pp. 2525-2539, Nov 1995.

[21] Y.-P. Lin and P. P. Vaidyanathan, "A Kaiser window approach for the design of prototype filters of cosine modulated filterbanks," *IEEE Signal Processing Lett.*, vol. 5, pp. 132134, June 1998.

[22] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optimization*, vol. 11, no. 3, pp. 796-817, 2001.

[23] J. B. Lasserre, "A semidefinite programming approach to the generalized problem of moments," *Mathematical Programming*, 112:65-92, 2008.

[24] W. Lawton and C.A. Michelli, "Design of conjugate quadrature filters having specified zeros," *Proc. ICASSP*, vol. 3, pp. 2069-2072, 1997.

[25] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Application*, vol. 248, pp. 193-228, 1998.

[26] R. D. C. Monteiro and T. Tsuchiya, "Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions," *Math. Programming*, vol. 88, pp. 6183, 2000.

[27] T. Q. Nguyen and R. D. Koilpillai, "The theory and design of arbitrary-length cosine-modulated filter banks and wavelets, satisfying perfect reconstruction," *IEEE Trans. Signal Processing*, vol. 44, pp. 473-483, Mar. 1996.

[28] T. Q. Nguyen, "A quadratic-constrained least-squares approach to the design of digital filter banks," *IEEE Int. Symp. Circuits Syst.*, San Diego, CA., May 1992.

[29] T. Q. Nguyen and P. N. Heller, "Biorthogonal cosine-modulated filter banks," *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1471-1474, Atlanta, CA, May 1996.

[30] T. Q. Nguyen, "Near-perfect-reconstruction pseudo-QMF banks," *IEEE Trans. Signal Processing*, vol. 42, pp. 65-76, Jan 1994.

[31] D. Pinchon, P. Siohan and C. Siclet, "Design techniques for orthogonal modulated filterbanks based on a compact representation," *Signal Processing, IEEE Transactions on*, vol.52, no.6, pp. 1682- 1692, June 2004.

[32] O. Rioul and P. Duhamel, "A Remez exchange algorithm for orthonormal wavelets," *IEEE Trans. Circuits Syst.*, II, vol. 41, pp. 550-560, Aug. 1994.

[33] M.J.T. Smith and T.P. Barnwell III, "Exact reconstruction techniques for tree-structured subband coders," *IEEE Trans. Signal Processing*, vol. 34, pp. 434-441, June 1986.

[34] T. Saramäki, "Designing prototype filters for perfect-reconstruction cosine-modulated filter banks," *IEEE Int. Symp. Circuits Syst.*, vol. 3, pp. 1605-1608, San Diego, May 1992.

[35] T. Saramäki and R. Bregović, "Multirate systems and filter banks," in *Multirate Systems: Design and Applications*, G. Jovanovic-Dolecek, Ed. Hershey, PA: Idea Group, 2002, ch. 2.

[36] G. Schuller, "New factorization and structure for cosine modulated filter banks with variable system delay," in *Proc. 30th Asilomar Conf. Signals, Systems, and Computers*, vol. 2, pp. 1310-1314, Pacific Grove, CA, Nov 1996.

[37] G. Schuller and T. Karp, "Modulated filter banks with arbitrary system delay: Efficient implementations and the time-varying case," *IEEE Trans. Signal Processing*, vol. 48, pp. 734-748, Mar 2000.

[38] D.B.H. Tay, "Zero-pinning the Bernstein polynomial: A simple design technique for orthogonal wavelets," *IEEE Signal Processing Letters*, vol. 12, no. 12, pp. 835-838, Dec. 2005.

[39] D.B.H. Tay, "Least squares design of orthogonal wavelets via the zeropinning technique," *Proc. ICASSP*, vol. 3, pp. 213-216, 2006.

[40] J. Tuqan and P.P. Vaidyanathan, "Globally optimal two channel FIR orthonormal filter banks adapted to the input signal statistics," *Proc. ICASSP*, pp. 1353-1356, 1998.

[41] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, 1993.

[42] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, pp. 49-95, 1996.

[43] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "SparsePOP: A sparse semidefinite programming relaxation of polynomial optimization problems," *Research Report*, Tokyo Institute of Technology, March 2005.

[44] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sum of squares and semidefinite programming relaxations for polynomial optimization problem with structured sparsity," *SIAM J. Optimization* 17 (2006) 218-242.

[45] H. Waki, S. Kim, M. Kojima, M. Muramatsu, and H. Sugimoto, "User manunal for SparsePOP: A sparse semidefinite programming relaxation of polynomial optimization problems," *Research Report*, Tokyo Institute of Technology, March 2005.

[46] J. Yan and W.-S. Lu, "Towards global design of orthogonal filter banks and wavelets," *PacRim* 2009, pp. 187-192, Victoria, BC, Aug 2009.

[47] J. Yan and W.-S. Lu, "Towards global design of orthogonal filter banks and wavelets," to be published in *Canadian J. Elec. Comp. Eng.* 2010.

[48] J. Yan and W.-S. Lu, "Global design of perfect-reconstruction orthogonal cosine-modulated filter banks," accepted for presentation at 2010 *CCECE*, Calgary, May 2010.

[49] Z.-J. Zhang, "Efficient design of cosine modulated milter banks based on gradient information," *Signal Processing Letters, IEEE*, vol.14, no.12, pp.940-943, Dec 2007.

[50] Optimization Toolbox V3.0.4 (R2006a) for MATLAB, The MathWorks, Natick, MA.

[51] SeDuMi version 1.1R3, available at `http://sedumi.mcmaster.ca/`.