Design of Finite Impulse Response Filters with Reduced Group Delays and Active/Robust Array Beamformers

Liu Yongzhi

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University in fulfillment of the requirement for the degree of Doctor of Philosophy

2007
Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me and has not been submitted for a higher degree to any other University or Institute.

Date

Liu Yongzhi
Acknowledgements

This research work received both financial and technical supports from the School of Electrical & Electronic Engineering, Nanyang Technological University.

First and foremost, I would like to thank Associate Prof. Zhiping Lin, my supervisor, from the bottom of my heart for his professional guidance, patience and encouragement throughout my research work. I have learned not only problem modelling and analytical capability, but also professional ethics from him, which I greatly appreciate.

I would also like to take this opportunity to thank Prof. Wu-Sheng Lu for the helpful discussion and for providing his design data for Example 3.3 in Chapter 3. Thanks go to Mr. Qiyue Zou for the extensive discussion and collaboration concerning the work presented in Chapter 5.

Special thanks are given to the laboratory technicians and my fellow lab-mates in the Information System Research Lab for making it a friendly and conducive working environment.

Last but not least, I would like to thank my lovely wife, Ms. Yin Mu, and our parents whose love and support have always been the greatest inspiration for me.
Summary

This thesis presents research work on two related subjects: finite impulse response (FIR) filter design and array beamforming. The focus is on the design of FIR filters with reduced passband group delay errors and efficient/robust array beamforming.

We first investigate the design of FIR filters with reduced passband group delays, which are very important in some applications. An analytic expression for the group delay of an FIR filter is derived, which shows that it is a nonlinear function with respect to the filter coefficients. Thus it is difficult to incorporate an additional group delay constraint into conventional FIR filter design methods, such as the Park-McClellan algorithm, the least-squares method and the eigenfilter approach. To meet the specifications in both magnitude response and group delay, we propose a new design method based on semidefinite programming (SDP). A group delay constraint is formulated as a linear matrix inequality (LMI) with some reasonable approximations. The group delay and magnitude constraints are then successfully incorporated into the SDP framework, leading to a new method that is particularly useful for designing band-selective FIR filters with reduced group delay errors.

The group delay performance is further investigated in a computationally
more efficient but structurally more complicated filter, namely, the frequency-
response masking (FRM) filter. Analytic expressions for the group delay of a
basic and a two-stage FRM filters are derived. Their gradients with respect to
the subfilter coefficients are obtained subsequently. In designing FRM filters
with reduced passband group delays, the coefficients of all subfilters are stacked
in a vector and optimized by a series of gradient-based linear updates, with each
update carried out by efficient second-order cone programming (SOCP). With
the magnitude constraint and the proposed group delay constraint, the group
delay error can be effectively reduced at the cost of slight increase in magnitude
ripples.

By analogy between temporal filtering by FIR filters and spatial filtering
by sensor arrays, the feasibility of applications of the FRM technique in array
beamforming is investigated in detail. On one hand, it is shown that there
is a limitation in applying the FRM technique in passive array beamforming.
On the other hand, for active array beamforming, a novel combination of the
concept of effective aperture and the FRM technique does lead to synthesis of
desirable beamformers, which have effective beampatterns with sharp transition
bands and low sidelobes, and can be implemented with fewer sensors compared
with other existing methods.

Unlike conventional beamformers which may use predetermined FIR filters
as their aperture weighting functions, adaptive beamformers can adjust the
beampatterns to maintain the prescribed frequency responses in the desired di-
rections while introducing nulls in the interference directions. The performance
of the general sidelobe cancellers (GSCs) in the presence of direction of arrival
(DOA) mismatch is investigated. To improve the robustness of GSCs against
DOA mismatch, a class of leakage constraints is derived by exploiting the statistical property of the leakage signals. Compared with the class of derivative constraints, fewer constraints are imposed.
Contents

Acknowledgements i

Summary ii

List of Figures xiv

1 Introduction 1

1.1 Background and Motivations . . . . . . . . . . . . . . . . . . . 1
1.2 Objectives . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
1.3 Major Contributions . . . . . . . . . . . . . . . . . . . . . . . . 8
1.4 Organizations . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9

List of Tables 1

2 FIR Filter Design with Group Delay Constraint 11

2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
2.2 Convex Optimization and SDP . . . . . . . . . . . . . . . . . . . 14
2.3 Group Delay of Complex FIR Filters . . . . . . . . . . . . . . . . 15
2.4 Problem Formulation and the Design Methodology . . . . . . . . 18
2.5 Design Examples .................................................. 24
  2.5.1 Example 2.1: Complex FIR Filter with Flatness Constraints 24
  2.5.2 Example 2.2: Complex Bandpass Filter .................... 27
  2.5.3 Example 2.3: Complex Bandpass Filter with Very Low Delay .................................................. 29
2.6 Discussions ...................................................... 31
  2.6.1 On Eigenvalues of \( \tilde{P}(\omega) \) ...................... 31
  2.6.2 Discrepancy in the Estimated Group Delay ............... 34
  2.6.3 On the Design of Real FIR Filters ..................... 35
2.7 Conclusions ..................................................... 39
2.8 Appendix: Analytic Expression for the Eigenvalues of \( \tilde{P}(0) \) . . . 40

3 FRM Filter Design with Group Delay Constraint 44
  3.1 Introduction .................................................... 44
  3.2 Preliminaries .................................................. 48
    3.2.1 FRM Technique ........................................ 48
    3.2.2 SOCP ................................................... 51
  3.3 Group Delay and Gradients for Basic FRM Filters .......... 54
    3.3.1 Group Delay for Basic FRM Filters .................. 54
    3.3.2 Gradients for Basic FRM Filters .................... 57
    3.3.3 Basic FRM Filters with Linear Phase Masking Filters . 61
  3.4 Design Methodology for Basic FRM Filters ................. 64
    3.4.1 Initial Design ....................................... 64
## Contents

3.4.2 Overall Optimization .......................... 65

3.5 Design of Multi-Stage FRM Filters .................. 70

3.5.1 Group Delay for Two-stage FRM Filters .......... 71

3.5.2 Gradients for Two-Stage FRM Filters .......... 74

3.5.3 Two-Stage FRM Filter Design .................. 76

3.6 Design Examples .................................. 76

3.6.1 Example 3.1: A Basic FRM Filter ............... 77

3.6.2 Example 3.2: Another Basic FRM Filter ........ 80

3.6.3 Example 3.3: A Two-Stage FRM Filter ........ 83

3.7 Conclusions ...................................... 87

3.8 Appendix: \( \frac{d\tau_f(\omega)}{dh_i} \) (\( i = 1, 6, 7 \)) for a two-stage FRM filter ..... 88

3.8.1 Common Components ........................... 88

3.8.2 Expressions for \( \frac{d\tau_f(\omega)}{dh_6} \) ............. 90

3.8.3 Expressions for \( \frac{d\tau_f(\omega)}{dh_7} \) ............. 90

3.8.4 Expressions for \( \frac{d\tau_f(\omega)}{dh_1} \) ............. 91

4 Application of the FRM Technique in Array Beamforming 92

4.1 Introduction ..................................... 93

4.2 Analysis of the FRM Technique in Time Domain .... 96

4.3 Passive Array Beamforming Using the FRM Technique .... 101

4.3.1 Passive Array Beamforming .................... 101

4.3.2 Interleaved Linear Array Beamformers .......... 102

4.3.3 Multiple Subarray Beamformers ................. 105
4.4 Active Array Beamforming Using the FRM Technique 107
  4.4.1 Active Interleaved Linear Array Beamformers and Effective Aperture 111
  4.4.2 Aperture Functions by the FRM Technique 112
  4.4.3 Computer Simulations 113
4.5 Active Two-Dimensional Array Beamforming 120
4.6 Conclusions 123

5 General Sidelobe Cancellers with Leakage Constraints 125
  5.1 Introduction 126
  5.2 Derivative Constraints for Broadband Beamformers 129
  5.3 Leakage Constraints for GSCs 132
    5.3.1 First Order Leakage Constraint 134
    5.3.2 Second Order Leakage Constraints 135
  5.4 Implementations 137
    5.4.1 On Estimation of $\frac{\partial^2 \text{R}_{ss}(\theta_s)}{\partial \theta_s^2}|_{\theta_s=0}$ and $\frac{\partial^4 \text{R}_{ss}(\theta_s)}{\partial \theta_s^4}|_{\theta_s=0}$ 137
    5.4.2 Formulation and Implementation in SOCP 139
    5.4.3 On Adaptive Algorithm 140
    5.4.4 On Norm Constrained Adaptive Filtering 141
  5.5 Computer Simulations 142
  5.6 Conclusions 144

6 Conclusions and Recommendations for Future Work 146
  6.1 Conclusions 146
6.2 Recommendations .......................... 147

Author’s Publications .......................... 149

Bibliography .......................... 157
List of Figures

2.1 Design results of the complex filter with flatness constraint in Example 2.1. The magnitude response and group delay of the filter obtained by the proposed method, of ‘Filter 1’ and of ‘Filter 2’ are indicated by solid lines, broken lines and dash-dot lines, respectively. (a) magnitude response, (b) passband ripple, (c) stopband attenuation and (d) group delay. . . . . . . . . . . . . . . . . . . . . . . . . . 25

2.2 Design results of the complex bandpass filter in Example 2.2. The magnitude response and group delay of the filter obtained by the proposed method, of ‘Filter 1’ and of ‘Filter 2’ are indicated by solid lines, broken lines and dash-dot lines, respectively. (a) magnitude response, (b) passband ripple, (c) stopband attenuation and (d) group delay. . . . . . . . . . . . . . . . . . . . . . . . . . 28

2.3 Design results of the complex bandpass filter in Example 2.3. The magnitude response and group delay of the filter obtained by the proposed method, of ‘Filter 1’ and of ‘Filter 2’ are indicated by solid lines, broken lines and dash-dot lines, respectively. (a) magnitude response, (b) passband ripple, (c) stopband attenuation and (d) group delay. . . . . . . . . . . . . . . . . . . . . . . . . . 30

2.4 Eigenvalues of $\tilde{P}(\omega)$ for complex filters with various filter lengths. 32
List of Figures

2.5  Eigenvalues of $\tilde{P}_1(\omega)$ for real filters with various filter lengths. 36

2.6  Design results of the real lowpass filter in Example 2.4. The magnitude response and group delay of the filters obtained by the proposed methods with iteration, without iteration, and the conventional SDP method are indicated by solid lines, broken lines and dash-dot lines, respectively. (a) magnitude response, (b) passband ripple, (c) stopband attenuation and (d) group delay. 38

3.1  A basic FRM filter structure. 48

3.2  A two-stage FRM filter structure. 49

3.3  A simplified basic FRM filter structure. 54

3.4  A simplified two-stage FRM filter structure. 71

3.5  Example 3.1. A basic FRM filter with linear phase masking filters when $\beta = 2 \times 10^{-3}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the initial FRM filter is in broken lines. 78

3.6  Example 3.1. A basic FRM filter with nonlinear phase masking filters when $\beta = 3 \times 10^{-3}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the initial FRM filter is in broken lines. 79

3.7  Example 3.2. Another basic FRM filter with nonlinear phase masking filters when $\beta = 1 \times 10^{-3}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the reference FRM filter is in broken lines. 81
3.8 Example 3.3. A two-stage FRM filter with linear phase masking filters when $\beta = 8.5 \times 10^{-4}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the initial FRM filter is in broken lines. 84

3.9 Example 3.3. A two-stage FRM filter with nonlinear phase masking filters when $\beta = 2 \times 10^{-3}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the initial FRM filter is in broken lines. 85

4.1 A realization structure for a filter using the FRM technique. 95

4.2 Condition for $N_m \geq M$. Maximum passband ripple $\delta_p$ and minimum stopband attenuation $\delta_s$ are expressed in decibels. 95

4.3 An interleaved linear array beamformer using one sparse ULA in solid circles and one dense ULA in squares with $M = 3$. 103

4.4 A typical multiple subarray beamformer with $N_m = M + 2$. The sensors shared by neighbouring subarrays are marked with a dot in a square. 105

4.5 (a) Example of a ‘desired effective aperture’ with element spacing $\lambda/2$. (b,c) Two different combinations of transmitting and receiving aperture functions that yield the same desired effective aperture in (a). 107

4.6 Aperture functions and effective aperture of a VSA with $p = 3$ and $\cos^2(\cdot)$ apodized functions. 108

4.7 An AILA beamformer using a sparse transmitting ULA (in solid circles) and a dense receiving ULA (in hollow squares). 108
4.8 Spatial filtering simulation of the AILA beamformer. (a) The beamformer output signal when a single object is present at 10° or 11° relative to broadside. (b) Scattered signals from two objects present simultaneously at the first excitation. (c) Scattered signals from two objects present simultaneously at the second excitation. (d) The AILA beamformer output when two objects present simultaneously. ........................................ 114

4.9 The effective beampattern of the AILA beamformer. .......... 115

4.10 Beampattern comparison between the AILA beamformer ($N_t = 10$, $N_r = 30$, $M = 10$) and the VSA beamformer with $p = 3$ and $\cos^2(\cdot)$ aperture function (24 sensors in the transmitting array and 24 sensors in the receiving array). ................. 116

4.11 A typical active 2D array comprises one $11 \times 7$ sparse transmitting array in solid circles and one $17 \times 9$ dense receiving array in blank squares with $M_1 = 3$ and $M_2 = 2$. ......................... 118

4.12 Block diagram of 2D FRM filter synthesis. .................. 119

4.13 Effective beampattern of the 2D array. ...................... 119

5.1 A broadband beamformer with $K$ sensors and $L$-tap delay lines, where $\tau_i (i = 1, \ldots, K)$ are presteering delays and $T_s$ is the sampling interval. ......................... 126

5.2 Block diagram of a generalized sidelobe canceller with $K$ sensors. 132

5.3 Spatial responses of various adaptive beamformers in Gaussian noisy environment. ................................. 144

5.4 Signal-to-interference ratio comparison among various adaptive beamformers. .......................... 145
5.5 Convergence of the proposed algorithm. The line on the top is for the GSC with derivative constraints, while the lines for the conventional GSC and the modified GSC almost overlap. . . . . 145
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Design results of the complex FIR filters in Example 2.1.</td>
<td>24</td>
</tr>
<tr>
<td>2.2</td>
<td>Design results of the complex FIR filters in Example 2.2.</td>
<td>27</td>
</tr>
<tr>
<td>2.3</td>
<td>Design results of the complex FIR filters in Example 2.3.</td>
<td>29</td>
</tr>
<tr>
<td>2.4</td>
<td>Non-zero eigenvalues of $\tilde{P}(\omega)$ in Figure 2.4.</td>
<td>32</td>
</tr>
<tr>
<td>2.5</td>
<td>Non-zero eigenvalues of $\tilde{P}_1(\omega)$ ($\omega = 0.12\pi$) in Figure 2.5.</td>
<td>35</td>
</tr>
<tr>
<td>2.6</td>
<td>Design results of the real FIR filters in Example 2.4.</td>
<td>38</td>
</tr>
<tr>
<td>3.1</td>
<td>The passband and stopband cutoff frequencies of the subfilters.</td>
<td>51</td>
</tr>
<tr>
<td>3.2</td>
<td>Important parameters for the basic FRM filters for Example 3.1.</td>
<td>82</td>
</tr>
<tr>
<td>3.3</td>
<td>Important parameters for the basic FRM filters for Example 3.2.</td>
<td>83</td>
</tr>
<tr>
<td>3.4</td>
<td>Important parameters for the two-stage FRM filters for Example 3.3.</td>
<td>87</td>
</tr>
<tr>
<td>4.1</td>
<td>Comparison between the FRM filters and the optimum filters.</td>
<td>98</td>
</tr>
<tr>
<td>4.2</td>
<td>Frequency specifications of the FRM subfilters with $M = 8$.</td>
<td>114</td>
</tr>
<tr>
<td>4.3</td>
<td>Frequency specifications of the aperture functions in a narrow-beamwidth beamformer synthesis.</td>
<td>115</td>
</tr>
<tr>
<td>4.4</td>
<td>Frequency specifications of the FRM subfilters with $M = 7$.</td>
<td>115</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Background and Motivations

This thesis presents research work on two related subjects: finite impulse response (FIR) filter design and array beamforming. FIR filters are widely used in many applications where the underlying object is signal in temporal domain, while beamformers provide spatial filtering or temporal-spatial filtering. The focus of the thesis is on the design of FIR filters with reduced passband group delay errors and efficient/robust array beamforming. A brief general review on related work is given below, which provides the background to carry out our research work. More detailed discussions on the related work will be given in the subsequent chapters.

- It is well known that linear phase FIR filters can be designed using conventional methods, such as the Park-McClellan algorithm [PM72], the
1.1. Background and Motivations

least-squares approach [TF70, Kel72], or the eigenfilter approach [VN87, Ngu93, PT01, TVN03], etc. Linear phase FIR filters are often the users’ first choice primarily because they are easy to design and implement due to the symmetric property of their filter coefficients. However, the group delay of a linear phase FIR filter, which is about half of the filter length, could be unacceptably large in some applications when the filter length is large, especially when it is designed to meet stringent requirements on the passband ripple, stopband attenuation and/or transition bandwidth.

Besides the desired magnitude response, low and accurate group delays are required for devices and digital filters to avoid long delay and minimize distortion of the transmitted information in certain applications, such as microwave and radar communications, optical communications and medical imaging, etc. [Cho95, VWHvO98, RO00, BGCL+03, FMC05]. There is a growing interest in designing FIR filters with reduced group delays in recent years. Conventional design methods, such as the Park-McClellan algorithm, the least-squares approach and the eigenfilter approach could be extended for this purpose [PS92, SCP93, Ngu93, KM95, KR95, KR96, TVN03, Lu04].

In general, the prescribed group delay is converted to the phase response and the desired complex frequency response, which combines the magnitude and phase responses, is then approximated. Although the estimation error, in the sense of either minimax or least-squares, is minimized, the group delay error is relatively large because the estimation error, which combines both magnitude and phase errors, is minimized instead of group delay error. Because of the nonlinear relationship between the group delay and the corresponding phase response, minimal group delay error is not
1.1. Background and Motivations

guaranteed even if the phase error is minimized.

This motivates the present investigation on designing FIR filters with reduced passband group delays and reduced group delay errors, or in short, FIR filters with reduced group delay errors throughout the thesis, by directly minimizing the group delay errors.

- A special type of FIR filters based on the frequency-response masking (FRM) technique [Lim86] has received much attention in the past two decades. FRM filters are often the designers’ primary choice due to their considerable complexity reduction in implementing FIR filters with sharp transition bands, compared to those direct-form FIR filters designed by conventional methods. Based on the concept of FRM, other novel structures to further reduce the implementation complexity were discussed in [LZK01, Lia03, LY03].

A basic FRM filter consists of a group of subfilters, namely, a prototype filter and a pair of masking filters. These subfilters are designed such that sharp transition band can be achieved when integrating under the FRM structure. Given the desired passband ripple, stopband attenuation and transition band, the bandedges of the subfilters can be determined, though not unique. The design of an FRM filter usually starts from designing masking filters by any conventional method. The prototype filter is then designed to minimize the ripples of the FRM filter around the transition band using, for example, linear programming [Lim86]. Since the ripple magnitudes of the masking filters are selected arbitrarily and the subfilters are separately designed, FRM filters designed in this way usually possess considerable design margin in passband ripple and stop-
band attenuation, which suggests that the design is suboptimal and the implementation complexity could be further reduced. A two-step optimization technique was proposed for simultaneously optimizing the sub-filters in [SYKJ03]. With a good initial FRM filter obtained at the first step, unconstrained nonlinear optimization is carried out at the second step. Using this method, the implementation complexity of the FRM filter was further reduced compared to the one in [Lim86]. Noted that all these FRM filters [Lim86, SYKJ03] as well as their subfilters are of linear phase. Although the implementation complexity of a linear phase FRM filter is successfully reduced, its group delay is even larger than that of the direct-form FIR filter with the same approximation accuracy designed by any conventional method [LL93].

Similar to the design of general FIR filters, the design of FRM filters with reduced passband group delays has also received growing interest in recent years [CL96, SJ02, LH03b, LL05]. Usually, the prescribed group delay of an FRM filter is distributed to the desired ones for the subfilters before each subfilter is designed. In the above methods, although the estimation error, which includes the magnitude and phase errors, is minimized, the group delay error is relatively large, noticeably, towards the bandedges.

This motivates us to investigate the efficient design of FRM filters with reduced group delay errors.

- A conventional beamformer, using a predetermined FIR filter as the aperture weighting function, linearly combines the spatially sampled time sequences from each sensor to obtain a scalar output time sequence. The correspondence between FIR filtering and beamforming is closest when
the beamformer operates at a single temporal frequency and the array geometry is linear and equi-spaced [VB88], or the so-called uniform linear array (ULA) beamformer. For a ULA beamformer, its response can be directly mapped to the frequency response of the corresponding FIR filter. The analogy between FIR filters and ULA beamformers led to the applications of FIR filter design methods in array beamformer synthesis [VB88, Van02]. To achieve desirable beampatterns with sharp transition bands and low sidelobes, a large number of sensors is usually required by beamformers, leading to high costs and heavy computational loads.

To reduce the system cost in the passive array beamformers, sparse arrays have received great attention from system designers [HBG97, WASH99]. In a uniform sparse array, grating lobes appear and cause spatial ambiguity. For active array beamformers that are commonly used in ultrasonic diagnostic systems and synthetic aperture radar systems, etc., several approaches have been proposed [VST75, LLOF96, AH00]. An approach was proposed to reduce the number of elements in a linear array while minimizing grating lobes with different spacings for the transmitting and receiving elements [VST75]. Continuing in this direction, four different strategies were proposed [LLOF96] for designing sparse arrays with emphasis on smoothness of the effective aperture function. For hardware simplicity, the transmitting and receiving aperture functions are simply unity, cosine or cosine-square functions. A novel approach was proposed [AH00] to further suppress the grating lobes by introducing convolution kernels for the transmitting and receiving aperture functions.

Because of the analogy between the FIR filtering and the conventional ULA beamforming, and the implementation efficiency of the FRM filters,
we focus on efficient beamforming by investigating the feasibility of the applications of the FRM technique in array beamforming, with emphasis on the reduction of the number of sensors and hence the computational complexity of the associated beamforming algorithm.

• Unlike conventional beamformers which may use predetermined FIR filters as their aperture weighting functions, adaptive beamformers are able to adjust the beampatterns in real time to maintain the prescribed frequency responses in the desired directions while introducing nulls in the interference directions. A typical adaptive array beamformer is the linear constrained minimum variance (LCMV) beamformer.

It is well known that LCMV beamformers suffer significant performance degradation when ideal assumptions do not exist in practice. Tremendous work has been done to improve the robustness of adaptive beamformers [God85, KU92, LL97b, LL97a, BEV00, ZYL04]. Typically, to suppress the signal cancellation caused by direction of arrival (DOA) mismatch, some constraints were proposed such as multipoint linear constraint [Nun83], soft quadratic response constraint [EC85], and maximally flat spatial power response derivative constraints [EC83]. With these constraints, the beamformer is robust in the vicinity of the assumed direction. The widened beamwidth is achieved at the cost of the decreased capability in interference suppression due to the reduction in the degree of freedom.

As some robust adaptive beamforming problems can be regarded as an convex optimization problems, they can be cast in the second-order cone programming (SOCP) framework and solved via the well-established interior point methods. This approach was adopted recently in robust narrow-
band beamforming [VGL03, VGLM04, YM05, LB05]. The advantage of this approach is its flexibility to incorporate more practical errors, which can be formulated as second-order cone constraints.

Motivated by the work done by other researchers, we focus on robust broadband beamforming against DOA mismatch for general sidelobe cancellers (GSCs), which is an alternative structure for implementing the LCMV beamformer.

1.2 Objectives

- To devise a new design method for general FIR filters with reduced passband group delays, such that the complex frequency responses can be achieved with reduced group delay errors.

- To devise an efficient design method for FRM FIR filters with reduced passband group delays, such that the prescribed group delays can be achieved with reduced group delay errors.

- To study in detail the feasibility of the applications of the FRM technique in array beamforming, with emphasis on the reduction of the number of sensors and hence the computational complexity of the associated beamforming algorithm.

- To find a new class of linear constraints to alleviate the signal cancellation due to DOA mismatch for broadband GSCs.
1.3 Major Contributions

1. An analytic expression of the group delay of a general FIR filter is derived. A group delay constraint is formulated as a linear matrix inequality (LMI) with some reasonable approximations. A new design method based on semidefinite programming (SDP) is proposed with the magnitude and group delay constraints. This method offers a designer flexibility in meeting the specifications in both magnitude response and group delay.

2. Analytic expressions of the group delay of an FRM filter are derived, as well as the gradient of the group delay with respect to the subfilter coefficients. The expressions are valid for both basic (or single-stage) and multi-stage FRM filters. An efficient design method for FRM filters with reduced passband group delays is proposed in which both the magnitude and an additional group delay constraints are incorporated into the framework of the second-order cone programming (SOCP). Using the proposed design method, the group delay error is effectively reduced at the cost of slight increase in magnitude ripples.

3. The feasibility of applications of the FRM technique in array beamforming is investigated in detail. On one hand, it is shown that there is a limitation in applying the FRM technique in passive array beamforming. On the other hand, for active array beamforming, a novel combination of the concept of effective aperture and the FRM technique does lead to synthesis of desirable beamformers with reduced number of sensors.

4. A class of leakage constraints for robust broadband GSCs to suppress signal cancellation due to DOA mismatch is derived by exploiting the
1.4. Organizations

statistical property of the leakage signals.

1.4 Organizations

The thesis is organized as follows.

We first discuss a new design method for FIR filters with reduced group delays in Chapter 2. An analytic expression of the group delay of a complex FIR filter is derived, which shows that the group delay is a nonlinear function with respect to the filter coefficients. To take the advantage of the flexibility of SDP, we subsequently formulate a group delay constraint as an LMI with some reasonable approximations. By balancing the complex magnitude error and the group delay error, the group delay constraint is then incorporated into the SDP framework, leading to a new design method for FIR filters with reduced group delays. The estimation error in group delay due to the approximations is discussed. Finally, the proposed design method for real FIR filters is briefly discussed.

An improved design method for FRM filters with reduced group delays is discussed in Chapter 3. For this structurally complicated filter, analytic expressions of the group delay of a basic (or single-stage) FRM filter is obtained, as well as the gradient of the group delay with respect to the subfilter coefficients. Subsequently, a group delay constraint is formulated and incorporated into the proposed gradient-based design method based on SOCP. This efficient design method is also extended to the design of multi-stage FRM filters.

In Chapter 4, the feasibility of applications of the FRM technique in array beamforming is investigated in detail with emphasis on the reduction of
the number of sensors with satisfactory beampattern. We first analyze the FRM filtering in temporal domain and show the difference between the temporal filtering and spatial filtering, which makes it infeasible to apply the FRM technique in passive array beamforming to reduce the number of sensors while maintaining satisfactory beampattern. Subsequently, we propose an active array beamforming method by novelly combining the FRM technique with the concept of effective aperture. The active array beamforming method is also generalized to the 2D case and illustrated by simulations.

Unlike conventional beamformers which may use predetermined FIR filters as their aperture weighting functions, GSCs can adjust the beampatterns to maintain the prescribed frequency responses in the desired directions while introducing nulls in the interference directions. To suppress signal cancellation due to DOA mismatch in GSCs, we present a class of linear constraints, namely, leakage constraints in Chapter 5. The leakage constraint is derived by exploiting the statistical property of the leakage signals. Subsequently, some implementation issues are discussed before presenting the computer simulations.

Finally, conclusions and recommendations for future work are given in Chapter 6.
Chapter 2

FIR Filter Design with Group Delay Constraint

2.1 Introduction

It is well known that the group delay of a linear phase FIR filter, which is about half of the filter length, could be unacceptably large when the filter length is large, especially when stringent requirements are imposed at the passband ripple, stopband attenuation and/or transition bandwidth. Besides the magnitude response, in certain applications, low passband group delay and low group delay error are required [Cho95, VWHvO98, RO00, BGCL+03, FMC05]. Traditionally, the prescribed group delay is converted to the phase response and the desired complex frequency response, which combines the magnitude and phase responses, was then approximated [PS92, Ngu93, KR95, LVKL96, TVN03] and the prescribed group delay was achieved indirectly.

Although the estimation error, either in the sense of minimax or least-
2.1. Introduction

squares, is minimized, the group delay error is relatively large because the estimation error, which combines both the magnitude and phase errors, is minimized instead of group delay error. Because of the nonlinear relationship between the group delay and the corresponding phase response, minimum group delay error cannot be guaranteed even if the phase error is minimized. In addition, extension of the above design methods to incorporate more constraints is not easy.

Recently, the design of various types of FIR and infinite impulse response (IIR) filters was formulated in the semidefinite programming (SDP) framework [Lu99, Lu00, Lu02, Lu04, CT04]. SDP is a relatively new optimization methodology which is primarily concerned with minimizing a linear or convex quadratic objective function subject to linear matrix inequality (LMI) constraints that depend on the design variable affinely [BV04]. There are several advantages in formulating a problem as an SDP problem. Two obvious advantages are (1) it is flexible to satisfy multiple objectives expressed in terms of a set of linear and convex quadratic constraints; (2) the problem can then be solved, reliably and efficiently, using an interior-point method. As an alternative and promising approach to FIR filter design, the design problem was formulated as an SDP problem with the constraint on the estimation error, in the sense of minimax, reformulated as an LMI [Lu99, Lu00, Lu02, Lu04]. This approach was also applied to the estimation error in the sense of least-squares [CT04]. Due to the flexibility of the SDP formulation, other LMI constraints can also be successfully incorporated, e.g., the magnitude flatness constraints and peak error constraints [CT04].

Because of the advantages of the SDP formulation, in this chapter, we pro-
pose a new method for designing FIR filters with reduced group delays and reduced group delay errors. A key step in our method is to directly impose a group delay constraint, which is formulated as an LMI with reasonable approximations. The main advantage of the proposed design method is that the significant reduction in the group delay error can be achieved at the cost of slight increase in the magnitude error.

In many signal processing applications, such as wavelet transform, channel equalization, digital video processing, etc. [ZDP99, PBN99, BB92], the need arises for the design of complex coefficient FIR filters to meet some specifications that cannot be achieved by real coefficient filters, such as asymmetric spectral response. Since real FIR filters can be considered as a special case of complex FIR filters, most of the discussion in this chapter is devoted to complex FIR filters and a brief discussion for real FIR filters is also presented at the end.

The rest of the chapter is organized as follows. Convex optimization and SDP are briefly reviewed in Section 2.2 where some notations are introduced that are to be used in the rest of the chapter. The group delay of a complex FIR filter is derived in Section 2.3. A new FIR filter design methodology is proposed in Section 2.4 and some design examples are given in Section 2.5. The estimation error caused by the approximations and the design of real FIR filters is briefly discussed in Section 2.6. Conclusions are given in Section 2.7, followed by an analytic expression for the eigenvalues of $\tilde{P}(0)$ in the Appendix (Section 2.8).
2.2 Convex Optimization and SDP

It is well known that linear programming, a well-established theory arising in a variety of applications, can be efficiently solved numerically. The same can be said for convex optimization, a special class of mathematical optimization problems [BV04]. A general problem of convex optimization is to find the minimum of a convex (or quasiconvex) function on a finite-dimensional convex set.

Semidefinite programming (SDP) is an optimization methodology which is primarily concerned with minimizing a linear or convex quadratic objective function subject to linear matrix inequality (LMI) type constraints that depend on the design variables affinely [VB96]. The class of SDP problem can be expressed as

\[
\begin{align*}
\text{minimize} & \quad f^T x \\
\text{subject to} : & \quad F(x) \succeq 0 \\
F(x) & = F_0 + \sum_{i=1}^{N} x_i F_i, \quad i = 1, 2, \ldots, N,
\end{align*}
\]

where \( f \in \mathbb{R}^{n \times 1} \) is the problem parameter, \( x = [x_1, x_2, \ldots, x_n] \) is the optimization variable, \( F_i \in \mathbb{R}^{m \times m} \) for \( i = 0, 1, \ldots, N \) are symmetric matrices and \( F(x) \succeq 0 \) denotes that \( F(x) \) is positive semidefinite at \( x \). Note that \( F(x) \) is affine with respect to \( x \). An SDP problem is a constrained optimization problem where the objective function is convex and the feasible region characterized by the constraint is a convex set. SDP includes both linear and convex quadratic programming as its special cases and represents a subclass of convex programming [LH03b]. As an important class of convex optimization prob-
2.3. Group Delay of Complex FIR Filters

Let
\[ H(z) = \sum_{n=0}^{N-1} h(n)z^{-n} \]
be the transfer function of a complex FIR filter of length \( N \), and \( h_r(n), h_i(n) \) be the real and imaginary parts of \( h(n) \), i.e., \( h(n) = h_r(n) + jh_i(n) \).

The frequency response \( H(\omega) \), phase response \( \phi(\omega) \) and group delay \( \tau(\omega) \) of \( H(z) \) are given by
2.3. Group Delay of Complex FIR Filters

\[ H(\omega) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n} = (h_r + jh_i)^T (c(\omega) + js(\omega)), \quad (2.2a) \]

\[ \phi(\omega) = \arctan 2 \left( h_r^T s(\omega) + h_i^T c(\omega), h_r^T c(\omega) - h_i^T s(\omega) \right), \quad (2.2b) \]

\[ \tau(\omega) = -\frac{d}{d\omega} \phi(\omega) = -\frac{d}{d\omega} \tan^{-1} \left( \frac{h_r^T s(\omega) + h_i^T c(\omega)}{h_r^T c(\omega) - h_i^T s(\omega)} \right), \quad (2.2c) \]

respectively, where

\[ h_r = \begin{bmatrix} h_r(0) & h_r(1) & \cdots & h_r(N-1) \end{bmatrix}^T, \]

\[ h_i = \begin{bmatrix} h_i(0) & h_i(1) & \cdots & h_i(N-1) \end{bmatrix}^T, \]

\[ c(\omega) = \begin{bmatrix} 1 & \cos \omega & \cdots & \cos(N-1)\omega \end{bmatrix}^T, \]

\[ s(\omega) = \begin{bmatrix} 0 & -\sin \omega & \cdots & -\sin(N-1)\omega \end{bmatrix}^T, \]

and

\[ \arctan 2(y, x) = \begin{cases} \arctan \left( \frac{y}{x} \right) & \text{for } x \geq 0, \\ \arctan \left( \frac{y}{x} \right) \pm \pi & \text{otherwise}. \end{cases} \]

Using the derivative identity

\[ \frac{d}{d\omega} \arctan 2(y, x) = \frac{d}{d\omega} \tan^{-1} \left( \frac{y}{x} \right) = \frac{1}{1 + (y/x)^2} \frac{d(y/x)}{d\omega}, \quad (2.3) \]
and letting

\[
|H(\omega)|^2 = (h^T_i s(\omega) + h^T_i c(\omega))^2 + (h^T_i c(\omega) - h^T_i s(\omega))^2,
\]

\[
\tilde{c}(\omega) = \frac{d c(\omega)}{d \omega} = \begin{bmatrix} 0 & -\sin \omega & \ldots & -(N-1)\sin(N-1)\omega \end{bmatrix}^T,
\]

\[
\tilde{s}(\omega) = \frac{d s(\omega)}{d \omega} = \begin{bmatrix} 0 & -\cos \omega & \ldots & -(N-1)\cos(N-1)\omega \end{bmatrix}^T,
\]

we arrive at the following analytic expression for the group delay

\[
\tau(\omega) = -\frac{(h^T_i c(\omega) - h^T_i s(\omega))^2}{|H(\omega)|^2} \frac{d}{d \omega} \left( \frac{h^T_i s(\omega) + h^T_i c(\omega)}{h^T_i c(\omega) - h^T_i s(\omega)} \right) \\
= -\frac{1}{|H(\omega)|^2} \left( (h^T_i \tilde{s}(\omega) + h^T_i \tilde{c}(\omega))(h^T_i c(\omega) - h^T_i s(\omega)) - (h^T_i s(\omega) + h^T_i c(\omega))(h^T_i \tilde{c}(\omega) - h^T_i \tilde{s}(\omega)) \right) \\
= \frac{h^T_i (s(\omega)\tilde{c}(\omega)^T - \tilde{s}(\omega)c(\omega)^T)h_r + h^T_i (c(\omega)\tilde{c}(\omega)^T - \tilde{c}(\omega)c(\omega)^T)h_i}{|H(\omega)|^2} \\
+ \frac{h^T_i (\tilde{s}(\omega)s(\omega)^T - s(\omega)\tilde{s}(\omega)^T)h_i + h^T_i (\tilde{c}(\omega)s(\omega)^T - c(\omega)\tilde{s}(\omega)^T)h_i}{|H(\omega)|^2} \\
= \frac{h^T_i \mathbf{P}(\omega)h_x}{|H(\omega)|^2},
\]

where

\[
h_x = [h^T_r h^T_i]^T,
\]

\[
\mathbf{P}_1(\omega) = s(\omega)\tilde{c}(\omega)^T - \tilde{s}(\omega)c(\omega)^T, \quad \mathbf{P}_2(\omega) = c(\omega)\tilde{c}(\omega)^T - \tilde{c}(\omega)c(\omega)^T,
\]

\[
\mathbf{P}_3(\omega) = \tilde{s}(\omega)s(\omega)^T - s(\omega)\tilde{s}(\omega)^T, \quad \mathbf{P}_4(\omega) = \tilde{c}(\omega)s(\omega)^T - c(\omega)\tilde{s}(\omega)^T,
\]
and the \((m, n)\)-th entries of \(P_i(\omega)\) \((m, n = 1, \ldots, N; i = 1, \ldots, 4)\) given by

\[
\begin{align*}
P_1(m, n) &= (n - 1) \sin(m - 1)\omega \sin(n - 1)\omega + (m - 1) \cos(m - 1)\omega \cos(n - 1)\omega \\
P_2(m, n) &= (m - 1) \sin(m - 1)\omega \cos(n - 1)\omega - (n - 1) \cos(m - 1)\omega \sin(n - 1)\omega \\
P_3(m, n) &= (m - 1) \cos(m - 1)\omega \sin(n - 1)\omega - (n - 1) \sin(m - 1)\omega \cos(n - 1)\omega \\
P_4(m, n) &= (m - 1) \sin(m - 1)\omega \sin(n - 1)\omega + (n - 1) \cos(m - 1)\omega \cos(n - 1)\omega.
\end{align*}
\]

### 2.4 Problem Formulation and the Design Methodology

The classical minimax problem in designing a general complex FIR filter can be stated as

\[
\begin{align*}
\text{minimize}_{h_r, h_i} \left\{ \max_{\omega \in \Omega_m} W_m(\omega) |H(\omega) - H_{\text{des}}(\omega)| \right\}, \quad (2.5)
\end{align*}
\]

where \(\Omega_m\) is the band of interest, which consists of the passband and stopband of the band-selective filter, \(W_m(\omega)\) is a positive weighting function and \(H_{\text{des}}(\omega)\) is a desired frequency response, which is usually complex-valued. The above minimax problem can be reformulated as a constrained minimization problem [Lu99, CT04]

\[
\begin{align*}
\text{minimize}_{h_r, h_i} \quad & \eta_m \\
\text{subject to : } \quad & W_m^2(\omega) |H(\omega) - H_{\text{des}}(\omega)|^2 \leq \eta_m \quad \text{for } \omega \in \Omega_m. \quad (2.6b)
\end{align*}
\]
Let $H_{\text{des}}(\omega) = H_{\text{dr}}(\omega) + jH_{\text{di}}(\omega)$. We have

$$W_m^2(\omega) |H(\omega) - H_{\text{des}}(\omega)|^2 = \alpha_1^2(h_{x, \omega}) + \alpha_2^2(h_{x, \omega}),$$

where

$$\alpha_1(h_{x, \omega}) = W_m(\omega) \left[ h_{T, c}(\omega) - s(\omega) - H_{\text{dr}}(\omega) \right],$$
$$\alpha_2(h_{x, \omega}) = W_m(\omega) \left[ h_{T, s}(\omega) + c(\omega) - H_{\text{di}}(\omega) \right].$$

Constraint (2.6b) is equivalent to [Lu99, CT04]

$$\hat{\Phi}_m(h_{x, \omega}) = \begin{bmatrix}
\eta_m & \alpha_1(h_{x, \omega}) & \alpha_2(h_{x, \omega}) \\
\alpha_1(h_{x, \omega}) & 1 & 0 \\
\alpha_2(h_{x, \omega}) & 0 & 1
\end{bmatrix} \succeq 0 \quad \text{for } \omega \in \Omega_m. \quad (2.7)$$

As (2.7) is in a form of LMI, the minimax problem can thus be cast as an SDP problem and solved using standard software packages such as the robust control toolbox from MathWorks Inc.[BCPS05]. Moreover, additional constraints such as magnitude flatness constraints could be imposed (see, e.g. [CT04]).

Although the complex magnitude error is minimized in (2.5), the group delay error may not be minimized. Throughout this chapter, we assume that the filter length is $N$ and the desirable group delay is $\tau_{\text{des}}$ with $\tau_{\text{des}} < (N - 1)/2$. It is well known that exact group delay equal to $\tau_{\text{des}}$ cannot be achieved in general and hence it is desirable to aim at minimizing the group delay error in addition to minimizing the complex magnitude error. Similarly to (2.5) and
(2.6), minimization of the maximum group delay error can be stated as

\[
\min_{h_r, h_i} \left\{ \max_{\omega \in \Omega_g} W_g(\omega) |\tau(\omega) - \tau_{des}| \right\},
\]

and then reformulated as the following constrained minimization problem

\[
\begin{align*}
\min_{h_r, h_i} & \quad \eta_g \\
\text{subject to} : & \quad W_g(\omega) |\tau(\omega) - \tau_{des}| \leq \eta_g \quad \text{for} \quad \omega \in \Omega_g,
\end{align*}
\]

where \( \Omega_g \) is the frequency band where the desired group delay is prescribed, which consists of the passband of the band-selective filter, and \( W_g(\omega) \) is a positive weighting function.

We now discuss in detail on how to formulate constraint (2.9b) as another LMI so that the combined minimax problem of (2.6) and (2.9) can be solved using SDP solver.

To reformulate the group delay constraint in (2.9b) as an LMI, \( P(\omega) \) in (2.4) has to be positive definite for any \( \omega \in \Omega_g \), which is however not the case. To overcome this difficulty, we first introduce a symmetric matrix \( \tilde{P}(\omega) \) such that

\[
h_x^T P(\omega) h_x = h_x^T \tilde{P}(\omega) h_x.
\]

Such a matrix can be readily obtained by \( \tilde{P}(\omega) = \frac{1}{2} (P + P^T) \). Defining \( \tilde{P}(\omega) = \begin{bmatrix} \tilde{P}_1(\omega) & \tilde{P}_3(\omega) \\ \tilde{P}_2(\omega) & \tilde{P}_4(\omega) \end{bmatrix} \), where the \((m,n)\)-th entries of \( \tilde{P}_i(\omega) \) \((m,n = 1, \ldots, N; i = 1, 3)\)
1, . . . , 4) are

\[ \tilde{P}_1(m, n) = \frac{1}{2}(P_1(m, n) + P_1(n, m)) = \frac{m + n - 2}{2} \cos(m - n)\omega, \]

\[ \tilde{P}_2(m, n) = \tilde{P}_3(n, m) = \frac{1}{2}(P_2(m, n) + P_3(n, m)) = \frac{m + n - 2}{2} \sin(m - n)\omega, \]

\[ \tilde{P}_4(m, n) = \frac{1}{2}(P_4(m, n) + P_4(n, m)) = \frac{m + n - 2}{2} \cos(m - n)\omega. \]

Although \( \tilde{P}(\omega) \) is now symmetrical, it is still not positive definite. Exhaustive simulations have demonstrated that for any arbitrarily fixed frequency \( \omega \), \( \tilde{P}(\omega) \) always has one positive eigenvalue \( \sigma \) of multiplicity 2, and one negative eigenvalue \( \nu \) of multiplicity 2, both of which are frequency independent, and satisfy \( \sigma \gg |\nu| \). The rest of the eigenvalues of \( \tilde{P}(\omega) \) are all zeros. In fact, it is proven in Appendix Section 2.8 that for \( \omega = 0 \) and \( N \geq 2 \), the two non-zero eigenvalues of \( \tilde{P}(\omega) \) are

\[ \sigma = \frac{N}{4} \left( N - 1 + \sqrt{\frac{4N^2 - 6N + 2}{3}} \right) > 0 \]

\[ \nu = \frac{N}{4} \left( N - 1 - \sqrt{\frac{4N^2 - 6N + 2}{3}} \right) < 0. \]

For \( N \geq 10 \), we have \( \frac{\sigma}{|\nu|} > 11.7 \) and

\[ \lim_{N \to \infty} \frac{\sigma}{|\nu|} = \frac{2 + \sqrt{3}}{2 - \sqrt{3}} \approx 13.93. \]

Thus, we can express \( \tilde{P}(\omega) \) as \( \tilde{P}(\omega) = U(\omega)\hat{D}U^T(\omega) \) by eigen-decomposition where \( \hat{D} = \text{diag}\{\sigma, \sigma, \nu, \nu, 0, \ldots, 0\} \). Since \( \sigma \gg |\nu| \), \( \tilde{P}(\omega) \) can be approximated by a positive semidefinite matrix as follows

\[ \tilde{P}(\omega) \approx \hat{P}(\omega) = U^T(\omega)\hat{D}U(\omega) \quad (2.11) \]
2.4. Problem Formulation and the Design Methodology

where \( \hat{D} = \text{diag}\{\sigma, \sigma, 0, \ldots, 0\} \).

With (2.4), (2.10) and (2.11), constraint (2.9b) can be approximated as

\[
|H(\omega)|^2 \left( \tau_{\text{des}} - \frac{\eta_g}{W_g(\omega)} \right) \leq h_x^T \hat{P}(\omega) h_x \leq |H(\omega)|^2 \left( \tau_{\text{des}} + \frac{\eta_g}{W_g(\omega)} \right) \quad \text{for } \omega \in \Omega_g.
\]

(2.12)

In the following, we replace \(|H(\omega)|\) with \(|H_{\text{des}}(\omega)|\) in (2.12) since both are very close in \(\Omega_g\), and adopt only the right hand side inequality since the left hand side inequality cannot be formulated as an LMI and exhaustive simulations have demonstrated that the left hand side inequality is also satisfied when the right hand side inequality is satisfied. With the above reasonable approximations, we finally arrive at the following group delay constraint

\[
h_x^T \hat{P}(\omega) h_x \leq |H_{\text{des}}(\omega)|^2 \left( \tau_{\text{des}} + \frac{\eta_g}{W_g(\omega)} \right) \quad \text{for } \omega \in \Omega_g.
\]

(2.13)

With simple algebra, it can be shown that (2.13) is equivalent to

\[
\hat{\Phi}_g(h_x, \omega) = \begin{bmatrix} |H_{\text{des}}(\omega)|^2 \left( \tau_{\text{des}} + \frac{\eta_g}{W_g(\omega)} \right) & h_x^T B(\omega) \\ B^T(\omega) h_x & I \end{bmatrix} \succeq 0 \quad \text{for } \omega \in \Omega_g,
\]

(2.14)

where \( I \) is a \(2N \times 2N\) identity matrix and \( B(\omega) \) is a \(2N \times 2N\) matrix satisfying \( B(\omega) = U(\omega) \hat{D}^{1/2} \).

Letting \( \beta \) be a real number balancing the complex magnitude error and the group delay error, and replacing constraint (2.6b) with (2.7) and (2.9b) with (2.14), respectively, the minimax problems of (2.6) and (2.9) can be combined...
as

\[
\begin{align*}
\text{minimize} & \quad \eta \\
\text{subject to :} & \quad \hat{\Phi}_m(h_x, \omega) = \begin{bmatrix} \eta & \alpha_1(h_x, \omega) & \alpha_2(h_x, \omega) \\ \alpha_1(h_x, \omega) & 1 & 0 \\ \alpha_2(h_x, \omega) & 0 & 1 \end{bmatrix} \succeq 0 \quad \text{for} \ \omega \in \Omega_m \\
& \quad \hat{\Phi}_g(h_x, \omega) = \begin{bmatrix} |H_{des}(\omega)|^2 \left( \tau_{des} + \frac{\eta}{\omega_{g} g(\omega)} \right) \ h_x^T B(\omega) \\ B^T(\omega) h_x \\ I \end{bmatrix} \succeq 0 \quad \text{for} \ \omega \in \Omega_g.
\end{align*}
\]

(2.15a

(2.15b

(2.15c

Define an augmented vector \( x = [\eta \ h_x^T]^T \), the objective function in (2.15a) can be expressed as \( \eta = f^T x \) with \( f = [1 \ 0 \ldots 0]^T \). By digitizing the frequency variable \( \omega \) over a dense set of frequencies \( \{\omega_{m_1}, \ldots, \omega_{m_K}\} \subset \Omega_m \) and \( \{\omega_{g_1}, \ldots, \omega_{g_K}\} \subset \Omega_g \), the discretized version of (2.15) can be written as

\[
\begin{align*}
\text{minimize} & \quad f^T x \\
\text{subject to :} & \quad G(x) \succeq 0,
\end{align*}
\]

(2.16a

(2.16b

where

\[
\begin{align*}
G(x) &= \mathrm{diag} \left\{ \Phi_m(x), \Phi_g(x) \right\}, \\
\Phi_m(x) &= \mathrm{diag} \left\{ \hat{\Phi}_m(h_x, \omega_{m_1}), \ldots, \hat{\Phi}_m(h_x, \omega_{m_K}) \right\}, \\
\Phi_g(x) &= \mathrm{diag} \left\{ \hat{\Phi}_g(h_x, \omega_{g_1}), \ldots, \hat{\Phi}_g(h_x, \omega_{g_K}) \right\}.
\end{align*}
\]

Note that matrix \( G(x) \) in (2.16b) is affine with respect to vector \( x \) and hence
Table 2.1: Design results of the complex FIR filters in Example 2.1.

| Design methods                          | max $|E_r(\omega)|$ | $\delta_p$(dB) | $\delta_s$(dB) |
|-----------------------------------------|-----------------|----------------|----------------|
| Minimax with unity $W_m(\omega)$ (Filter 1) | 1.0712          | 0.4053         | -30.0266       |
| Minimax with $W_m(\omega)$ in (2.17) (Filter 2) | 0.9332          | 0.3728         | -27.8237       |
| Proposed method                         | 0.7033          | 0.3622         | -27.3725       |

(2.16) is an SDP problem.

Regarding the selection of $\beta$, it can be approximated by ratio of $\eta_m$ and $\eta_g$.

Noted in this chapter, $\eta_m$ denotes the square of the complex magnitude error. Both $\eta_m$ and $\eta_g$ can be computed from a reference design using conventional SDP method. To improve the performance of the group delay, slightly larger $\beta$ should be applied.

## 2.5 Design Examples

To demonstrate the effectiveness of the proposed method, three illustrative examples are given in this section.

### 2.5.1 Example 2.1: Complex FIR Filter with Flatness Constraints

A complex filter of length $N = 31$, with passband cutoff frequencies $\omega_{p1} = -0.1\pi$, $\omega_{p2} = 0.3\pi$, stopband cutoff frequencies $\omega_{s1} = -0.2\pi$, $\omega_{s2} = 0.4\pi$ and desirable constant passband group delay $\tau_{des} = 12$ over the passband is considered here [CT04]. For comparison with [CT04], the magnitude flatness constraint of order 2 at $\omega = 0$ and another magnitude constraint with two zeros at $\omega = \pi$ are incorporated into the proposed design method. The numbers of frequency grids
Figure 2.1: Design results of the complex filter with flatness constraint in Example 2.1. The magnitude response and group delay of the filter obtained by the proposed method, of ‘Filter 1’ and of ‘Filter 2’ are indicated by solid lines, broken lines and dash-dot lines, respectively. (a) magnitude response, (b) passband ripple, (c) stopband attenuation and (d) group delay.
are $m_K = 220$ and $g_K = 80$. The magnitude responses and the group delays of the designed filters are shown in Figure 2.1. The proposed filter is designed with $\beta = 4 \times 10^{-4}$, the magnitude weighting function given by

$$W_m(\omega) = \begin{cases} 
1.1, & \text{for } \omega \in [-\pi, -0.2\pi] \cup [0.4\pi, \pi], \\
1.0, & \text{for } \omega \in (-0.05\pi, 0.1\pi], \\
2.0, & \text{for } \omega \in [-0.1\pi, -0.05\pi] \cup (0.1\pi, 0.3\pi], 
\end{cases} \tag{2.17}$$

and the group delay weighting function given by

$$W_g(\omega) = \begin{cases} 
2, & \text{for } \omega \in [-0.1\pi, -0.05\pi] \cup (0.2\pi, 0.3\pi], \\
1, & \text{for } \omega \in (-0.05\pi, 0.2\pi].
\end{cases}$$

For comparison, we also design two filters using the conventional SDP method. ‘Filter 1’ is designed with the unity magnitude weighting function $W_m(\omega)$, while ‘Filter 2’ is designed with $W_m(\omega)$ in (2.17). Note that ‘Filter 1’ is the result presented in [CT04].

Some important design parameters, such as passband ripple $\delta_p$, stopband attenuation $\delta_s$, and maximum absolute group delay error $\max |E_\tau(\omega)|$, with $E_\tau(\omega) = \tau(\omega) - \tau_{des}$, are measured and tabulated in Table 2.1. Compared with ‘Filter 1’, the filter designed by the proposed method achieves better passband ripple and poorer stopband attenuation. However, our proposed filter outperforms ‘Filter 1’ in the group delay with a reduction of about 35% in $\max |E_\tau(\omega)|$. Compared with ‘Filter 2’, the performance of the filter obtained by the proposed method is much better in group delay with a reduction of about 25% in $\max |E_\tau(\omega)|$, while the magnitude response in the passband is about the
Table 2.2: Design results of the complex FIR filters in Example 2.2.

| Design methods                        | max $|E_{\tau}(\omega)|$ | $\delta_p$(dB) | $\delta_s$(dB) |
|---------------------------------------|--------------------------|----------------|----------------|
| Minimax with unity $W_m(\omega)$ (Filter 1) | 0.7337                  | 0.1268         | -41.3279       |
| Minimax with $W_m(\omega)$ in (2.18) (Filter 2) | 0.7907                  | 0.1650         | -42.9363       |
| Proposed method                       | 0.4744                   | 0.1509         | -42.1034       |

same but with an increase of about 5% in stopband error.

### 2.5.2 Example 2.2: Complex Bandpass Filter

A complex bandpass FIR filter of length $N = 45$ with cutoff frequencies $\omega_{s1} = 0.1\pi$, $\omega_{p1} = 0.2\pi$, $\omega_{p2} = 0.6\pi$ and $\omega_{s2} = 0.7\pi$, and constant passband group delay $\tau_{des} = 15$ is considered here. The numbers of frequency grids are $m_K = 121$ and $g_K = 27$.

The magnitude responses and the group delays of the designed filters are shown in Figure 2.2. The magnitude response and the group delay of the filter designed using the proposed method are indicated by solid lines in the figures. This filter is designed with $\beta = 8 \times 10^{-5}$, the magnitude weighting function given by

$$W_m(\omega) = \begin{cases} 
1 & \text{for } \omega \text{ in the passband,} \\
1.4 & \text{for } \omega \text{ in the stopband,}
\end{cases} \quad (2.18)$$

and the group delay weighting function given by

$$W_g(\omega) \equiv 1 \text{ for } \omega \text{ in passband.}$$

As can be seen from Table 2.2, compared with ‘Filter 2’, the performance of the filter obtained by the proposed method is much better in group delay with a reduction of about 40% in max $|E_{\tau}(\omega)|$, while the magnitude response in the
2.5. Design Examples

Figure 2.2: Design results of the complex bandpass filter in Example 2.2. The magnitude response and group delay of the filter obtained by the proposed method, of ‘Filter 1’ and of ‘Filter 2’ are indicated by solid lines, broken lines and dash-dot lines, respectively. (a) magnitude response, (b) passband ripple, (c) stopband attenuation and (d) group delay.
2.5. Design Examples

Table 2.3: Design results of the complex FIR filters in Example 2.3.

| Design methods                  | max $|E_r(\omega)|$ | $\delta_p$(dB) | $\delta_s$(dB) |
|---------------------------------|----------------|---------------|---------------|
| Minimax with unity $W_m(\omega)$ (Filter 1) | 0.7849         | 0.1979        | -31.1200      |
| Minimax with $W_m(\omega)$ in (2.19) (Filter 2) | 0.7908         | 0.2086        | -31.3542      |
| Proposed method                 | 0.7779         | 0.2081        | -31.3531      |

passband is slightly better but with an increase of about 9% in stopband error.

2.5.3 Example 2.3: Complex Bandpass Filter with Very Low Delay

A complex bandpass FIR filter of length $N = 51$ with cutoff frequencies $\omega_{s1} = 0.1\pi$, $\omega_{p1} = 0.2\pi$, $\omega_{p2} = 0.6\pi$ and $\omega_{s2} = 0.7\pi$, and constant passband group delay $\tau_{des} = 5$ is considered here. Note that $\tau_{des} \ll \frac{N-1}{2}$ in this example. The numbers of frequency grids are the same with the ones in the last example.

The magnitude responses and the group delays of the designed filters are shown in Figure 2.3. The magnitude response and the group delay of the filter designed using the proposed method are indicated by solid lines in the figures. This filter is designed with $\beta = 1 \times 10^{-4}$, the magnitude weighting function given by

$$W_m(\omega) = \begin{cases} 
1 & \text{for } \omega \text{ in the passband,} \\
1.1 & \text{for } \omega \text{ in the stopband,}
\end{cases} \quad (2.19)$$

and the group delay weighting function given by

$$W_g(\omega) \equiv 1 \quad \text{for } \omega \text{ in passband.}$$

As can be seen from Table 2.3, compared with ‘Filter 2’, the performance
Figure 2.3: Design results of the complex bandpass filter in Example 2.3. The magnitude response and group delay of the filter obtained by the proposed method, of ‘Filter 1’ and of ‘Filter 2’ are indicated by solid lines, broken lines and dash-dot lines, respectively. (a) magnitude response, (b) passband ripple, (c) stopband attenuation and (d) group delay.
2.6. Discussions

of the filter obtained by the proposed method is slightly better in group delay
with a reduction of about 2% in max $|E_r(\omega)|$, while the magnitude responses
in both passband and stopband remain unchanged, which is attributed to the
very low group delay. Such design example is not very typical in filter design.

All three design examples also verify that although only the right hand
side inequality constraint in (2.12) is imposed, it achieves the same effect as if
both the left and right hand side inequality constraints are imposed since the
maximum absolute group delay error, max $|E_r(\omega)|$, is in fact contributed by the
maximum positive group delay error, max $E_r(\omega)$.

The proposed method has also been applied successfully in designing high-
pass complex FIR filters, and the MATLAB programs are available at: http://www.ntu.edu.sg/home/ezplin/SPL.htm.

2.6 Discussions

In order to formulate the group delay constraint as an LMI, some approxima-
tions have been used in the previous sections. For completeness, the approxi-
mations are justified and investigated in this section.

2.6.1 On Eigenvalues of $\tilde{P}(\omega)$

To justify the validity of approximating $\tilde{P}(\omega)$ with $\hat{P}(\omega)$, we analyze the eigen-
values of $\tilde{P}(\omega)$ for any $\omega \in (-\pi, \pi]$ here. Denote $\lambda$ the eigenvalue of $\tilde{P}(\omega)$. For
2.6. Discussions

Figure 2.4: Eigenvalues of $\tilde{P}(\omega)$ for complex filters with various filter lengths.

Table 2.4: Non-zero eigenvalues of $\tilde{P}(\omega)$ in Figure 2.4.

| N  | $\nu$    | $\sigma$  | $\sigma/|\nu|$ |
|----|----------|-----------|----------------|
| 31 | -38.1959 | 503.1959  | 13.1741        |
| 45 | -79.8152 | 1069.8152 | 13.4037        |
| 80 | -250.1912| 3410.1912 | 13.6303        |
| 101| -397.8988| 5447.8988 | 13.6917        |
2.6. Discussions

\( N = 2 \), the characteristic equation of

\[
\tilde{P}(\omega) = \frac{1}{2} \begin{bmatrix}
0 & \cos \omega & 0 & \sin \omega \\
\cos \omega & 1 & -\sin \omega & 0 \\
0 & -\sin \omega & 0 & \cos \omega \\
\sin \omega & 0 & \cos \omega & 1
\end{bmatrix}
\]

is \( \det(\lambda I - \tilde{P}(\omega)) = 0 \), i.e.

\[
\det(\lambda I - \tilde{P}(\omega)) = \frac{1}{4} \lambda^2 - \lambda^3 + \frac{1}{4} \lambda \cos^2 \omega + \lambda^4 - \frac{1}{2} \lambda^2 \cos^2 \omega + \frac{1}{4} \lambda \sin^2 \omega \\
- \frac{1}{2} \lambda^2 \sin^2 \omega + \frac{1}{16} \cos^4 \omega + \frac{1}{8} \sin^2 \omega \cos^2 \omega + \frac{1}{16} \sin^4 \omega \\
= \lambda^4 - \lambda^3 + \frac{1}{4} \lambda^2 + \frac{1}{4} \lambda + \frac{1}{16} = (\lambda^2 - \lambda - \frac{1}{4})^2 = 0.
\]

We have \( \nu = \lambda_{1,2} = -0.2071 \) and \( \sigma = \lambda_{3,4} = 1.2071 \), which are frequency independent. Although we cannot find the analytic expression for the eigenvalues of \( \tilde{P}(\omega) \) for any arbitrary \( N \), exhaustive computer simulation is conducted to find the eigenvalues of \( \tilde{P}(\omega) \). Figure 2.4 presents the eigenvalues of \( \tilde{P}(\omega) \) for the complex filters with various filter lengths. It shows that for any arbitrarily fixed frequency \( \omega \), \( \tilde{P}(\omega) \) always has one positive eigenvalue \( \sigma \) of multiplicity of 2, and one negative eigenvalue \( \nu \) of multiplicity 2, both of which are frequency independent, and satisfy \( \sigma \gg \left| \nu \right| \), as seen from Table 2.4. Moreover, for \( \omega = 0 \), an analytic expression of the eigenvalues of \( \tilde{P}(\omega) \) is presented in the Appendix (Section 2.8).
2.6.2 Discrepancy in the Estimated Group Delay

The LMI constraint on group delay in (2.14) is obtained by two approximations, i.e., (1) replacing $|H(\omega)|$ with $|H_{\text{des}}(\omega)|$ and (2) replacing $\hat{\mathbf{P}}(\omega)$ with $\tilde{\mathbf{P}}(\omega)$. These approximations lead to some discrepancy in group delay.

Define the approximated group delay $\hat{\tau}(\omega) = \frac{\mathbf{h}_x^T \hat{\mathbf{P}}(\omega) \mathbf{h}_x}{|H_{\text{des}}(\omega)|^2}$, the actual group delay $\tau(\omega) = \frac{\mathbf{h}_x^T \mathbf{P}(\omega) \mathbf{h}_x}{|H(\omega)|^2}$ and their difference $\tau_\Delta(\omega) = \hat{\tau}(\omega) - \tau(\omega)$. The LMI constraint on group delay in (2.14) should be exactly expressed as

$$\hat{\Phi}_g(h_x, \omega) = \begin{bmatrix} |H(\omega)|^2 \left( \tau_{\text{des}} + \tau_\Delta(\omega) + \frac{\eta_g}{W_g(\omega)} \right) & \mathbf{h}_x^T \mathbf{B}(\omega) \\ \mathbf{B}^T(\omega) \mathbf{h}_x & \mathbf{I} \end{bmatrix} \succeq 0, \quad \text{for } \omega \in \Omega_g.$$  

(2.20)

The discrepancy $\tau_\Delta(\omega)$ can be reduced through a series of iterative updates with each carried out with the updated LMI constraint given by

$$\hat{\Phi}^{(i+1)}_g(h_x, \omega) = \begin{bmatrix} |H^{(i)}(\omega)|^2 \left( \tau_{\text{des}} + \mu \tau_\Delta^{(i)}(\omega) + \frac{\eta_g}{W_g(\omega)} \right) & \mathbf{h}_x^T \mathbf{B}(\omega) \\ \mathbf{B}^T(\omega) \mathbf{h}_x & \mathbf{I} \end{bmatrix} \succeq 0, \quad \text{for } \omega \in \Omega_g,$$  

(2.21)

where $\mu$ is a constant less than 1, $H^{(i)}(\omega)$ and $\tau_\Delta^{(i)}$ are evaluated with $h_x^{(i)}$. The iterative process can be described as

1. Let $i = 0$, $\mu = 0.7$, $|H^{(0)}(\omega)|^2 \equiv |H_{\text{des}}(\omega)|^2$ for $\omega \in \Omega_m$, and $\tau_\Delta^{(0)}(\omega) \equiv 0$ for $\omega \in \Omega_g$.

2. Design the filter using both the magnitude and group delay constraints in (2.15b) and (2.15c), resulting in $h_x^{(0)}$ and $\eta^{(0)}$. 

2.6. Discussions

Table 2.5: Non-zero eigenvalues of $\tilde{P}_1(\omega)$ ($\omega = 0.12\pi$) in Figure 2.5.

| $N$ | $\nu_{1,2}$ | $\sigma_{1,2}$ | $\min \sigma_{1,2}/|\max \nu_{1,2}|$ |
|-----|-------------|---------------|------------------|
| 31  | -22.2496, -15.6222 | 228.1187, 274.7532 | 12.3487 |
| 45  | -46.6252, -33.0809 | 498.9387, 570.7675 | 12.2416 |
| 80  | -136.4155, -113.6486 | 1640.3733, 1769.6909 | 12.9728 |
| 101 | -206.3361, -191.0405 | 2657.0172, 2790.3593 | 13.5234 |

3. Let $i = i + 1$ and evaluate $|H^{(i)}(\omega)|^2$ and $\tau^{(i)}(\omega)$ using $h_x^{(i-1)}$.

4. Design the filter using both the magnitude and group delay constraints in (2.15b) and (2.21), leading to $h_x^{(i)}$ and $\eta^{(i)}$.

5. Go to (3) if $\kappa = \frac{|\eta^{(i)} - \eta^{(i-1)}|}{\eta^{(i)} + \eta^{(i-1)}} > 0.01$ or the predefined number of iterations is reached.

The effectiveness of the above iterative method will be demonstrated through an example of real FIR filter design in the next subsection.

2.6.3 On the Design of Real FIR Filters

Since real FIR filters can be considered as a special case of complex filters, i.e., $h_i = 0$, the proposed design method for real FIR filters is briefly discussed in the following. For the real FIR filters, the group delay is

$$\tau(\omega) = \frac{h_x^T \tilde{P}_1(\omega) h_x}{|H(\omega)|^2} = \frac{h_x^T \tilde{P}_1(\omega) h_x}{|H(\omega)|^2}$$  (2.22)
Figure 2.5: Eigenvalues of $\tilde{P}_1(\omega)$ for real filters with various filter lengths.
where $\tilde{P}_1(\omega) = \frac{1}{2}(P_1(\omega) + P_1^T(\omega))$ is a symmetric $N \times N$ matrix given by

$$
\tilde{P}_1(\omega) = \frac{1}{2} \begin{bmatrix}
0 & \cos \omega & 2 \cos 2\omega & \ldots & (N-1) \cos(N-1)\omega \\
\cos \omega & 2 & 3 \cos \omega & \ldots & N \cos(N-2)\omega \\
2 \cos 2\omega & 3 \cos \omega & 4 & \ldots & (N+1) \cos(N-3)\omega \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
(N-1) \cos(N-1)\omega & \ldots & \ldots & \ldots & 2N - 2
\end{bmatrix}.
$$

For $N = 2$, we have

$$
\tilde{P}_1(\omega) = \frac{1}{2} \begin{bmatrix}
0 & \cos \omega \\
\cos \omega & 2
\end{bmatrix}
$$

and the eigenvalues of $\tilde{P}_1$ satisfy $\lambda^2 - \lambda - \frac{1}{4} \cos^2 \omega = 0$. Therefore we have $\lambda_{1,2} = \frac{1}{2}(1 \pm \sqrt{1 + \cos^2 \omega})$, which are frequency dependent.

The eigenvalues of $\tilde{P}_1(\omega)$ for real FIR filters of different lengths are exhaustively simulated and some examples are presented in Figure 2.5. Noted that $\tilde{P}_1(\omega)$ in general has two distinct positive eigenvalues $\{\sigma_1, \sigma_2\}$ and two distinct negative ones $\{\upsilon_1, \upsilon_2\}$ for $N \geq 4$, all of which are frequency dependent, as showed in Figure 2.5 and Table 2.5. Similarly to $\tilde{P}(\omega)$ for complex FIR filters, we also observe that $\min\{\sigma_1, \sigma_2\} \gg \max\{|\upsilon_1|, |\upsilon_2|\}$.

**Example 2.4: Real FIR Filter Design**

A lowpass FIR filter of length $N = 31$ with cutoff frequencies $\omega_p = 0.12\pi$ and $\omega_s = 0.24\pi$, and constant passband group delay $\tau_{des} = 12$ in the passband is considered here [PS92]. The number of frequency grids are $m_K = 120$ and...
2.6. Discussions

Figure 2.6: Design results of the real lowpass filter in Example 2.4. The magnitude response and group delay of the filters obtained by the proposed methods with iteration, without iteration, and the conventional SDP method are indicated by solid lines, broken lines and dash-dot lines, respectively. (a) magnitude response, (b) passband ripple, (c) stopband attenuation and (d) group delay.

| Design methods                              | max $|E_r(\omega)|$ | $\delta_p$(dB) | $\delta_s$(dB) |
|---------------------------------------------|------------------|----------------|----------------|
| Minimax with $W_m(\omega)$ in (2.23)       | 0.3388           | 0.0356         | -31.1643       |
| Proposed method without iteration           | 0.1405           | 0.0446         | -28.1186       |
| Proposed method with iteration              | 0.2619           | 0.0386         | -31.0966       |
$g_K = 80$.

The proposed filters with and without iteration are designed with $\beta = 3 \times 10^{-3}$, the magnitude weighting function given by

$$W_m(\omega) = \begin{cases} 
5.0, & \text{for } \omega \text{ in the passband}, \\
1.0, & \text{for } \omega \text{ in the stopband},
\end{cases} \quad (2.23)$$

and the group delay weighting function given by $W_g(\omega) \equiv 1$. It takes 4 iterations for the proposed method with iteration to converge. The magnitude responses and the group delays of the designed filters are shown in Figure 2.6. For comparison, we also design one filter using the conventional SDP method with $W_m(\omega)$ in (2.23).

As tabulated in Table 2.6, compared with the filter obtained by the conventional SDP method, the proposed one with iteration is slightly poorer in magnitude response with about 1% increase in stopband attenuation but is much better in the group delay with a reduction of about 20% in $\max |E_\tau(\omega)|$. Compared with the proposed filter without iteration, the ripple magnitude is decreased at the cost of the increased group delay error.

### 2.7 Conclusions

In this chapter, we have proposed a new design method for FIR filters with reduced group delays. To take the advantage of the flexibility of the SDP formulation, a group delay constraint is formulated to an LMI constraint with some reasonable approximations. Subsequently, both the magnitude and the group delay constraints are successfully incorporated into the design problem.
As shown by the examples, the group delay error can be effectively reduced as the expense of slight increase in ripple magnitude. These approximations are shown to be reasonable through extensive computer simulations and some analytical derivations.

2.8 Appendix: Analytic Expression for the Eigenvalues of $\tilde{P}(0)$

For a complex FIR filter of length $N$, $\tilde{P}_i(\omega)$ (i=1, \ldots, 4) are

$$
\tilde{P}_1(\omega) = \tilde{P}_4(\omega) = \frac{1}{2} \begin{bmatrix}
0 & \cos \omega & 2 \cos 2\omega & 3 \cos 3\omega & \ldots & (N - 1) \cos(N - 1)\omega \\
\cos \omega & 2 & 3 \cos \omega & 4 \cos 2\omega & \ldots & N \cos(N - 2)\omega \\
2 \cos 2\omega & 3 \cos \omega & 4 & 5 \cos \omega & \ldots & (N + 1) \cos(N - 3)\omega \\
3 \cos 3\omega & 4 \cos 2\omega & 5 \cos \omega & 6 & \ldots & (N + 2) \cos(N - 4)\omega \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
(N - 1) \cos(N - 1)\omega & \ldots & \ldots & \ldots & \ldots & 2N - 2
\end{bmatrix}
$$

$$
\tilde{P}_2(\omega) = \tilde{P}_3^T(\omega) = \frac{1}{2} \begin{bmatrix}
0 & -\sin \omega & -2 \sin 2\omega & -3 \sin 3\omega & \ldots & -(N - 1) \sin(N - 1)\omega \\
\sin \omega & 0 & -3 \sin \omega & -4 \sin 2\omega & \ldots & -(N + 1) \sin(N - 2)\omega \\
2 \sin 2\omega & 3 \sin \omega & 0 & -5 \sin \omega & \ldots & -(N + 1) \sin(N - 3)\omega \\
3 \sin 3\omega & 4 \sin 2\omega & 5 \sin \omega & 0 & \ldots & -(N + 2) \sin(N - 4)\omega \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
(N - 1) \sin(N - 1)\omega & \ldots & \ldots & \ldots & \ldots & 0
\end{bmatrix}
$$
As a special case when $\omega = 0$, we have $\tilde{P}_2(0) = \tilde{P}_3(0) = 0$ and

$$\tilde{P}(0) = \begin{bmatrix} \tilde{P}_1(0) & 0 \\ 0 & \tilde{P}_1(0) \end{bmatrix}$$

Let $A = \tilde{P}_1(0)$. The characteristic equation of $\tilde{P}(0)$ is

$$\det(\lambda I_{2N} - \tilde{P}(0)) = \det(\lambda I_N - A)^2 = 0,$$

which means the eigenvalues of $\tilde{P}(0)$ are the same as those of $A$ with multiplicity of 2.

Next, we are to find the eigenvalues of $A$. We first show that $\text{Rank}(A) = 2$. Given

$$A = \frac{1}{2} \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & \ldots & N - 1 \\ 1 & 2 & 3 & 4 & 5 & \ldots & N \\ 2 & 3 & 4 & 5 & 6 & \ldots & N + 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ N - 1 & N & N + 1 & N + 2 & N + 3 & \ldots & 2N - 2 \end{bmatrix},$$

by subtracting the $m$-th row from the $(m+1)$-th row for $m = N-1, N-2, \ldots, 1$, we have

$$A' = \frac{1}{2} \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & \ldots & N - 1 \\ 1 & 1 & 1 & 1 & 1 & \ldots & 1 \\ 1 & 1 & 1 & 1 & 1 & \ldots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & 1 & \ldots & 1 \end{bmatrix}.$$
Hence, \( \text{Rank}(A) = \text{Rank}(A') = 2 \).

An explicit form of the characteristic equation is given by [Gan59]:

\[
\det (\lambda I_N - A) = \lambda^N - S_1 \lambda^{N-1} + S_2 \lambda^{N-2} - \ldots + (-1)^N S_N = 0, \tag{2.25}
\]

where \( S_n \) \((n = 1, 2, \ldots, N)\) is the sum of the principal minors of order \( n \) of \( A \).

We have

\[
S_1 = \sum_{n=0}^{N-1} n = \frac{1}{2} N(N-1) \tag{2.26}
\]

\[
S_2 = \frac{1}{4} \begin{vmatrix} 0 & 1 & 1 & 1 & \ldots & 1 \\ 1 & 2 & 3 & 4 & \ldots & N-1 \\ 2 & 4 & 6 & 8 & \ldots & 2N-2 \\ 3 & 6 & \ldots & 2N-3 \end{vmatrix}
\]

\[
= \frac{1}{4} \left( 1^2 + 2^2 + 3^2 + \ldots + (N-1)^2 \right) - \frac{1}{4} \left( 1^2 + 2^2 + \ldots + (N-2)^2 \right) - \ldots - \frac{1}{4} \times 1^2
\]

\[
= -\frac{1}{4} \left[ (N-1) \left( 1^2 + 2^2 + 3^2 + \ldots + (N-1)^2 \right) - \sum_{n=1}^{N-2} n(n+1)^2 \right]
\]

\[
= -\frac{1}{4} \left[ \frac{1}{6} (N-1)^2 N(2N-1) - \frac{1}{12} (N-2)(N-1)N(3N-1) \right]
\]

\[
= -\frac{1}{48} N^2 (N^2 - 1), \tag{2.27}
\]
using the identities

\[
\sum_{n=1}^{N} n^2 = \frac{1}{6} N(N + 1)(2N + 1),
\]
\[
\sum_{n=1}^{N} n(n + 1)^2 = \frac{1}{12} N(N + 1)(N + 2)(3N + 5).
\]

Since \( \text{Rank}(A) = 2 \), \( S_n = 0 \) for \( n \geq 3 \). Hence, the characteristic equation in (2.25) becomes

\[
\lambda^N - S_1 \lambda^{N-1} + S_2 \lambda^{N-2} = 0.
\]

(2.28)

Substituting (2.26) and (2.27) in (2.28) gives

\[
\lambda^{N-2} \left( \lambda^2 - \frac{1}{2} N(N - 1) \lambda - \frac{1}{48} N^2(N^2 - 1) \right) = 0.
\]

Therefore the eigenvalues of \( A (\tilde{P}_1(0)) \) are given by

\[
\lambda_{1,2} = \frac{1}{2} \left( \frac{1}{2} N(N - 1) \pm \sqrt{\frac{1}{4} N^2(N - 1)^2 + \frac{1}{12} N^2(N^2 - 1)} \right)
\]
\[
= \frac{1}{2} \left( \frac{1}{2} N(N - 1) \pm \sqrt{\frac{1}{12} N^2[3(N - 1)^2 + N^2 - 1]} \right)
\]
\[
= \frac{N}{4} \left( N - 1 \pm \sqrt{\frac{4N^2 - 6N + 2}{3}} \right),
\]

\( \lambda_n = 0, \) for \( n = 3, \ldots, N. \)

For example, for \( N = 2 \), \( \lambda_{1,2} = \{-0.2071, 1.2071\} \), which agree with those given in Section 2.6. For \( N = 45 \), \( \lambda_{1,2} = \{-79.8152, 1069.8152\} \), which also agree with the simulation results in Table 2.4.
Chapter 3

FRM Filter Design with Group Delay Constraint

3.1 Introduction

Compared with conventional methods for designing direct-form FIR filters, such as the Park-McClellan algorithm [PM72], the least-squares method [TF70, Kel72], and the eigenfilter approach [VN87, Ngu93, PT01, TVN03], the frequency-response masking (FRM) technique, invented by Y. C. Lim in 1986 [Lim86], is well known to be more efficient in the implementation of FIR filters with sharp transition bands.

The design of an FRM filter proposed in [Lim86] started from its masking filters. Given a scaling factor, the bandedges of the masking filters can be determined. With ripples slightly smaller than those of the desired FRM filter, the masking filters can be designed using any conventional FIR filter design method. Subsequently, the prototype filter was designed such that the ripples
of the FRM filter around the transition band were minimum [Lim86]. The FRM filter designed in this way is suboptimum and a concept of overall optimization was firstly proposed in a two-step optimization technique [SYKJ03]. In the first step, a suboptimal FRM filter was obtained by iteratively and alternatively optimizing the prototype filter and masking filters, with each emphasizing on different frequency bands. In the second step, the suboptimal FRM filter was used as an initial solution for unconstrained nonlinear optimization. Using this method, the implementation complexity can be considerably reduced. Note that all these FRM filters as well as their subfilters [Lim86, SYKJ03] are of linear phase. Since a linear phase FRM filter is equivalent to a linear phase FIR filter that is slightly longer than a direct-form FIR filter designed by conventional methods with the same approximation accuracy [LL93], it has a large constant group delay in the entire frequency band.

As we have mentioned earlier, low passband group delay and low group delay error are required in certain applications for devices and digital filters to avoid long delay and minimize distortion of the transmitted information [Cho95, VWHvO98, BGCL+03]. In the context of FRM filters with reduced group delays, an intuitive design method was first proposed in [CL96]. The prescribed group delay was distributed to the ones of the subfilters such that they were inversely proportional to the corresponding transition bands. Similar to the design method in [Lim86], the masking filters and the prototype filter were designed individually and sequentially based on a weighted least-squares technique, leading to a suboptimal FRM filter.

The design of various FRM filters with different passband bandwidths and low delays was discussed [SJ02]. The prescribed group delay and the ripples of
the desired FRM filter were distributed among the subfilters, which were then
designed individually using linear programming (LP), resulting in a suboptimal
FRM filter. Inspired by the overall optimization proposed in [SYKJ03], the
suboptimal FRM filter was then used as an initial solution for the unconstraint
nonlinear optimization.

An alternative and more promising approach was proposed in [LH03b, LH03a]
in which all subfilter coefficients were stacked as a design vector and optimized
through a sequence of linear updates. With a reasonable initial design, the
overall optimization problem was formulated and carried out in semidefinite
programming (SDP) framework [LH03b] or second-order cone programming
(SOCP) framework [LH03a].

Similar to the design methods in [SCP93, KM95, KR95, KR96] for FIR filters
with reduced group delays, the prescribed group delays are indirectly achieved
and the complex magnitude errors in [CL96, SJ02, LH03b, LH03a], combining
both magnitude and phase errors, are minimized. Without any constraint in
group delay, it is not easy to control the group delay error, leading to relatively
large group delay error, noticeably towards the bandedges.

In this chapter, we propose an improved method for designing real FRM
filters with reduced group delays by directly imposing a group delay constraint.
The proposed design method may be considered as a generalization of the
method in [LH03b, LH03a] by introducing an additional group delay constraint
in the overall optimization. Similar to the FIR filter design formulation de-
scribed in Chapter 2, the filter design problem here is formulated by taking
into account both the complex magnitude error and the group delay error. The
key step is the derivation of the group delay of FRM filters and its gradient
with respect to the subfilter coefficients. Based on the derivation, a group delay constraint is formulated and incorporated into the overall optimization process. The optimization is formulated as an SOCP problem to reduce the design time. This is important since the optimization problem formulated as an SOCP problem can be solved more efficiently than in an equivalent SDP setting, especially when the dimension of the design vector increases (e.g., for FRM filters with nonlinear phase masking filters), or the number of frequency grids increases. Another advantage of the proposed method is that we address the design of general FRM filters with nonlinear phase masking filters. Although it was mentioned in [LH03b], only FRM filters with linear phase masking filters were considered there. As it can be seen from subsequent discussions, the design of FRM filters with nonlinear phase masking filters is much more complicated than that of FRM filters with linear phase masking filters. As will be shown by the design examples taken from [CL96] and [LH03b], starting with a reasonable initial design, the proposed method converges efficiently to an optimized FRM filter with considerable reduction in group delay error at the expense of slight increase in magnitude error.

The rest of this chapter is organized as follows. Prior to the presentation of the proposed design method, the FRM technique and SOCP are briefly reviewed in Section 3.2. Without loss of generality, all subfilters are assumed to be of nonlinear phase. The group delay of a basic FRM filter and its gradient with respect to the subfilter coefficients are derived in Section 3.3, where the derivation for a basic FRM filter with linear phase masking filters is also given as a special case. In Section 3.4, the proposed design method for basic FRM filters is described and the group delay constraint is formulated and incorporated into the overall optimization that takes into account both complex magnitude error
and group delay error. The derivation for the group delay and its gradient for multi-stage FRM filters is presented in Section 3.5. The effectiveness of the proposed design method for designing various FRM filters is illustrated by three examples in Section 3.6. Conclusions are given in Section 3.7, followed by some expressions for the gradient of the group delay for a two-stage FRM filter in the Appendix (Section 3.8).

Throughout this chapter, we assume that the lengths of the subfilters are all odd. Moreover, for the sake of notation simplicity, the frequency variable $\omega$ is omitted in the following derivation unless otherwise specified.

### 3.2 Preliminaries

Since the FRM technique is used in both Chapters 3 and 4, and SOCP is used in Chapters 3 and 5, a brief review on the FRM technique and SOCP is presented in this section.

#### 3.2.1 FRM Technique

![Figure 3.1: A basic FRM filter structure.](image)
3.2. Preliminaries

A basic FRM filter consists of a prototype filter $H_a(z)$, a pair of masking filters $\{H_{ma}(z), H_{mc}(z)\}$, and a delay line that together with the prototype filter, forms a complementary pair $\{H_a(z^M), H_c(z^M)\}$, as illustrated in Figure 3.1 [Lim86], where $M$ and $d_a$ are the scaling factor and the group delay of $H_a(z)$, respectively. For additional reduction of realization complexity, the prototype filter may be realized with a second basic FRM filter, leading to a two-stage FRM filter [Lim86], see Figure 3.2, where $M_1, M_2$ are the scaling factors of stage 1 and 2, respectively, and $d_{m2}$ is the group delay of $\{H_{ma2}(z), H_{mc2}(z)\}$. The process can be repeated for a general multi-stage FRM filter.

For a basic FRM filter, the transfer functions of the subfilters are given by

$$H_a(z) = \sum_{n=0}^{N_a-1} h_a(n)z^{-n},$$

$$H_{ma}(z) = \sum_{n=0}^{N_{ma}-1} h_{ma}(n)z^{-n},$$

$$H_{mc}(z) = \sum_{n=0}^{N_{mc}-1} h_{mc}(n)z^{-n}. \quad (3.1)$$
As we know, the transition bandwidth of $H_a(z^M)$ is $1/M$ of that of $H_a(z)$ and multiple spectrum replicas appear in $[0, 2\pi)$. Using the properly designed masking filters, some unwanted spectrum replicas are removed while the rest are integrated in passband synthesis. Let $N_m = \max\{N_{ma}, N_{mc}\}$ and denote $d_m$ the group delay of the masking filters. For a basic linear phase FRM filter, whose subfilters are all of linear phase, we have $d_a = \frac{N_a-1}{2}$ and $d_m = \frac{N_m-1}{2}$. The frequency response and the group delay of the linear phase FRM filter are

\[
H_f(\omega) = H_a(M\omega)H_{ma}(\omega) + \left(e^{-j\omega M \frac{N_a-1}{2}} - H_a(M\omega)\right)H_{mc}(\omega),
\]

\[
\tau_f(\omega) = Md_a + d_m = M\frac{N_a - 1}{2} + \frac{N_m - 1}{2}.
\]

To design an linear phase FRM filter, the following conditions must satisfied

1. all subfilters are linear phase filters;
2. $M(N_a - 1)$ is even;
3. the lengths of both masking filters are even or odd [Lim86].

To design a basic linear phase FRM filter with the desired frequency response, $H_{des}(\omega)$, the passband cutoff frequency, $\omega_p$, and the stopband cutoff frequency, $\omega_s$, the following steps are involved [Lim86]:

1. Given $M$, the passband and stopband cutoff frequencies of the subfilters may be determined from one of the two sets given below, depending on which set of $\{\omega_{ap}, \omega_{as}\}$ satisfies $0 < \omega_{ap} < \omega_{as} < \pi$.

   where $[x]$ and $\lceil x \rceil$ denote the largest integer less than $x$ and the smallest integer larger than $x$, respectively.

2. With ripples about $10\% - 15\%$ smaller than those of the desired FRM filter and the cutoff frequencies given in Table 3.1, the linear phase masking filters, $H_{ma}(z)$ and $H_{mc}(z)$, can be designed using any conventional FIR
3.2. Preliminaries

Table 3.1: The passband and stopband cutoff frequencies of the subfilters

<table>
<thead>
<tr>
<th></th>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>$m = \lfloor \omega_p M / 2 \pi \rfloor$</td>
<td>$m = \lfloor \omega_s M / 2 \pi \rfloor$</td>
</tr>
<tr>
<td>$H_a(z)$</td>
<td>$\omega_{ap} = \omega_p M - 2m\pi$</td>
<td>$\omega_{ap} = 2m\pi - \omega_s M$</td>
</tr>
<tr>
<td></td>
<td>$\omega_{as} = \omega_s M - 2m\pi$</td>
<td>$\omega_{as} = 2m\pi - \omega_p M$</td>
</tr>
<tr>
<td>$H_{ma}(z)$</td>
<td>$\omega_{map} = \frac{2m\pi + \omega_{ap}}{M}$</td>
<td>$\omega_{map} = \frac{2(m-1)\pi + \omega_{as}}{M}$</td>
</tr>
<tr>
<td></td>
<td>$\omega_{mas} = \frac{2(m+1)\pi - \omega_{as}}{M}$</td>
<td>$\omega_{mas} = \frac{2m\pi - \omega_{ap}}{M}$</td>
</tr>
<tr>
<td>$H_{mc}(z)$</td>
<td>$\omega_{mcp} = \frac{2m\pi - \omega_{ap}}{M}$</td>
<td>$\omega_{mcp} = \frac{2m\pi - \omega_{as}}{M}$</td>
</tr>
<tr>
<td></td>
<td>$\omega_{mcs} = \frac{2m\pi + \omega_{as}}{M}$</td>
<td>$\omega_{mcs} = \frac{2m\pi + \omega_{ap}}{M}$</td>
</tr>
</tbody>
</table>

filter design method.

3. Since $H_{des}(\omega)$ and all subfilters in a linear phase FRM filter are of linear phase, zero-phase responses of them are used for simplicity. The frequency response in (3.2) is simplified to the zero-phase response given by

$$G_f(\omega) = G_a(M\omega) (G_{ma}(\omega) - G_{mc}(\omega)) + G_{mc}(\omega),$$

where $G_a(\omega)$, $\{G_{ma}(\omega), G_{mc}(\omega)\}$ and $G_f(\omega)$ are the zero-phase responses of the prototype filter, masking filters and the FRM filter, respectively.

The prototype filter $H_a(z)$ can then be optimized such that the ripples are minimized around the transition band, i.e., $\frac{2m\pi - \omega_{ap}}{M} < \omega < \frac{2(m-1)\pi + \omega_{as}}{M}$ for Case I and $\frac{2(m+1)\pi - \omega_{as}}{M} < \omega < \frac{2m\pi + \omega_{ap}}{M}$ for Case II.

3.2.2 SOCP

Second-order cone program (SOCP) is a class of convex programming problems where a linear function is minimized over the intersection of an affine set and the product of second-order cones [LVBL98]. They can be generally expressed
as

\[
\text{minimize } \quad f^T x \\
\text{subject to : } \quad \|A_i x + b_i\| \leq c_i^T x + d_i, \quad i = 1, 2, \ldots, N
\]

where \(x \in \mathbb{R}^{n \times 1}, \ f \in \mathbb{R}^{n \times 1}, \ A_i \in \mathbb{R}^{(n_i-1) \times n}, \ b_i \in \mathbb{R}^{(n_i-1) \times 1}, \ c_i \in \mathbb{R}^{n \times 1}, \ d_i \in \mathbb{R},\)

and \(\| \cdot \|\) denotes the standard Euclidean norm. The norm appearing in (3.4b) is a standard Euclidean norm and each constraint in (3.4b) is a second-order cone constraint of dimension \(n_i\) because the set of points satisfying a second-order cone constraint is the inverse image of the unit second-order cone under an affine mapping:

\[
\|A_i x + b_i\| \leq c_i^T x + d_i \iff [A_i; \ c_i^T] x + [b_i; \ d_i] \in C_{n_i},
\]

where \(C_{n_i}\) is a second-order cone in \(\mathbb{R}^{n_i}, \ i.e.,\)

\[
C_{n_i} = \left\{ \begin{bmatrix} t & u \end{bmatrix} : u \in \mathbb{R}^{(n_i-1) \times 1}, t \geq 0, \|u\| \leq t \right\}.
\]

Second-order cone constraints can be used to represent several common convex constraints. For example, they reduce to the linear constraints when \(n_i = 1\) for \(i = 1, \ldots, N\); they reduce to quadratic constraint \(\|A_i x + b_i\|^2 \leq d_i^2\) when \(c_i = 0\). Thus quadratic programs, quadratically constrained quadratic programs, and many other nonlinear convex optimization problems can be reformulated as second-order cone programs.

Regarding the place of SOCP in convex optimization relative to other program classes, SOCP includes several important standard classes of convex opti-
mization problems, such as linear programs, quadratic programs and quadrat-
cally constrained quadratic programs. On the other hand, it is less general than
SDP, because the second-order cone can be embedded in the cone of positive
semidefinite matrices since

\[ \|u\| \leq t \iff \begin{bmatrix} tI & u \\ x^T & t \end{bmatrix} \succeq 0, \quad (3.7) \]

i.e., a second-order cone constraint is equivalent to a linear matrix inequality.
Using this property, the SOCP in (3.4) can be expressed as an SDP as

\[
\begin{align*}
\text{minimize} & \quad f^T x \\
\text{subject to} : & \quad \begin{bmatrix} (c_i^T x + d_i)I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \succeq 0, \quad i = 1, \ldots, N. \quad (3.8b)
\end{align*}
\]

However solving SOCPs via SDPs is not a good idea. Interior-point methods
have a better worst-case complexity when applying in an SOCP setting than
in an equivalent SDP setting: (1) the number of iterations to decrease
the duality gap to a constant fraction of itself is bounded above by \( O(\sqrt{N}) \)
for the SOCP algorithm, and by \( O(\sqrt{\sum_i n_i}) \) for the SDP algorithm, where \( O() \)
means the order of the computational complexity. More importantly in practice,
each iteration is much faster; (2) the amount of computation per iteration is
\( O(n^2 \sum_i n_i) \) in the SOCP algorithm and \( O(n^2 \sum_i n_i^2) \) for the SDP algorithm.
The difference between these numbers is significant if the dimensions \( n_i \) of the
second-order constraints are large [LVBL98].

Many interior-point methods are applicable to solve SOCP, e.g. robust con-
trol toolbox [BCPS05], SeDuMi[Stu99], and Yalmip[Yal05].
3.3 Group Delay and Gradients for Basic FRM Filters

3.3.1 Group Delay for Basic FRM Filters

For the sake of convenience in the derivation of many expressions presented in this chapter, we adopt a simplified notation for the prototype filter and masking filters for a basic FRM filter, as shown in Figure 3.3. Let \( h_1 = [h_1(0), h_1(1), \ldots, h_1(N_1 - 1)] \), \( h_3 = [h_3(0), h_3(1), \ldots, h_3(N_3 - 1)] \) and \( h_4 = [h_4(0), h_4(1), \ldots, h_4(N_4 - 1)] \) denote the vectors of the coefficients for the prototype filter and the masking filters, which correspond to \( h_a \), \( h_{ma} \) and \( h_{mc} \) (\( H_a(z) \), \( H_{ma}(z) \) and \( H_{mc}(z) \)) in Figure 3.1, respectively.

Let \( N_m \) be the maximum of the lengths of the masking filters, i.e., \( N_m = \max\{N_3, N_4\} \). Denote \( d_a \) and \( d_m \) the group delays of the prototype filter and the masking filters, respectively. A dummy vector \( h_2 = u_{d_a+1} - h_1 \) of dimension \( N_2 = N_1 \), is introduced for the complimentary filter of \( H_1(z) \) in Figure 3.3, where \( u_{d_a+1} \) is an all-zero vector except having 1 at the \((d_a+1)\)-th entry. Thus we have \( H_2(z) = z^{-d_a} - H_1(z) \).
In order to derive a general expression of the group delay of a basic FRM filter, we assume that all subfilters are of nonlinear phase, and magnitude and phase responses of the subfilters and at various nodes in Figure 3.3 are examined.

Let $A_1(\omega)e^{j\phi_1(\omega)}$, $A_2(\omega)e^{j\phi_2(\omega)}$, $A_3(\omega)e^{j\phi_3(\omega)}$ and $A_4(\omega)e^{j\phi_4(\omega)}$ be the frequency responses of $H_1(z^M)$, $H_2(z^M)$, $H_3(z)$ and $H_4(z)$, respectively. The magnitude and phase responses are given by

$$A_i(\omega) = \sqrt{(h_i^Tc_i)^2 + (h_i^Ts_i)^2}, \quad \phi_i(\omega) = \arctan 2(h_i^Ts_i, h_i^Tc_i), \quad i = 1, \ldots, 4,$$

where $c_i$ and $s_i$ are trigonometric vectors of the same dimension of $h_i$ as

$$c_i = \begin{bmatrix} 1 & \cos M\omega & \cos 2M\omega & \ldots & \cos(N_i - 1)M\omega \end{bmatrix}^T, \quad i = 1, 2, \quad (3.9a)$$

$$s_i = \begin{bmatrix} 0 & -\sin M\omega & -\sin 2M\omega & \ldots & -\sin(N_i - 1)M\omega \end{bmatrix}^T, \quad i = 1, 2, \quad (3.9b)$$

$$c_l = \begin{bmatrix} \cos\left(\frac{N_m}{2} - \frac{N_i}{2}\right)\omega & \cos\left(\frac{N_m}{2}\right)\omega & \ldots & \cos\left(\frac{N_m}{2} + \frac{N_i}{2}\right)\omega \end{bmatrix}^T, \quad l = 3, 4, \quad (3.9c)$$

$$s_l = \begin{bmatrix} -\sin\left(\frac{N_m}{2} - \frac{N_i}{2}\right)\omega & -\sin\left(\frac{N_m}{2}\right)\omega & \ldots & -\sin\left(\frac{N_m}{2} + \frac{N_i}{2}\right)\omega \end{bmatrix}^T, \quad l = 3, 4. \quad (3.9d)$$

The frequency responses at nodes ‘d’, ‘e’ and ‘f’ can be obtained as

$$A_d e^{j\phi_d} = A_1 A_3 e^{j(\phi_1 + \phi_3)},$$

$$A_e e^{j\phi_e} = A_2 A_4 e^{j(\phi_2 + \phi_4)},$$

$$A_f e^{j\phi_f} = A_d e^{j\phi_d} + A_e e^{j\phi_e} = (A_d \cos \phi_d + A_e \cos \phi_e) + j(A_d \sin \phi_d + A_e \sin \phi_e).$$
3.3. Group Delay and Gradients for Basic FRM Filters

Let $N_f = A_d \sin \phi_d + A_e \sin \phi_e$ and $D_f = A_d \cos \phi_d + A_e \cos \phi_e$. The phase response of the basic FRM filter is

$$\phi_f = \arctan \frac{N_f}{D_f}.$$  

Using (2.2c) and (2.3), the group delay of the FRM filter is

$$\tau_f(\omega) = -\frac{d}{d\omega} \left( \tan^{-1} \frac{N_f}{D_f} \pm \pi \right) = -\frac{1}{D_f^2 + N_f^2} \left( \frac{dN_f}{d\omega} D_f - N_f \frac{dD_f}{d\omega} \right)$$  

(3.10)

where

$$\frac{dN_f}{d\omega} = \frac{dA_d}{d\omega} \sin \phi_d + A_d \frac{d\phi_d}{d\omega} + \frac{dA_e}{d\omega} \sin \phi_e + A_e \frac{d\phi_e}{d\omega},$$  

(3.11a)

$$\frac{dD_f}{d\omega} = \frac{dA_d}{d\omega} \cos \phi_d - A_d \frac{d\phi_d}{d\omega} + \frac{dA_e}{d\omega} \cos \phi_e - A_e \frac{d\phi_e}{d\omega}.$$  

(3.11b)

$\frac{dA_d}{d\omega}$, $\frac{dA_e}{d\omega}$, $\frac{d\phi_d}{d\omega}$ and $\frac{d\phi_e}{d\omega}$ can each be obtained as a function of $\frac{dA_i}{d\omega}$ and $\frac{d\phi_i}{d\omega}$ ($i = 1, \ldots, 4$) given by

$$\frac{dA_d}{d\omega} = \frac{dA_1}{d\omega} A_3 + A_1 \frac{dA_3}{d\omega}, \quad \frac{d\phi_d}{d\omega} = \frac{d\phi_1}{d\omega} + \frac{d\phi_3}{d\omega},$$  

(3.12a)

$$\frac{dA_e}{d\omega} = \frac{dA_2}{d\omega} A_4 + A_2 \frac{dA_4}{d\omega}, \quad \frac{d\phi_e}{d\omega} = \frac{d\phi_2}{d\omega} + \frac{d\phi_4}{d\omega},$$  

(3.12b)

where

$$\frac{dA_i}{d\omega} = \frac{1}{A_i} h_i^T (c_i \tilde{c}_i^T + s_i \tilde{s}_i^T) h_i = \frac{1}{A_i} h_i^T \tilde{P}_{mwi} h_i, \quad i = 1, \ldots, 4,$$

$$\frac{d\phi_i}{d\omega} = \frac{1}{A_i} h_i^T (\tilde{s}_i c_i^T - \tilde{c}_i s_i^T) h_i = \frac{1}{A_i} h_i^T \tilde{P}_{pwi} h_i, \quad i = 1, \ldots, 4,$$

with $\tilde{s}_i = \frac{ds_i}{d\omega}$ and $\tilde{c}_i = \frac{dc_i}{d\omega}$. Both $\tilde{P}_{mwi}$ and $\tilde{P}_{pwi}$ are square matrices of compatible dimensions with $h_i$. Similarly to the technique used in Chapter 2, it can be
3.3. Group Delay and Gradients for Basic FRM Filters

easily shown that

\[
\frac{dA_i}{d\omega} = \frac{1}{A_i} h_i^T P_{m\omega i} h_i, \quad i = 1, \ldots, 4, \quad (3.13a)
\]
\[
\frac{d\phi_i}{d\omega} = \frac{1}{A_i^2} h_i^T P_{p\omega i} h_i, \quad i = 1, \ldots, 4, \quad (3.13b)
\]

where \( P_{m\omega i} \) and \( P_{p\omega i} \) are symmetric matrices obtained by

\[
P_{m\omega i} = \frac{1}{2} (\tilde{P}_{m\omega i} + \tilde{P}_{m\omega i}^T), \quad i = 1, \ldots, 4, \quad (3.14a)
\]
\[
P_{p\omega i} = \frac{1}{2} (\tilde{P}_{p\omega i} + \tilde{P}_{p\omega i}^T), \quad i = 1, \ldots, 4. \quad (3.14b)
\]

The analytic expression of the group delay of a basic FRM filter given in (3.10)-(3.14) will facilitate the derivation of gradient of the group delay with respect to the subfilter coefficients, which is given next.

3.3.2 Gradients for Basic FRM Filters

For an FRM filter with nonlinear phase masking filters, both the frequency response and the group delay are nonlinear with respect to the subfilter coefficients. Gradients of the frequency response and group delay are derived here, as they are the key components of the proposed design method to be presented in the next section.

Gradient of the Frequency Response
3.3. Group Delay and Gradients for Basic FRM Filters

The frequency response of the basic FRM filter is

$$H_f(\omega) = H_1(M\omega)H_3(\omega) + (e^{-jM_d\omega} - H_1(M\omega))H_4(\omega)$$

$$= h_1^T e_1(h_3^T e_3 - h_4^T e_4) + e^{-jM_d\omega}h_4^T e_4,$$  \hspace{1cm} (3.15)

where

$$e_1 = \begin{bmatrix} 1 & e^{-jM\omega} & e^{-j2M\omega} & \ldots & e^{-j(N_1-1)M\omega} \end{bmatrix}^T,$$

$$e_i = \begin{bmatrix} e^{-j\frac{N_m-N_i}{2}\omega} & e^{-j\frac{N_m-N_i+2}{2}\omega} & \ldots & e^{-j\frac{N_m+N_i-2}{2}\omega} \end{bmatrix}^T, \hspace{0.5cm} i = 3, 4.$$

When the design variables are put together as a design vector $x = [h_1^T \ h_3^T \ h_4^T]^T$ of dimension $N_T = N_1 + N_3 + N_4$, the gradient of the frequency response with respect to $x$ is complex-valued given by

$$g_m(\omega) = \begin{bmatrix} \frac{dH_1(\omega)}{dh_1} \\
\frac{dH_3(\omega)}{dh_3} \\
\frac{dH_4(\omega)}{dh_4} \end{bmatrix} = \begin{bmatrix} (H_3(\omega) - H_4(\omega))e_1 \\
H_1(M\omega)e_3 \\
(e^{-jM_d\omega} - H_1(M\omega))e_4 \end{bmatrix}.$$  \hspace{1cm} (3.16)

Gradient of the Group Delay

Using the group delay expressions given in (3.10)-(3.14), the gradient of the group delay with respect to $x$ is defined as

$$g_\tau(\omega) = \begin{bmatrix} \frac{d\tau_f(\omega)}{dh_1} \\
\frac{d\tau_f(\omega)}{dh_3} \\
\frac{d\tau_f(\omega)}{dh_4} \end{bmatrix},$$  \hspace{1cm} (3.17)
whose components, the gradient of $\tau_f(\omega)$ with respect to $h_i$, $h_3$ and $h_4$ can be written as

\[
\frac{d\tau_f(\omega)}{dh_i} = \frac{2}{(D_f^2 + N_f^2)^2} \left( D_f \frac{dD_f}{dh_i} + N_f \frac{dN_f}{dh_i} \right) \left( \frac{dN_f}{d\omega} D_f - N_f \frac{dD_f}{d\omega} \right)
\]

\[
- \frac{1}{D_f^2 + N_f^2} \left[ \frac{d^2N_f}{d\omega dh_i} D_f + \frac{dN_f}{d\omega} \frac{dD_f}{dh_i} - \frac{dD_f}{d\omega} \frac{dN_f}{dh_i} \right],
\]

\[i = 1, 3, 4. \quad (3.18)\]

\[
\frac{dN_f}{dh_i}, \frac{dD_f}{dh_i}, \frac{d^2N_f}{d\omega dh_i} \quad \text{and} \quad \frac{d^2D_f}{d\omega dh_i} \quad (i = 1, 3, 4) \quad \text{can be obtained in the following.}
\]

- Components of $\frac{dN_f}{dh_i}$, $\frac{dD_f}{dh_i}$, $\frac{d^2N_f}{d\omega dh_i}$ and $\frac{d^2D_f}{d\omega dh_i}$ ($i = 3, 4$) are given by

\[
\frac{dN_f}{dh_i} = \frac{dA_q}{dh_i} \sin \phi_q + A_q \cos \phi_q \frac{d\phi_q}{dh_i}, \quad \frac{dA_q}{dh_i} = A_i \frac{dA_i}{dh_i}, \quad (3.19a)
\]

\[
\frac{dD_f}{dh_i} = \frac{dA_q}{dh_i} \cos \phi_q - A_q \sin \phi_q \frac{d\phi_q}{dh_i}, \quad \frac{dA_q}{dh_i} = \frac{d\phi_i}{dh_i}, \quad (3.19b)
\]

\[
\frac{d^2N_f}{d\omega dh_i} = \frac{d^2A_q}{d\omega dh_i} \sin \phi_q + \frac{dA_q}{d\omega} \cos \phi_q \frac{d\phi_q}{dh_i} + \frac{dA_q}{dh_i} \cos \phi_q \frac{d\phi_q}{d\omega} - A_q \sin \phi_q \frac{d\phi_q}{dh_i}, \quad (3.19c)
\]

\[
\frac{d^2D_f}{d\omega dh_i} = \frac{d^2A_q}{d\omega dh_i} \cos \phi_q - \frac{dA_q}{d\omega} \sin \phi_q \frac{d\phi_q}{dh_i} - \frac{dA_q}{dh_i} \sin \phi_q \frac{d\phi_q}{d\omega} - A_q \cos \phi_q \frac{d\phi_q}{dh_i}, \quad (3.19d)
\]

\[
\frac{d^2A_q}{d\omega dh_i} = \frac{dA_i}{d\omega} \frac{dA_i}{dh_i} + A_i \frac{d^2A_i}{d\omega dh_i}, \quad \frac{d^2\phi_q}{d\omega dh_i} = \frac{d^2\phi_i}{d\omega dh_i}, \quad (3.19e)
\]

for $(i, l, q) = (3, 1, d)$ or $(i, l, q) = (4, 2, e)$.  

3.3. Group Delay and Gradients for Basic FRM Filters

- Components of \( \frac{dN_f}{dh_1}, \frac{dP_f}{dh_1}, \frac{d^2N_f}{dh_1^2} \) and \( \frac{d^2P_f}{dh_1^2} \) are given by

\[
\begin{align*}
\frac{dN_f}{dh_1} &= \frac{dA_d}{dh_1} \sin \phi_d + A_d \cos \phi_d \frac{d\phi_d}{dh_1} + \frac{dA_e}{dh_1} \sin \phi_e + A_e \cos \phi_e \frac{d\phi_e}{dh_1}, \\
\frac{dP_f}{dh_1} &= \frac{dA_d}{dh_1} \cos \phi_d - A_d \sin \phi_d \frac{d\phi_d}{dh_1} + \frac{dA_e}{dh_1} \cos \phi_e - A_e \sin \phi_e \frac{d\phi_e}{dh_1}, \\
\frac{d^2N_f}{dh_1^2} &= \frac{d^2A_d}{dh_1^2} \sin \phi_d + \frac{dA_d}{dh_1} \cos \phi_d \frac{d\phi_d}{dh_1} + \frac{dA_d}{dh_1} \cos \phi_d \frac{d^2\phi_d}{dh_1^2} - A_d \sin \phi_d \frac{d\phi_d}{dh_1} + A_d \cos \phi_d \frac{d^2\phi_d}{dh_1^2} \\
&\quad + \frac{d^2A_e}{dh_1^2} \sin \phi_e + \frac{dA_e}{dh_1} \cos \phi_e \frac{d\phi_e}{dh_1} + \frac{dA_e}{dh_1} \cos \phi_e \frac{d^2\phi_e}{dh_1^2} - A_e \sin \phi_e \frac{d\phi_e}{dh_1} + A_e \cos \phi_e \frac{d^2\phi_e}{dh_1^2}, \\
\frac{d^2P_f}{dh_1^2} &= \frac{d^2A_d}{dh_1^2} \cos \phi_d - \frac{dA_d}{dh_1} \sin \phi_d \frac{d\phi_d}{dh_1} + \frac{dA_d}{dh_1} \cos \phi_d \frac{d^2\phi_d}{dh_1^2} - A_d \sin \phi_d \frac{d\phi_d}{dh_1} + A_d \cos \phi_d \frac{d^2\phi_d}{dh_1^2} \\
&\quad + \frac{d^2A_e}{dh_1^2} \cos \phi_e - \frac{dA_e}{dh_1} \sin \phi_e \frac{d\phi_e}{dh_1} + \frac{dA_e}{dh_1} \cos \phi_e \frac{d^2\phi_e}{dh_1^2} - A_e \sin \phi_e \frac{d\phi_e}{dh_1} + A_e \cos \phi_e \frac{d^2\phi_e}{dh_1^2}.
\end{align*}
\]

(3.20a) (3.20b) (3.20c) (3.20d) (3.20e) (3.20f) (3.20g) (3.20h) (3.20i)
Note that \( \frac{dA_i}{dh_i}, \frac{d\phi_i}{dh_i}, \frac{d^2A_i}{d\omega dh_i} \) and \( \frac{d^2\phi_i}{d\omega dh_i} \) \((i = 1, \ldots, 4)\) in (3.19) and (3.20) are given by

\[
\frac{dA_i}{dh_i} = \frac{1}{A_i} (c_i c_i^T + s_i s_i^T) h_i = \frac{1}{A_i} P_{mhi} h_i \quad i = 1, \ldots, 4, \quad (3.21a)
\]

\[
\frac{d\phi_i}{dh_i} = \frac{1}{A_i^2} (s_i c_i^T - c_i s_i^T) h_i = \frac{1}{A_i^2} P_{phi_i} h_i \quad i = 1, \ldots, 4, \quad (3.21b)
\]

\[
\frac{d^2A_i}{d\omega dh_i} = -\frac{1}{A_i} \frac{dA_i}{dh_i} \frac{dA_i}{d\omega} + \frac{2}{A_i} P_{m\omega i} h_i \quad i = 1, \ldots, 4, \quad (3.21c)
\]

\[
\frac{d^2\phi_i}{d\omega dh_i} = -\frac{2}{A_i} \frac{dA_i}{dh_i} \frac{d\phi_i}{d\omega} + \frac{2}{A_i^2} P_{p\omega i} h_i \quad i = 1, \ldots, 4. \quad (3.21d)
\]

### 3.3.3 Basic FRM Filters with Linear Phase Masking Filters

In the above derivation of the group delay and the gradients of the frequency response and the group delay, we assume that all subfilters are of nonlinear phase. For cases when either the prototype filter or the masking filters are of linear phase, the derivation can be simplified by exploiting the symmetric properties of the impulse responses.

As we have explained in Section 3.2, the group delay of a linear phase FRM filter in (3.3) is mainly contributed by the prototype filter. To effectively reduce the group delay, the prototype filter should be of nonlinear phase with \( d_a < \frac{N_i-1}{2} \) while the masking filters could be of linear phase. Therefore, in this subsection, we only briefly discuss the derivation for an FRM filter with linear phase masking filters (and a nonlinear phase prototype filter). We assume that the masking filters satisfy the symmetric condition, \( h_i(n) = h_i(N_i - 1 - n) \) \((n = 0, 1, \ldots, \frac{N_i-1}{2}; i = 3, 4)\). The anti-symmetric condition is similar and omitted here.
3.3. Group Delay and Gradients for Basic FRM Filters

Group Delay

As $N_3$ may be different from $N_4$, the length of a masking filter can be adjusted to $N_m$ by appropriate zero-padding. Using the symmetric properties of the masking filters, the frequency responses of the masking filters are

$$H_i(\omega) = e^{-j\omega \frac{N_m-1}{2}} \left[ \sum_{n=0}^{\frac{N_i-1}{2}} h_i(n) \cos n\omega \right] = e^{-j\omega_d m} h_i^T \tilde{y_i}$$

where

$$h_i = \begin{bmatrix} h_i(\frac{N_i-1}{2}) & 2h_i(\frac{N_i-1}{2}+1) & \cdots & 2h_i(N_i-1) \end{bmatrix}^T, \quad i = 3, 4,$$

$$\tilde{y_i} = \begin{bmatrix} 1 & \cos \omega & \cos 2\omega & \cdots & \cos \frac{N_i-1}{2}\omega \end{bmatrix}^T, \quad i = 3, 4.$$  

The magnitude and phase responses of the masking filters become

$$A_i = \sqrt{(h_i^T \tilde{y_i})^2} = |h_i^T \tilde{y_i}|, \quad i = 3, 4,$$

$$\phi_i = -\left(\frac{N_m-1}{2}\right)\omega \pm \pi, \quad i = 3, 4.$$  

The expressions of the group delay of the FRM filter are the same as (3.10)-(3.12) with $\frac{dA_i}{d\omega}$ and $\frac{d\phi_i}{d\omega}$ ($i = 1, 2$) given by (3.13), and

$$\frac{dA_i}{d\omega} = \text{Sign} \left( h_i^T \tilde{y_i} \right) h_i^T \tilde{y_i}, \quad \frac{d\phi_i}{d\omega} = -\frac{N_m-1}{2}, \quad i = 3, 4,$$  

(3.22)

where $\text{Sign}(x)$ means the sign of $x$ and $\tilde{y_i} = \frac{dy_i}{d\omega} = -[0 \sin \omega \sin 2\omega \cdots \sin \frac{N_i-1}{2}\omega].$

Gradient of the Frequency Response and Group Delay
Taking advantage of the symmetric properties of the impulse responses of
the masking filters, the dimension of \( x = [h_1 \ h_3 \ h_4] \) becomes \( N_T = N_1 + \frac{N_4 - 1}{2} + \frac{N_3 - 1}{2} \). In what follows, we focus on the gradient of the group delay and
refer the reader to [LH03b] for the gradient of the frequency response given by

\[
 g_m(\omega) = \begin{bmatrix}
 \frac{dH_f(\omega)}{dh_1} \\
 \frac{dH_f(\omega)}{dh_3} \\
 \frac{dH_f(\omega)}{dh_4}
\end{bmatrix} = \begin{bmatrix}
 (H_3(\omega) - H_4(\omega)) e_1 \\
 e^{-j\omega d_m} H_1(M\omega) y_3 \\
 e^{-j\omega d_m} (e^{-jM\omega} - H_1(M\omega)) y_4
\end{bmatrix}.
\]

Similar to (3.17), the gradient of the group delay in this case is

\[
 g_\tau(\omega) = \begin{bmatrix}
 \frac{d\tau_f(\omega)}{dh_1} \\
 \frac{d\tau_f(\omega)}{dh_3} \\
 \frac{d\tau_f(\omega)}{dh_4}
\end{bmatrix},
\]

\[\frac{d\tau_f(\omega)}{dh_i} \quad (i = 3, 4) \text{ have the same expression as (3.18). And the components}
\]

\[
 \frac{dN_f}{dh_1}, \ \frac{dN_f}{dh_3}, \ \frac{dN_f}{dh_4} \ \text{and} \ \frac{d^2N_f}{d\omega d h_i} \ \text{are given in (3.20), while the components}
\]

\[
 \frac{dD_f}{dh_1}, \ \frac{d^2D_f}{d\omega d h_i} \ \text{and} \ \frac{d^2D_f}{d\omega d h_i} \ (i = 3, 4) \text{ are given by}
\]

\[
 \frac{dN_f}{dh_i} = A_l \frac{dA_i}{dh_i} \sin \phi_q, \quad \frac{dD_f}{dh_i} = A_l \frac{dA_i}{dh_i} \cos \phi_q, \quad (3.23a)
\]

\[
 \frac{d^2N_f}{d\omega d h_i} = \left( \frac{dA_i}{d\omega} \frac{dA_i}{dh_i} + A_l \frac{d^2A_i}{d\omega d h_i} \right) \sin \phi_q + A_l \frac{dA_i}{dh_i} \cos \phi_q \left( \frac{d\phi_i}{d\omega} + \frac{d\phi_i}{d\omega} \right), \quad (3.23b)
\]

\[
 \frac{d^2D_f}{d\omega d h_i} = \left( \frac{dA_i}{d\omega} \frac{dA_i}{dh_i} + A_l \frac{d^2A_i}{d\omega d h_i} \right) \cos \phi_q - A_l \frac{dA_i}{dh_i} \sin \phi_q \left( \frac{d\phi_i}{d\omega} + \frac{d\phi_i}{d\omega} \right), \quad (3.23c)
\]

for \((i, l, q) = (3, 1, d)\) or \((i, l, q) = (4, 2, e)\).
3.4 Design Methodology for Basic FRM Filters

Noted that \( \frac{dA_i}{d\omega h_i} \) and \( \frac{d^2A_i}{d\omega^2 h_i} \) \((i = 3, 4)\) in (3.23) are given as

\[
\frac{dA_i}{d\omega h_i} = y_i, \quad \frac{d^2A_i}{d\omega^2 h_i} = \tilde{y}_i, \quad i = 3, 4.
\] (3.24a)

3.4 Design Methodology for Basic FRM Filters

The proposed design method for FRM filters with reduced group delays can be carried out in two steps. In the first step, a good initial FRM filter is obtained, which is then used as an initial solution for the overall optimization in the second step.

3.4.1 Initial Design

As an initial design, we assume that the masking filters are of linear phase with \( d_m = \frac{N_m - 1}{2} \), while the prototype filter is of nonlinear phase with reduced group delay \( d_a < \frac{N_a - 1}{2} \). Given the masking filters, the prototype filter is to be designed to partially compensate the ripples contributed from the masking filters. Different from the linear phase FRM filter design, the phase responses of the subfilters have to be taken into account. The frequency response of the FRM filter satisfies

\[
(H_f(\omega) - H_4(\omega) e^{-j\omega M d_a}) e^{-j\omega d_m} = (G_3(\omega) - G_4(\omega)) H_1(M\omega, h_1).
\] (3.25)

Note that \( H_1(M\omega, h_1) \) is a function of \( h_1 \) that is to be designed. Replace \( H_f(\omega) \) with \( H_{des}(\omega) \) and denote \( F_d^{(1)} = \left( H_{des}(\omega) - H_4(\omega) e^{-jM d_a} \right) e^{-j\omega d_m} \). The design problem can be converted into the following constrained minimization problem:
3.4. Design Methodology for Basic FRM Filters

\[
\text{minimize}_{\mathbf{h}_1} \quad \eta_m^{(1)} \quad (3.26a)
\]
subject to : \[
\left| W_m^{(1)}(\omega)F_1(M\omega, \mathbf{h}_1) - F_d^{(1)}(\omega) \right| \leq \eta_m^{(1)} \quad \text{for} \quad \omega \in \Omega_m^{(1)}, \quad (3.26b)
\]

where \( W_m^{(1)}(\omega) = G_3(\omega) - G_4(\omega) \), \( \Omega_m^{(1)} \) is the band of interest near the transition band of the FRM filter \([\text{Lim86}]\). The minimax problem in (3.26) can be solved using LP \([\text{Lim86}]\). The FRM filter obtained in this way is a suboptimal filter, and is used as an initial solution for the overall optimization next.

### 3.4.2 Overall Optimization

Firstly, the overall optimization problem taking into account both the complex magnitude error and the group delay error is formulated. Secondly, a group delay constraint is formulated and incorporated into the overall optimization problem in the SOCP framework.

Let \( \tau_{des}(\omega) \) be the desired passband group delay, and \( H_f(\omega, \mathbf{x}) \) and \( \tau_f(\omega, \mathbf{x}) \) be the designed frequency response and passband group delay that depend on a real-valued design vector \( \mathbf{x} \).

**Problem Formulation**

The classical minimax problem to approximate \( H_{des}(\omega) \) can be stated as

\[
\text{minimize}_{\mathbf{x}} \quad \left\{ \max_{\omega \in \Omega_m} \left| W_m(\omega) \right| H_f(\omega, \mathbf{x}) - H_{des}(\omega) \right\}, \quad (3.27)
\]

where \( \Omega_m \) is the band of interest, which consists of the passband and stopband, and \( W_m(\omega) \) is a positive weighting function in magnitude.
Similarly, the minimax problem to approximate \( \tau_{\text{des}}(\omega) \) can be stated as

\[
\minimize_{x} \left\{ \maximize_{\omega \in \Omega_g} W_g(\omega) |\tau_f(\omega, x) - \tau_{\text{des}}(\omega)| \right\}, \tag{3.28}
\]

where \( \Omega_g \) is the band of interest, which is the passband of the FRM filter, and \( W_g(\omega) \) is a positive weighting function in group delay.

Denote the complex magnitude error, \( \eta_m(\omega, x) = W_m(\omega)|H_f(\omega, x) - H_{\text{des}}(\omega)| \) for \( \omega \in \Omega_m \) and the group delay error, \( \eta_g(\omega, x) = W_g(\omega)|\tau_f(\omega, x) - \tau_{\text{des}}(\omega)| \) for \( \omega \in \Omega_g \). Similar to the formulation in Chapter 2, the design of an FRM filter with reduced group delay error can be formulated as

\[
\minimize_{x} \left\{ \maximize_{\omega} \{ \eta_m(\omega, x), \beta \eta_g(\omega, x) \} \right\}, \tag{3.29}
\]

where \( \beta \) is a positive constant balancing the complex magnitude error and the group delay error. Larger \( \beta \) is applied when smaller group delay error is required.

**Estimation Errors for FRM Filters**

Suppose we have a reasonable initial filter \( x_0 \), which is either obtained as described above or from the FRM filter in [LH03b], to start optimization, and we are in the \( k \)-th iteration. For a nonlinear and smooth \( H_f(\omega, x) \) at the vicinity of \( x_k \), i.e., at \( x_k + v \) where \( v \in \mathbb{R}^{N_T \times 1} \) is a small vector, we can write [LH03b] \( H_f(\omega, x_k + v) = H_f(\omega, x_k) + g_{m, k}^T(\omega)v + o(\|v\|) \),

where \( o(\|v\|) \) represents the summation of the higher order terms of the Taylor expansion, and \( g_{m, k}(\omega) \) is the gradient of \( H_f(\omega, x) \) with respect to \( x \) in (3.16) and evaluated at \( x_k \). Hence, provided that \( \|v\| \) is small with \( x = x_k + v \), the
3.4. Design Methodology for Basic FRM Filters

complex magnitude error becomes

\[
\eta_m(\omega, x_k + v) \approx W_m(\omega) \left| g_{m,k}^T(\omega)v + [H_f(\omega, x_k) - H_{des}(\omega)] \right|
\]

\[
= \left| (g_{mr,k}(\omega)v + e_{mr,k}(\omega)) + j (g_{mi,k}(\omega)v + e_{mi,k}(\omega)) \right|
\]

(3.30)

where \( g_{mr,k} \) and \( g_{mi,k} \) are the real and imaginary parts of \( W_m(\omega)g_{m,k} \), respectively, and

\[
e_{mr,k}(\omega) = W_m(\omega) [H_{rf}(\omega, x_k) - H_{rd}(\omega)],
\]

\[
e_{mi,k}(\omega) = W_m(\omega) [H_{if}(\omega, x_k) - H_{id}(\omega)],
\]

with \( H_{rf}(\omega, x_k) \), \( H_{rd}(\omega) \), \( H_{if}(\omega, x_k) \) and \( H_{id}(\omega) \) being the real and imaginary parts of \( H_f(\omega, x_k) \) and \( H_{des}(\omega) \), respectively [LH03b].

Similarly, the group delay error at the \( k \)-th iteration can be obtained through nonlinear approximation

\[
\eta_g(\omega, x_k + v) \approx W_g(\omega) \left| g_{r,k}^T(\omega)v + \tau_f(\omega, x_k) - \tau_{des}(\omega) \right|
\]

\[
= \left| g_{rr,k}^T(\omega)v + e_{rr,k}(\omega) \right|,
\]

(3.31)

where \( g_{r,k}(\omega) \) is the gradient of the group delay with respect to \( x \) in (3.17) and evaluated at \( x_k \), and

\[
e_{rr,k}(\omega) = W_g(\omega) (\tau_f(\omega, x_k) - \tau_{des}(\omega)),
\]

\[
g_{rr,k}(\omega) = W_g(\omega)g_{r,k}(\omega).
\]

SOCP Formulation
Using (3.30) and (3.31), it follows that an approximated solution of (3.29) at the $k$-th iteration can be obtained by solving the following problem

\[
\begin{align*}
\text{minimize} & \quad \eta \\
\text{subject to :} & \quad \|G_{m,k}(\omega)v + e_{m,k}(\omega)\| \leq \eta \quad \text{for } \omega \in \Omega_m, \\
& \quad \beta \|\left(g_{r,r,k}^T(\omega)v + e_{r,r,k}(\omega)\right)\| \leq \eta \quad \text{for } \omega \in \Omega_g, \\
& \quad \|v\| \leq b_v.
\end{align*}
\]

where $b_v$ is a prescribed bound for $v$,

\[
G_{m,k}(\omega) = \begin{bmatrix} g_{mr,k}^T \\ g_{mi,k}^T \end{bmatrix}, \quad \text{and} \quad e_{m,k}(\omega) = \begin{bmatrix} e_{mr,k}(\omega) \\ e_{mi,k}(\omega) \end{bmatrix},
\]

Noted that (3.32c) is the group delay constraint.

If we treat the upper bound $\eta$ as an additional design variable and define an augmented vector as $y = [\eta \ v]^T$, the objective function in (3.32a) can be expressed as $\eta = u_1^T y$ with $u_1 = [1 \ 0 \ \ldots \ 0]^T$. The constrained optimization problem defined by (3.32) can then be formulated as an SOCP problem by digitizing the frequency variable $\omega$ over a dense set of frequencies in the bands of interest. With $m_K$ discrete frequency grids $\{\omega_1, \ldots, \omega_{m_K}\} \subset \Omega_m$ and $g_K$ discrete frequency grids $\{\omega_1, \ldots, \omega_{g_K}\} \subset \Omega_g$, the $k$-th iteration problem in
(3.32) becomes

\[
\begin{align*}
\text{minimize} & \quad u_1^T y \\
\text{subject to} : & \\
& \begin{bmatrix} 1 & 0^T \\ 0 & G_{m,k}(\omega_l) \end{bmatrix} y + \begin{bmatrix} 0 \\ e_{m,k}(\omega_l) \end{bmatrix} \in C_l, \quad l = 1, \ldots, m_K, \\
& \begin{bmatrix} 1 & 0^T \\ 0 & \beta g_{rr,k}(\omega_i) \end{bmatrix} y + \begin{bmatrix} 0 \\ \beta e_{rr,k}(\omega_i) \end{bmatrix} \in C_i, \quad i = 1, \ldots, g_K, \\
& \begin{bmatrix} 0 & 0^T \\ 0 & I \end{bmatrix} y + \begin{bmatrix} b_v \\ 0 \end{bmatrix} \in C_p,
\end{align*}
\]  

where \( I \) is an identity matrix of size \( N_T \), and \( C_l, C_i \) and \( C_p \) are the second-order cones in \( \mathcal{R}^3, \mathcal{R}^2 \) and \( \mathcal{R}^{N_T+1} \), respectively. Noted that there are \( m_K + g_K + 1 \) second-order cone constraints in (3.33), which can be solved efficiently by any SOCP solvers, such as SeDuMi used in the simulations.

Having solved the problem in (3.33) for a minimizer

\[ y_k^* = \begin{bmatrix} \eta_k^* \\ v_k^* \end{bmatrix}, \]

vector \( v_k^* \) is used to update \( x_k \) as

\[ x_{k+1} = x_k + v_k^*. \]

The iteration continues until \( \|v_k^*\| \) becomes insignificantly small compared to a prescribed tolerance.
3.5 Design of Multi-Stage FRM Filters

With regards to the proposed design method, some points are highlighted in the following.

- Concerning the convergence of the proposed method, although a rigorous proof is presently not available, in our simulations when the method was applied to design a variety of FRM filters, it always converges. The number of updates depends on the initial design.

- Bound $b_v$ in (3.33d) determines the deviation of $v_k$ in successive update. We found the proposed design method is sensitive to $b_v$, which is probably due to the incorporation of the group delay constraint. Based on our numerical simulations, we recommend it in the range between $2 \times 10^{-4} N_T$ and $5 \times 10^{-4} N_T$, which is somewhat smaller than the one suggested in [LH03b].

- Regarding the selection of $\beta$, it can be approximated by ratio of $\max_{\omega \in \Omega_m} \{\eta_m(\omega, x_0)\}$ and $\max_{\omega \in \Omega_g} \{\eta_g(\omega, x_0)\}$. Both can be computed from the initial design. To improve the performance of the group delay, slightly larger $\beta$ should be applied. Typically, $\beta \in [1 \times 10^{-3}, 5 \times 10^{-3}]$ is used in the design examples.

The method developed above is applicable to the design of general basic FRM filters with linear or nonlinear phase masking filters, as to be demonstrated in Section 3.6.

3.5 Design of Multi-Stage FRM Filters

The design method for a basic FRM filter presented in the previous sections is now extended to that for a multi-stage FRM filter. For the sake of illustration,
the derivation for two-stage FRM filters is given in details in this section. It can be extended in principle to a higher-stage FRM filter but the expressions will become very involved and hence omitted here. In the following, we focus on the derivation of the group delay and the gradients of the frequency response and the group delay for a two-stage FRM filter with nonlinear phase masking filters.

### 3.5.1 Group Delay for Two-stage FRM Filters

A simplified two-stage FRM filter consists of a prototype filter \( H_1(z) \) and two pairs of masking filters \( \{ H_6(z), H_7(z) \} \) and \( \{ H_3(z), H_4(z) \} \), as shown in Figure 3.4. The vectors of the coefficients of the subfilters are \( \mathbf{h}_1 \in \mathbb{R}^{N_1 \times 1} \), \( \mathbf{h}_6 \in \mathbb{R}^{N_6 \times 1} \), \( \mathbf{h}_7 \in \mathbb{R}^{N_7 \times 1} \), \( \mathbf{h}_3 \in \mathbb{R}^{N_3 \times 1} \), and \( \mathbf{h}_4 \in \mathbb{R}^{N_4 \times 1} \), which correspond to \( \mathbf{h}_a \), \( \mathbf{h}_{ma2} \), \( \mathbf{h}_{mc2} \), \( \mathbf{h}_{ma1} \) and \( \mathbf{h}_{mc1} \) (\( H_a(z) \), \( H_{ma2}(z) \), \( H_{mc2}(z) \), \( H_{ma1}(z) \) and \( H_{mc1}(z) \)) in Figure 3.2, respectively.

Let \( N_{m2} = \max\{N_6, N_7\} \) and \( N_{m1} = \max\{N_3, N_4\} \). Denote \( d_a \), \( d_{m2} \) and \( d_{m1} \) the group delays of \( H_1(z) \), \( \{ H_6(z), H_7(z) \} \) and \( \{ H_3(z), H_4(z) \} \), respectively.
A dummy vector $h_2 = u_{d_i+1} - h_1$ for the complementary filter of $H_1(z)$ is introduced and we have $H_2(z) = z^{-d_i} - H_1(z)$. Let $d_5 = M_1M_2d_a + M_2d_{m_2}$.

Let $A_1(\omega)e^{j\phi_1(\omega)}$, $A_2(\omega)e^{j\phi_2(\omega)}$, $A_6(\omega)e^{j\phi_6(\omega)}$, $A_7(\omega)e^{j\phi_7(\omega)}$, $A_3(\omega)e^{j\phi_3(\omega)}$ and $A_4(\omega)e^{j\phi_4(\omega)}$ be the frequency responses of $H_1(z^{M_1M_2})$, $H_2(z^{M_1M_2})$, $H_6(z^{M_2})$, $H_7(z^{M_2})$, $H_3(z)$ and $H_4(z)$. We have

$$A_i = \sqrt{(h_i^Tc_i)^2 + (h_i^Ts_i)^2}, \quad \phi_i = \arctan 2 \left( h_i^Ts_i, h_i^Tc_i \right), \quad i = 1, 2, 6, 7, 3, 4,$$

where

$$c_i = \begin{bmatrix} 1 & \cos M_1M_2\omega & \cos 2M_1M_2\omega & \cdots & \cos(N_1 - 1)M_1M_2\omega \end{bmatrix}^T, \quad i = 1, 2, \quad (3.34a)$$

$$s_i = \begin{bmatrix} 0 & -\sin M_1M_2\omega & -\sin 2M_1M_2\omega & \cdots & -\sin(N_1 - 1)M_1M_2\omega \end{bmatrix}^T, \quad i = 1, 2, \quad (3.34b)$$

$$c_l = \begin{bmatrix} \cos \left( \frac{N_{m_2}-1}{2} - \frac{N_l-1}{2} \right) M_2\omega & \cdots & \cos \frac{N_{m_2}-1}{2} M_2\omega & \cdots & \cos \left( \frac{N_{m_2}-1}{2} + \frac{N_l-1}{2} \right) M_2\omega \end{bmatrix}^T, \quad l = 6, 7, \quad (3.34c)$$

$$s_l = \begin{bmatrix} -\sin \left( \frac{N_{m_2}-1}{2} - \frac{N_l-1}{2} \right) M_2\omega & \cdots & -\sin \frac{N_{m_2}-1}{2} M_2\omega & \cdots & -\sin \left( \frac{N_{m_2}-1}{2} + \frac{N_l-1}{2} \right) M_2\omega \end{bmatrix}^T, \quad l = 6, 7, \quad (3.34d)$$

$$c_q = \begin{bmatrix} \cos \left( \frac{N_{m_1}-1}{2} - \frac{N_q-1}{2} \right)\omega & \cdots & \cos \frac{N_{m_1}-1}{2}\omega & \cdots & \cos \left( \frac{N_{m_1}-1}{2} + \frac{N_q-1}{2} \right)\omega \end{bmatrix}^T, \quad q = 3, 4, \quad (3.34e)$$

$$s_q = \begin{bmatrix} -\sin \left( \frac{N_{m_1}-1}{2} - \frac{N_q-1}{2} \right)\omega & \cdots & -\sin \frac{N_{m_1}-1}{2}\omega & \cdots & -\sin \left( \frac{N_{m_1}-1}{2} + \frac{N_q-1}{2} \right)\omega \end{bmatrix}^T, \quad q = 3, 4. \quad (3.34f)$$
3.5. Design of Multi-Stage FRM Filters

The frequency responses at nodes ‘a’, ‘b’, ‘c’, ‘5’, ‘d’, ‘e’ and ‘f’ are

\[ A_a e^{j\phi_a} = A_1 A_6 e^{j(\phi_1 + \phi_6)}, \quad A_b e^{j\phi_b} = A_2 A_7 e^{j(\phi_2 + \phi_7)}, \]
\[ A_c e^{j\phi_c} = A_a e^{j\phi_a} + A_b e^{j\phi_b}, \quad A_5 e^{j\phi_5} = e^{-j\omega d_5} - A_c e^{j\phi_c}, \]
\[ A_d = A_c A_3 e^{j(\phi_c + \phi_3)}, \quad A_e = A_5 A_4 e^{j(\phi_5 + \phi_4)}, \]
\[ A_f e^{j\phi_f} = A_d e^{j\phi_d} + A_e e^{j\phi_e}. \]

Let \( D_f, D_5, D_c \) and \( N_f, N_5, N_c \) be the real and imaginary parts of \( A_f e^{j\phi_f}, A_5 e^{j\phi_5} \) and \( A_c e^{j\phi_c} \), respectively, i.e.,

\[ N_f = A_d \sin \phi_d + A_c \sin \phi_c, \quad N_5 = -\sin d_5 \omega - A_c \sin \phi_c, \quad N_c = A_d \sin \phi_a + A_b \sin \phi_b, \]
\[ D_f = A_d \cos \phi_d + A_c \cos \phi_c, \quad D_5 = \cos d_5 \omega - A_c \cos \phi_c, \quad D_c = A_a \cos \phi_a + A_b \cos \phi_b. \]

We have

\[ A_f = \sqrt{N_f^2 + D_f^2}, \quad A_5 = \sqrt{D_5^2 + N_5^2}, \quad A_c = \sqrt{N_c^2 + D_c^2}, \]
\[ \phi_f = \arctan 2(N_f, D_f), \quad \phi_5 = \arctan 2(N_5, D_5), \quad \phi_c = \arctan 2(N_c, D_c). \]

Then the phase response and group delay of the two-stage FRM filter are

\[ \phi_f = \arctan 2(N_f, D_f), \quad (3.35a) \]
\[ \tau_f(\omega) = -\frac{1}{D_f^2 + N_f^2} \left( \frac{dN_f}{d\omega} D_f - N_f \frac{dD_f}{d\omega} \right), \quad (3.35b) \]

where \( \frac{dN_f}{d\omega} \) and \( \frac{dD_f}{d\omega} \) as functions of \( \frac{dA_d}{d\omega}, \frac{d\phi_d}{d\omega}, \frac{dA_c}{d\omega} \) and \( \frac{d\phi_c}{d\omega} \) are given in (3.11).
3.5. Design of Multi-Stage FRM Filters

Subsequently, \( \frac{dA_d}{d\omega}, \frac{d\phi_d}{d\omega}, \frac{dA_e}{d\omega} \) and \( \frac{d\phi_e}{d\omega} \) can be obtained as

\[
\begin{align*}
\frac{dA_d}{d\omega} &= \frac{dA_c}{d\omega} A_3 + A_e \frac{dA_3}{d\omega}, \\
\frac{dA_e}{d\omega} &= A_3 \frac{dA_3}{d\omega}, \\
\frac{dA_3}{d\omega} &= \frac{1}{A_c} \left( N_c \frac{dN_c}{d\omega} + D_c \frac{dD_c}{d\omega} \right), \\
\frac{d\phi_d}{d\omega} &= \frac{d\phi_c}{d\omega} + \frac{d\phi_3}{d\omega}, \\
\frac{d\phi_c}{d\omega} &= \frac{d\phi_e}{d\omega} + \frac{d\phi_4}{d\omega}, \\
\frac{d\phi_e}{d\omega} &= \frac{d\phi_5}{d\omega} + \frac{d\phi_4}{d\omega}.
\end{align*}
\]

Noted that \( \frac{dA_i}{d\omega} \) and \( \frac{d\phi_i}{d\omega} (i = 1, 2, 6, 7, 3, 4) \) are given in (3.13) and (3.14).

3.5.2 Gradients for Two-Stage FRM Filters

For a two-stage FRM filter, the frequency response is

\[
H_c(\omega) = H_1(M_1\omega)H_6(\omega) + \left( e^{-jM_1d_\omega} - H_1(M_1\omega) \right) H_7(\omega), \tag{3.37a}
\]

\[
H_f(\omega) = H_c(M_2\omega)H_8(\omega) + \left( e^{-jM_2d_\omega} - H_c(M_2\omega) \right) H_4(\omega). \tag{3.37b}
\]

The gradient \( g_m(\omega) \) of the frequency response with respect to the design vector

\[
x = [h_1^T \ h_6^T \ h_7^T \ h_3^T \ h_4^T]^T \]

is a complex-valued vector of dimension \( N_T = \)
3.5. Design of Multi-Stage FRM Filters

\[ N_1 + N_6 + N_7 + N_3 + N_4 \]
given by

\[
\mathbf{g}_m(\omega) = \begin{bmatrix}
\frac{dH_f(\omega)}{dh_1} \\
\frac{dH_f(\omega)}{dh_6} \\
\frac{dH_f(\omega)}{dh_7} \\
\frac{dH_f(\omega)}{dh_3} \\
\frac{dH_f(\omega)}{dh_4}
\end{bmatrix} =
\begin{bmatrix}
(H_3(\omega) - H_4(\omega))(H_6(M_2\omega) - H_7(M_2\omega))\mathbf{e}_1 \\
(H_3(\omega) - H_4(\omega))H_1(M_1M_2\omega)\mathbf{e}_6 \\
(H_3(\omega) - H_4(\omega))(e^{-jM_1M_2d_\omega} - H_1(M_1M_2\omega))\mathbf{e}_7 \\
H_c(M_2\omega)\mathbf{e}_3 \\
(e^{-jd_\omega} - H_c(M_2\omega))\mathbf{e}_4
\end{bmatrix},
\]

(3.38)

where

\[ \mathbf{e}_1 = \begin{bmatrix} 1 & e^{-jM_1M_2\omega} & e^{-j2M_1M_2\omega} & \ldots & e^{-j(N_1-1)M_1M_2\omega} \end{bmatrix}^T, \]

\[ \mathbf{e}_i = \begin{bmatrix} e^{-j\frac{N_2-N_1}{2}M_2\omega} & e^{-j\frac{N_2-N_1+2}{2}M_2\omega} & \ldots & e^{-j\frac{N_2+N_1-2}{2}M_2\omega} \end{bmatrix}^T, \quad i = 6, 7, \]

\[ \mathbf{e}_l = \begin{bmatrix} e^{-j\frac{N_1-N_2}{2}\omega} & e^{-j\frac{N_1-N_2+2}{2}\omega} & \ldots & e^{-j\frac{N_1+N_2-2}{2}\omega} \end{bmatrix}^T, \quad l = 3, 4. \]

Using the group delay in (3.35), its gradient with respect to \( \mathbf{x} \) is defined as

\[
\mathbf{g}_\tau(\omega) = \begin{bmatrix}
\frac{d\tau_f(\omega)}{dh_1} \\
\frac{d\tau_f(\omega)}{dh_6} \\
\frac{d\tau_f(\omega)}{dh_7} \\
\frac{d\tau_f(\omega)}{dh_3} \\
\frac{d\tau_f(\omega)}{dh_4}
\end{bmatrix}.
\]

(3.39)

The derivation to obtain \( \frac{d\tau_f(\omega)}{dh_i} \) (\( i = 1, 6, 7, 3, 4 \)) is similar to that for a basic
FRM filter. In general, we have
\[
\frac{d\tau_f}{dh_i} = \frac{2}{(D_f^2 + N_f^2)^2} \left( D_f \frac{dD_f}{dh_i} + N_f \frac{dN_f}{dh_i} \right) \left( \frac{dN_f}{d\omega} D_f - N_f \frac{dD_f}{d\omega} \right)
- \frac{1}{D_f^2 + N_f^2} \left( \frac{d^2N_f}{d\omega dh_i} D_f + \frac{dN_f}{d\omega} \frac{dD_f}{dh_i} - \frac{dN_f}{dh_i} \frac{dD_f}{d\omega} - N_f \frac{d^2D_f}{d\omega dh_i} \right),
\]
i = 1, 6, 7, 3, 4.

Using \( \frac{dA_i}{d\omega}, \frac{d\phi_i}{d\omega}, \frac{dA_i}{dh_i}, \frac{d\phi_i}{dh_i}, \frac{d^2A_i}{d\omega dh_i}, \) and \( \frac{d^2\phi_i}{d\omega dh_i}, \) \( i = 1, 6, 7, 3, 4 \) given by (3.13) and (3.21), we have the same expressions for \( \frac{dN_f}{dh_i}, \frac{dD_f}{dh_i}, \frac{d^2N_f}{d\omega dh_i}, \) and \( \frac{d^2D_f}{d\omega dh_i}, \) \( l = 3, 4 \) given by (3.19). Detailed derivation for \( \frac{dN_f}{dh_i}, \frac{dD_f}{dh_i}, \frac{d^2N_f}{d\omega dh_i}, \) and \( \frac{d^2D_f}{d\omega dh_i}, \) \( i = 1, 6, 7 \) is given in the Appendix (Section 3.8).

### 3.5.3 Two-Stage FRM Filter Design

The design of a two-stage FRM filter is similar to that for a basic FRM filter described in Section 3.4. In the initial design, with passband and stopband cutoff frequencies, and ripples slightly smaller than those of the desired FRM filter, the two pairs of masking filters, \( \{H_6(z), H_7(z)\} \) and \( \{H_3(z), H_4(z)\} \), can be designed separately, before the prototype filter \( H_1(z) \) being designed to partially offset the ripples around the transition band. In the overall optimization, the same design methodology can be applied here with the gradients in (3.38) and (3.39).

### 3.6 Design Examples

In this section, two design examples for basic FRM filters and one for a two-stage FRM filter are given. To compare the performance of the designed filters with
existing methods, passband ripple, $\delta_p$, stopband attenuation, $\delta_s$, both in decibels, and relative deviation in group delay, defined as $D_{rd} = \max \left| \frac{\tau_f(\omega) - \tau_{des}(\omega)}{\tau_{des}(\omega)} \right|$, are adopted.

### 3.6.1 Example 3.1: A Basic FRM Filter

A basic FRM filter with cutoff frequencies $\omega_p = 0.6\pi$, $\omega_s = 0.61\pi$ [Lim86] and desired passband group delay $\tau_{des}(\omega) = 182$ is considered here. Given $M = 9$, the subfilters are designed with $N_1 = 45$, $N_3 = 41$, $N_4 = 33$, $d_m = 20$ and $d_a = 18$, resulting in an initial FRM filter with the reduced group delay at around 182 and $D_{rd} = 0.0332$. For the overall optimization, 700 and 500 frequency grids ($m_K = 1200$ and $g_K = 700$) are selected in the passband and the stopband, respectively, and relatively dense grid points are placed in the regions near the bandedges. Both weighting functions, $W_m(\omega)$ and $W_g(\omega)$, are unity. Some important parameters of the designed FRM filters are measured and tabulated in Table 3.3.

**Case with linear phase masking filters:** With $\beta = 2 \times 10^{-3}$ and $b_v = 2 \times 10^{-4}N_T$ ($N_T = 83$), the proposed design method converges in 8 iterations. The performance of the proposed FRM filter is presented in solid lines in Figure 3.5, as compared with the initial FRM filter in broken lines. The passband ripple and the stopband attenuation of the proposed FRM filter are about 0.002 dB and 0.07 dB smaller than those of the initial FRM filter. Moreover, the group delay error is reduced by 0.007, which corresponds to 22% reduction in group delay error. Thus the proposed FRM filter outperforms the initial FRM filter in all aspects. Further reduction in group delay error can also be achieved by increasing $\beta$. For example, when $\beta = 3 \times 10^{-3}$, the group delay error is
Figure 3.5: Example 3.1. A basic FRM filter with linear phase masking filters when $\beta = 2 \times 10^{-3}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the initial FRM filter is in broken lines.
Figure 3.6: Example 3.1. A basic FRM filter with nonlinear phase masking filters when $\beta = 3 \times 10^{-3}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the initial FRM filter is in broken lines.
reduced by 0.015 compared to the initial FRM filter, which corresponds to 46% reduction in group delay error. However, the improvement in group delay error is achieved at the cost of increase in magnitude ripples, i.e., the passband ripple and stopband attenuation are increased by 0.001 dB and 0.2 dB, respectively in this case.

*Case with nonlinear phase masking filters:* With $\beta = 2 \times 10^{-3}$, $b_v = 2 \times 10^{-4} N_T$ ($N_T = 83$) and the same initial FRM filter, the proposed design method converges in 19 iterations. The passband ripple and the stopband attenuation of the proposed FRM filter are about 0.013 dB and 0.3 dB smaller than those of the initial FRM filter. Moreover, the group delay error is reduced by 0.011, which corresponds to 33% reduction in group delay error. The better performance of the proposed FRM filter is partially attributed to the coefficients of the nonlinear phase masking filter. Further reduction in group delay error can also be achieved by increasing $\beta$. For example, when $\beta = 3 \times 10^{-3}$, the group delay error is reduced by 0.018 compared to the initial FRM filter, which corresponds to 54% reduction in group delay error. However, the improvement in group delay error is achieved at the cost of increase in magnitude ripples. In this case, the passband ripple is reduced by 0.011 dB while the stopband attenuation is increased by 0.09 dB. The performance of the proposed FRM filter is presented in solid lines in Figure 3.6, as compared with the initial FRM filter in broken lines.

### 3.6.2 Example 3.2: Another Basic FRM Filter

Another basic FRM filter with cutoff frequencies $\omega_p = 0.6\pi$, $\omega_s = 0.61\pi$ and the passband group delay $\tau_{des}(\omega) = 120$ is considered here [CL96]. To be comparable
3.6. Design Examples

Figure 3.7: Example 3.2. Another basic FRM filter with nonlinear phase masking filters when $\beta = 1 \times 10^{-3}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the reference FRM filter is in broken lines.
### Table 3.2: Important parameters for the basic FRM filters for Example 3.1.

<table>
<thead>
<tr>
<th>Filter</th>
<th>β</th>
<th>$D_{rd}$</th>
<th>$\delta_p$ (dB)</th>
<th>$\delta_s$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial FRM Filter</td>
<td>(NA)</td>
<td>0.0332</td>
<td>0.0845</td>
<td>-40.2190</td>
</tr>
<tr>
<td>Proposed FRM (linear masking)</td>
<td>$2 \times 10^{-3}$</td>
<td>0.0259</td>
<td>0.0824</td>
<td>-40.2878</td>
</tr>
<tr>
<td>Proposed FRM (linear masking)</td>
<td>$3 \times 10^{-3}$</td>
<td>0.0178</td>
<td>0.0856</td>
<td>-39.9822</td>
</tr>
<tr>
<td>Proposed FRM (nonlinear masking)</td>
<td>$2 \times 10^{-3}$</td>
<td>0.0224</td>
<td>0.0710</td>
<td>-40.5055</td>
</tr>
<tr>
<td>Proposed FRM (nonlinear masking)</td>
<td>$3 \times 10^{-3}$</td>
<td>0.0154</td>
<td>0.0733</td>
<td>-40.1331</td>
</tr>
</tbody>
</table>

in implementation complexity with the FRM filter presented at Example 3 in [CL96], the subfilters are designed with $M = 9$, $N_1 = 51$, $N_3 = 43$, $N_4 = 35$, $d_a = 11$ and $d_m = 21$, leading to an initial FRM filter with the reduced group delay at about 120. For the overall optimization, 700 and 500 frequency grids ($m_K = 1200$ and $g_K = 700$) are selected for the passband and the stopband, respectively, and relatively dense grid points are placed in the regions near the bandedges. Both weighting functions, $W_m(\omega)$ and $W_g(\omega)$, are unity. Some important parameters are measured for different $\beta$ in Table 3.3.

With $\beta = 1 \times 10^{-3}$ and $b_v = 2 \times 10^{-4}N_T$ ($N_T = 129$), the proposed design method converges in 17 iterations. The performance of the proposed FRM filter is presented in solid lines in Figure 3.7, as compared with the reference FRM filter obtained by the weighted least-squares [CL96] in broken lines. The passband ripple and the stopband attenuation of the proposed FRM filter are about 0.011 dB and 0.8 dB smaller than those of the reference filter. Moreover, the group delay error is reduced by 0.021, which corresponds to 15% reduction in group delay error. Thus the proposed FRM filter outperforms the reference FRM filter in all aspects. Further reduction in group delay error can also be achieved by increasing $\beta$. For example, when $\beta = 2 \times 10^{-3}$, the group delay error is reduced by 0.041 compared to the initial FRM filter, which corresponds to 50% reduction in group delay error. However, the improvement in group delay error...
delay error is achieved at the cost of increase in magnitude ripples. In this case, the passband ripple is reduced by 0.0007 dB while the stopband attenuation is increased by 0.38 dB.

<table>
<thead>
<tr>
<th>Filter</th>
<th>(\beta)</th>
<th>(D_{rd})</th>
<th>(\delta_p) (dB)</th>
<th>(\delta_s) (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference FRM filter [CL96]</td>
<td>(NA)</td>
<td>0.0816</td>
<td>0.0840</td>
<td>-39.7767</td>
</tr>
<tr>
<td>Proposed FRM filter</td>
<td>(1 \times 10^{-3})</td>
<td>0.0695</td>
<td>0.0732</td>
<td>-40.5803</td>
</tr>
<tr>
<td>Proposed FRM filter</td>
<td>(2 \times 10^{-3})</td>
<td>0.0399</td>
<td>0.0833</td>
<td>-39.3936</td>
</tr>
</tbody>
</table>

### 3.6.3 Example 3.3: A Two-Stage FRM Filter

A two-stage FRM filter with cutoff frequencies \(\omega_p = 0.6\pi\), \(\omega_s = 0.61\pi\) and passband group delay \(\tau_{des}(\omega) = 207\) is considered here [LH03b]. In this example, the optimal FRM filter from [LH03b] is used as the initial design for the overall optimization. Some important parameters are \(M_1 = M_2 = 4\), \(N_1 = 27\), \(N_6 = 13\), \(N_7 = 27\), \(N_3 = 15\), \(N_4 = 23\), \(d_a = 9\), \(d_{m1} = 11\) and \(d_{m2} = 13\). For the overall optimization, 700 and 500 frequency grids \((m_K = 1200\) and \(g_K = 700\)) are selected for the passband and the stopband, respectively, and relatively dense grid points are placed in the regions near the bandedges. Both weighting functions, \(W_m(\omega)\) and \(W_g(\omega)\), are unity. Some other results with different \(\beta\) are also tabulated in 3.4.

**Case with linear phase masking filters:** With \(\beta = 8.5 \times 10^{-4}\) and \(b_v = 2 \times 10^{-4}N_T\) \((N_T = 68)\), the proposed design method converges in 3 iterations. The performance of the proposed FRM filter is presented in solid lines in Figure 3.8, as compared with the initial FRM filter in broken lines. The passband ripple is slightly smaller and the stopband attenuation is about 0.6 dB smaller than that of the initial FRM filter. Moreover, the group delay error is reduced by 0.0006,
3.6. Design Examples

Figure 3.8: Example 3.3. A two-stage FRM filter with linear phase masking filters when $\beta = 8.5 \times 10^{-4}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the initial FRM filter is in broken lines.
Figure 3.9: Example 3.3. A two-stage FRM filter with nonlinear phase masking filters when $\beta = 2 \times 10^{-3}$. (a) magnitude response; (b) passband ripple; (c) stopband attenuation; (d) group delay. The proposed FRM filter is indicated in solid lines while the initial FRM filter is in broken lines.
which corresponds to 3% reduction in group delay error. Thus the proposed FRM filter slightly better than the initial FRM filter. Further reduction in group delay error can also be achieved by increasing $\beta$. For example, when $\beta = 2 \times 10^{-3}$, the group delay error is reduced by 0.013 compared to the initial FRM filter, which corresponds to 52% reduction in group delay error. However, the improvement in group delay error is achieved at the cost of increase in magnitude ripples, \textit{i.e.}, the passband ripple and stopband attenuation are increased by 0.006 dB and 0.5 dB, respectively in this case.

\textit{Case with nonlinear phase masking filters:} With $\beta = 8.5 \times 10^{-4}$, $b_v = 2 \times 10^{-4} N_T$ ($N_T = 105$) and the same initial FRM filter, the proposed design method converges in 14 iterations. The passband ripple and the stopband attenuation of the proposed FRM filter are about 0.004 dB and 1.5 dB smaller than those of the initial FRM filter. Moreover, the group delay error is reduced by 0.002, which corresponds to 10\% reduction in group delay error. The better performance of the proposed FRM filter is partially attributed to the coefficients of the nonlinear phase masking filter. Further reduction in group delay error can also be achieved by increasing $\beta$. For example, when $\beta = 2 \times 10^{-3}$, the group delay error is reduced by 0.013 compared to the initial FRM filter, which corresponds to 47\% reduction in group delay error. However, the improvement in group delay error is achieved at the cost of increase in magnitude ripples. In this case, the passband ripple and the stopband attenuation are increased by 0.02 dB and 0.1 dB, respectively. The performance of the proposed FRM filter is presented in solid lines in Figure 3.9, as compared with the initial FRM filter in broken lines.

It should be pointed out that the relative deviation in group delay, $D_{rd}$, using
the method of [LH03b] should be 0.0264 instead of 0.0132 given on page 566 of [LH03b], the latter being arrived using an approximated method for evaluating $D_{rd}$. We are grateful to Prof. W.-S. Lu for the discussion [Lu05].

Table 3.4: Important parameters for the two-stage FRM filters for Example 3.3.

<table>
<thead>
<tr>
<th>Filter</th>
<th>$\beta$</th>
<th>$D_{rd}$</th>
<th>$\delta_p$ (dB)</th>
<th>$\delta_s$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial FRM filter [LH03b]</td>
<td>(NA)</td>
<td>0.0264</td>
<td>0.0396</td>
<td>-45.9001</td>
</tr>
<tr>
<td>Proposed FRM filter (linear masking)</td>
<td>$8.5 \times 10^{-4}$</td>
<td>0.0258</td>
<td>0.0395</td>
<td>-46.4996</td>
</tr>
<tr>
<td>Proposed FRM filter (linear masking)</td>
<td>$2 \times 10^{-3}$</td>
<td>0.0127</td>
<td>0.0457</td>
<td>-45.4341</td>
</tr>
<tr>
<td>Proposed FRM filter (nonlinear masking)</td>
<td>$8.5 \times 10^{-4}$</td>
<td>0.0237</td>
<td>0.0356</td>
<td>-47.4207</td>
</tr>
<tr>
<td>Proposed FRM filter (nonlinear masking)</td>
<td>$2 \times 10^{-3}$</td>
<td>0.0139</td>
<td>0.0418</td>
<td>-45.8203</td>
</tr>
</tbody>
</table>

### 3.7 Conclusions

In this chapter, we have proposed an improved method for designing real FRM filters with reduced passband group delays. To directly and effectively control the group delay error, the FRM filter design is formulated by taking into account both complex magnitude and group delay errors. One of the key steps is the derivation of the expressions of group delay and its gradients by analyzing the magnitude and phase responses of the subfilters and at various nodes, which is applicable to various FRM filters with linear or nonlinear phase masking filters. A group delay constraint based on the derivation is formulated and incorporated into the overall optimization. Another key step is that the FRM filter design problem, formulated as an SOCP, can be efficiently solved by any existing interior-point method, leading to the reduction in design time.

The effectiveness of the proposed design method has been illustrated by some design examples. For the three examples presented, we are able to obtain new FRM filters that outperform those by other existing FRM filter design methods.
3.8 Appendix: $\frac{d\tau_f(\omega)}{dh_i} (i = 1, 6, 7)$ for a two-stage FRM filter

Further reduction in group delay error has also been achieved by increasing $\beta$ at the cost of slight increase in magnitude ripples, unlike existing methods for designing FRM filters with reduced group delays. It is therefore concluded that the proposed method is very flexible in balancing complex magnitude error and group delay error.

3.8 Appendix: $\frac{d\tau_f(\omega)}{dh_i} (i = 1, 6, 7)$ for a two-stage FRM filter

3.8.1 Common Components

Common components for $\frac{dN_f}{dh_i}$ and $\frac{dD_f}{dh_i}$: $i = 1, 6, 7.$

$$\begin{align*}
\frac{dN_f}{dh_i} &= \frac{dA_d}{dh_i} \sin \phi_d + A_d \cos \phi_d \frac{d\phi_d}{dh_i} + \frac{dA_e}{dh_i} \sin \phi_e + A_e \cos \phi_e \frac{d\phi_e}{dh_i}, \\
\frac{dD_f}{dh_i} &= \frac{dA_d}{dh_i} \cos \phi_d - A_d \sin \phi_d \frac{d\phi_d}{dh_i} + \frac{dA_e}{dh_i} \cos \phi_e - A_e \sin \phi_e \frac{d\phi_e}{dh_i}, \\
\frac{dA_d}{dh_i} &= \frac{dA_e}{dh_i} = \frac{dA_6}{dh_i}, \quad \frac{d\phi_d}{dh_i} = \frac{d(\phi_e + \phi_6)}{dh_i} = \frac{d\phi_e}{dh_i}, \\
\frac{dA_e}{dh_i} &= \frac{dA_5}{dh_i} = \frac{dA_7}{dh_i} = \frac{d\phi_5}{dh_i}, \\
\frac{dA_5}{dh_i} &= \frac{1}{A_5} \left( N_5 \frac{dN_5}{dh_i} + D_5 \frac{dD_5}{dh_i} \right), \quad \frac{d\phi_5}{dh_i} = \frac{1}{A_5^2} \left( \frac{dN_5}{dh_i} D_5 - N_5 \frac{dD_5}{dh_i} \right), \\
\frac{dN_5}{dh_i} &= -\frac{dA_c}{dh_i} \sin \phi_c - A_c \cos \phi_c \frac{d\phi_c}{dh_i}, \quad \frac{dD_5}{dh_i} = -\frac{dA_c}{dh_i} \cos \phi_c + A_c \sin \phi_c \frac{d\phi_c}{dh_i}, \\
\frac{dA_c}{dh_i} &= \frac{1}{A_c} \left( N_c \frac{dN_c}{dh_i} + D_c \frac{dD_c}{dh_i} \right), \quad \frac{d\phi_c}{dh_i} = \frac{1}{A_c^2} \left( \frac{dN_c}{dh_i} D_c - N_c \frac{dD_c}{dh_i} \right). 
\end{align*}$$
3.8. Appendix: $\frac{d^2 \tau_f(\omega)}{d\omega h_i}$ (i = 1, 6, 7) for a two-stage FRM filter

Common components for $\frac{d^2 N_i}{d\omega h_i}$ and $\frac{d^2 D_i}{d\omega h_i}$, i = 1, 6, 7.

$$\frac{d^2 N_f}{d\omega h_i} = \frac{d^2 A_d}{d\omega h_i} \sin \phi_d + \frac{dA_d}{dh_i} \cos \phi_d \frac{d\phi_d}{dh_i} + \frac{dA_d}{dh_i} \cos \phi_d \frac{d\phi_d}{dh_i} - A_d \sin \phi_d \frac{d\phi_d}{dh_i} \frac{d\phi_d}{dh_i}$$
$$\quad + A_d \cos \phi_d \frac{d^2 \phi_d}{d\omega h_i} + \frac{d^2 A_e}{d\omega h_i} \sin \phi_e + \frac{dA_e}{dh_i} \cos \phi_e \frac{d\phi_e}{dh_i} + \frac{dA_e}{dh_i} \cos \phi_e \frac{d\phi_e}{dh_i}$$
$$\quad - A_e \sin \phi_e \frac{d\phi_e}{dh_i} - A_e \cos \phi_e \frac{d^2 \phi_e}{d\omega h_i},$$

$$\frac{d^2 D_f}{d\omega h_i} = \frac{d^2 A_d}{d\omega h_i} \cos \phi_d - \frac{dA_d}{dh_i} \sin \phi_d \frac{d\phi_d}{dh_i} - \frac{dA_d}{dh_i} \sin \phi_d \frac{d\phi_d}{dh_i} - A_d \cos \phi_d \frac{d\phi_d}{dh_i} \frac{d\phi_d}{dh_i}$$
$$\quad - A_d \sin \phi_d \frac{d^2 \phi_d}{d\omega h_i} + \frac{d^2 A_e}{d\omega h_i} \cos \phi_e - \frac{dA_e}{dh_i} \sin \phi_e \frac{d\phi_e}{dh_i} - \frac{dA_e}{dh_i} \sin \phi_e \frac{d\phi_e}{dh_i}$$
$$\quad - A_e \cos \phi_e \frac{d\phi_e}{dh_i} \frac{d\phi_e}{dh_i} - A_e \sin \phi_e \frac{d^2 \phi_e}{d\omega h_i},$$

$$\frac{d^2 A_d}{d\omega h_i} = \frac{d^2 A_d}{d\omega h_i} A_6 + \frac{dA_d}{dh_i} A_6 + \frac{d^2 A_d}{dh_i} A_6$$
$$\frac{d^2 A_e}{d\omega h_i} = \frac{d^2 A_e}{d\omega h_i} A_7 + \frac{dA_e}{dh_i} A_7 + \frac{d^2 A_e}{dh_i} A_7$$
$$\frac{d^2 A_5}{d\omega h_i} = \frac{d^2 A_5}{d\omega h_i} A_7 + \frac{dA_5}{dh_i} A_7 + \frac{d^2 A_5}{dh_i} A_7$$

$$\frac{d^2 A_5}{d\omega h_i} = \frac{d^2 A_5}{d\omega h_i} A_7 + \frac{dA_5}{dh_i} A_7 + \frac{d^2 A_5}{dh_i} A_7$$

$$\frac{d^2 A_5}{d\omega h_i} = \frac{d^2 A_5}{d\omega h_i} A_7 + \frac{dA_5}{dh_i} A_7 + \frac{d^2 A_5}{dh_i} A_7$$

$$\frac{d^2 N_5}{d\omega h_i} = \frac{d^2 A_6}{d\omega h_i} \sin \phi_e - \frac{dA_6}{dh_i} \cos \phi_e \frac{d\phi_e}{dh_i} - \frac{dA_6}{dh_i} \cos \phi_e \frac{d\phi_e}{dh_i} + A_e \sin \phi_e \frac{d\phi_e}{dh_i} \frac{d\phi_e}{dh_i}$$
$$\quad - A_e \cos \phi_e \frac{d^2 \phi_e}{d\omega h_i},$$

$$\frac{d^2 D_5}{d\omega h_i} = \frac{d^2 A_6}{d\omega h_i} \cos \phi_e + \frac{dA_6}{dh_i} \sin \phi_e \frac{d\phi_e}{dh_i} + \frac{dA_6}{dh_i} \sin \phi_e \frac{d\phi_e}{dh_i} + A_e \cos \phi_e \frac{d\phi_e}{dh_i} \frac{d\phi_e}{dh_i}$$
$$\quad + A_e \sin \phi_e \frac{d^2 \phi_e}{d\omega h_i},$$

$$\frac{d^2 A_e}{d\omega h_i} = \frac{d^2 A_e}{d\omega h_i} \frac{dA_e}{dh_i} + \frac{dA_e}{dh_i} \frac{d^2 A_e}{dh_i} \frac{dA_e}{dh_i} + \frac{dA_e}{dh_i} \frac{d^2 A_e}{dh_i} \frac{dA_e}{dh_i}$$

$$\frac{d^2 \phi_e}{d\omega h_i} = \frac{d^2 \phi_e}{d\omega h_i} \frac{dA_e}{dh_i} + \frac{dA_e}{dh_i} \frac{d^2 \phi_e}{d\omega h_i} \frac{dA_e}{dh_i} + \frac{dA_e}{dh_i} \frac{d^2 \phi_e}{d\omega h_i} \frac{dA_e}{dh_i}.$$
3.8.2 Expressions for $\frac{d\tau_f(\omega)}{dh_6}$

\[
\frac{dN_c}{dh_6} = \frac{dA_a}{dh_6} \sin \phi_a + A_a \cos \phi_a \frac{d\phi_a}{dh_6}, \quad \frac{dD_c}{dh_6} = \frac{dA_a}{dh_6} \cos \phi_a - A_a \sin \phi_a \frac{d\phi_a}{dh_6}, \\
\frac{dA_a}{dh_6} = A_1 \frac{dA_a}{dh_6}, \quad \frac{d\phi_a}{dh_6} = \frac{d\phi_a}{dh_6}, \\
\frac{d^2N_c}{d\omega dh_6} = \frac{d^2A_a}{d\omega dh_6} \sin \phi_a + \frac{dA_a}{d\omega} \cos \phi_a \frac{d\phi_a}{dh_6} + \frac{dA_a}{dh_6} \cos \phi_a \frac{d\phi_a}{d\omega} - A_a \sin \phi_a \frac{d\phi_a}{dh_6} \frac{d\phi_a}{d\omega} \\
+ A_a \cos \phi_a \frac{d^2\phi_a}{d\omega dh_6}, \\
\frac{d^2D_c}{d\omega dh_6} = \frac{d^2A_a}{d\omega dh_6} \cos \phi_a - \frac{dA_a}{d\omega} \sin \phi_a \frac{d\phi_a}{dh_6} - \frac{dA_a}{dh_6} \sin \phi_a \frac{d\phi_a}{d\omega} - A_a \cos \phi_a \frac{d\phi_a}{dh_6} \frac{d\phi_a}{d\omega} \\
- A_a \sin \phi_a \frac{d^2\phi_a}{d\omega dh_6}, \\
\frac{d^2A_a}{d\omega dh_6} = \frac{dA_1 dA_a}{d\omega dh_6} + A_1 \frac{d^2A_a}{d\omega dh_6}, \quad \frac{d^2\phi_a}{d\omega dh_6} = \frac{d^2\phi_a}{d\omega dh_6}.
\]

3.8.3 Expressions for $\frac{d\tau_f(\omega)}{dh_7}$

\[
\frac{dN_c}{dh_7} = \frac{dA_b}{dh_7} \sin \phi_b + A_b \cos \phi_b \frac{d\phi_b}{dh_7}, \quad \frac{dD_c}{dh_7} = \frac{dA_b}{dh_7} \cos \phi_b - A_b \sin \phi_b \frac{d\phi_b}{dh_7}, \\
\frac{dA_b}{dh_7} = A_2 \frac{dA_b}{dh_7}, \quad \frac{d\phi_b}{dh_7} = \frac{d\phi_b}{dh_7}, \\
\frac{d^2N_c}{d\omega dh_7} = \frac{d^2A_b}{d\omega dh_7} \sin \phi_b + \frac{dA_b}{d\omega} \cos \phi_b \frac{d\phi_b}{dh_7} + \frac{dA_b}{dh_7} \cos \phi_b \frac{d\phi_b}{d\omega} \\
- A_b \sin \phi_b \frac{d\phi_b}{dh_7} \frac{d\phi_b}{d\omega} + A_b \cos \phi_b \frac{d^2\phi_b}{d\omega dh_7}, \\
\frac{d^2D_c}{d\omega dh_7} = \frac{d^2A_b}{d\omega dh_7} \cos \phi_b - \frac{dA_b}{d\omega} \sin \phi_b \frac{d\phi_b}{dh_7} - \frac{dA_b}{dh_7} \sin \phi_b \frac{d\phi_b}{d\omega} \\
- A_b \cos \phi_b \frac{d\phi_b}{dh_7} \frac{d\phi_b}{d\omega} - A_b \sin \phi_b \frac{d^2\phi_b}{d\omega dh_7}, \\
\frac{d^2A_b}{d\omega dh_7} = \frac{dA_2 dA_b}{d\omega dh_7} + A_2 \frac{d^2A_b}{d\omega dh_7}, \quad \frac{d^2\phi_b}{d\omega dh_7} = \frac{d^2\phi_b}{d\omega dh_7}.
\]
3.8.4 Expressions for $\frac{d\tau_f(\omega)}{dh_i}$

\[
\begin{align*}
\frac{dN_c}{dh_1} &= dA_a \sin \phi_a + A_a \cos \phi_a \frac{d\phi_a}{dh_1} + dA_b \sin \phi_b + A_b \cos \phi_b \frac{d\phi_b}{dh_1} , \\
\frac{dD_c}{dh_1} &= dA_a \cos \phi_a - A_a \sin \phi_a \frac{d\phi_a}{dh_1} + dA_b \cos \phi_b - A_b \sin \phi_b \frac{d\phi_b}{dh_1} , \\
\frac{dA_a}{dh_1} &= A_3 \frac{dA_1}{dh_1} = A_4 \frac{dA_2}{dh_1} , \quad \frac{d\phi_a}{dh_1} = \frac{d\phi_1}{dh_1} = \frac{d\phi_2}{dh_1} , \\
\frac{dA_2}{dh_1} &= -\frac{1}{A_2} P_{mh_2} h_2 , \quad \frac{d\phi_2}{dh_1} = -\frac{1}{A_2^2} P_{ph_2} h_2 , \quad \\
\frac{d^2 N_c}{d\omega dh_1} &= \frac{d^2 A_a}{d\omega dh_1} \sin \phi_a + \frac{d^2 A_a}{d\omega dh_1} \cos \phi_a \frac{d\phi_a}{dh_1} + \frac{d^2 A_a}{d\omega dh_1} \cos \phi_a \frac{d\phi_a}{dh_1} - A_a \sin \phi_a \frac{d\phi_a}{dh_1} \frac{d\phi_a}{d\omega} \\
&+ A_a \cos \phi_a \frac{d^2 \phi_a}{d\omega dh_1} + \frac{d^2 A_b}{d\omega dh_1} \sin \phi_b + \frac{d^2 A_b}{d\omega dh_1} \cos \phi_b \frac{d\phi_b}{dh_1} + \frac{d^2 A_b}{d\omega dh_1} \cos \phi_b \frac{d\phi_b}{dh_1} \\
&- A_b \sin \phi_b \frac{d\phi_b}{dh_1} \frac{d\phi_b}{d\omega} + A_b \cos \phi_b \frac{d^2 \phi_b}{d\omega dh_1} , \\
\frac{d^2 D_c}{d\omega dh_1} &= \frac{d^2 A_a}{d\omega dh_1} \cos \phi_a - \frac{d^2 A_a}{d\omega dh_1} \sin \phi_a \frac{d\phi_a}{dh_1} - \frac{d^2 A_a}{d\omega dh_1} \sin \phi_a \frac{d\phi_a}{dh_1} - A_a \cos \phi_a \frac{d\phi_a}{dh_1} \frac{d\phi_a}{d\omega} \\
&- A_a \sin \phi_a \frac{d^2 \phi_a}{d\omega dh_1} + \frac{d^2 A_b}{d\omega dh_1} \cos \phi_b - \frac{d^2 A_b}{d\omega dh_1} \sin \phi_b \frac{d\phi_b}{dh_1} - \frac{d^2 A_b}{d\omega dh_1} \sin \phi_b \frac{d\phi_b}{dh_1} \\
&- A_b \cos \phi_b \frac{d\phi_b}{dh_1} \frac{d\phi_b}{d\omega} - A_b \sin \phi_b \frac{d^2 \phi_b}{d\omega dh_1} , \\
\frac{d^2 A_b}{d\omega dh_1} &= \frac{d^2 A_2}{d\omega dh_1} A_1 + \frac{d^2 A_2}{d\omega dh_1} A_4 + \frac{d^2 \phi_2}{d\omega dh_1} , \\
\frac{d^2 A_a}{d\omega dh_1} &= \frac{d^2 A_1}{d\omega dh_1} A_3 + \frac{d^2 A_1}{d\omega dh_1} A_4 + \frac{d^2 \phi_1}{d\omega dh_1} , \\
\frac{d^2 A_2}{d\omega dh_1} &= -\frac{1}{A_2} \frac{dA_2}{dh_1} \frac{dA_2}{d\omega} - \frac{2}{A_2} P_{m_{\omega h_2}} h_2 , \quad \frac{d^2 \phi_2}{d\omega dh_1} = -\frac{2}{A_2} \frac{dA_2}{dh_1} \frac{d\phi_2}{d\omega} - \frac{2}{A_2^2} P_{p_{\omega h_2}} h_2 .
\end{align*}
\]
Chapter 4

Application of the FRM Technique in Array Beamforming
4.1 Introduction

Digital array beamforming has been widely and successfully deployed in military and commercial applications [LL96, FL99, Ang00]. By exploiting the spatial diversity of sensors, digital sensor array beamformers receive/transmit signals from/to specific directions and attenuate signals from/to other directions, even if both the desired signals and the interferences occupy the same temporal frequency band. A conventional beamformer linearly combines the spatially sampled time sequences from each sensor to obtain a scalar output time sequence in the same manner that an FIR filter linearly combines temporally sampled data. The correspondence between FIR filtering and beamforming is closest when the beamformer operates at a single temporal frequency and the array geometry is linear and equi-spaced [VB88], or the so-called uniform linear array (ULA) beamformer. For a ULA beamformer, its response can be directly mapped to the frequency response of a corresponding FIR filter. The analogy between FIR filters and ULA beamformers led to the applications of FIR filter design methods in array beamformer synthesis [VB88, Van02].

Beampattern characteristics fundamentally affect how the acquired sensor data are processed. The most important parameters are mainlobe width, sidelobe level and transition bandwidth. In general sharp transition bands and low sidelobes are desirable properties of a beamformer as transition bandwidth determines the spatial discrimination capability and sidelobe level determines the interferences and noise suppression capability. Analogous to FIR filter design, the spatial discrimination capability of a ULA beamformer depends on the size of the spatial aperture. To achieve a desirable beampattern with sharp transition bands and low sidelobes, a large number of sensors is required by
beamformers designed by conventional methods, leading to high costs and heavy computational loads. Although FIR filter design methods, such as windowing methods, have been applied to array beamformer synthesis in the aspect of sidelobe level reduction [VB88], no attempt has been made in reducing the number of sensors while maintaining the same transition bandwidth and sidelobe level. This has motivated the present feasibility study of the applications of the FRM technique in efficient array beamforming, with emphasis on the reduction of the number of sensors and hence the computational complexity of the associated beamforming algorithm.

To facilitate the discussion on array beamforming using the FRM technique, some FRM filter implementation issues such as the computational complexity, the required memory and the effective filter length are studied in detail Section 4.2. Passive array beamforming using the FRM technique is presented in Section 4.3 and an analysis is carried out to show the infeasibility of reducing the number of sensors. Subsequently, active array beamforming based on the concept of effective aperture [GA81] and the FRM technique is discussed in Section 4.4. Using the proposed method, it is possible to reduce the number of sensors for active array beamforming, as verified by simulations. Generalizations to two-dimensional (2D) active array beamforming is discussed in Section 4.5 and conclusions are given in Section 4.6.
4.2. Analysis of the FRM Technique in Time Domain

\[ H_a(z^M) \rightarrow v(k) \rightarrow H_{ma}(z) \rightarrow \begin{array}{c} \oplus \end{array} \rightarrow y(k) \]

\[ z^{-\frac{(N_m-1)M}{2}} \rightarrow u(k) \rightarrow H_{mc}(z) \rightarrow \begin{array}{c} \oplus \end{array} \rightarrow y_{bc}(k) \]

Figure 4.1: A realization structure for a filter using the FRM technique.

\[ \delta \leq 1.384 \times 10^{-2.03} \]

Figure 4.2: Condition for \( N_m \geq M \). Maximum passband ripple \( \delta_p \) and minimum stopband attenuation \( \delta_s \) are expressed in decibels.
4.2 Analysis of the FRM Technique in Time Domain

The FRM technique [Lim86] is an efficient method for implementing linear phase FIR filters with sharp transition bands. As the FRM technique is getting mature, it is of particular interest to look for applications in other areas. To facilitate the discussion on the FRM applications in efficient array beamforming in Section 4.3, implementation issues such as the computational complexity, the required memory and the effective filter length are studied in this section. The study of these issues is of interest in itself as well.

A realization structure for a filter using the FRM technique is illustrated in Figure 4.1, in which a number of subfilters are involved. They are a prototype filter \( \{ h_a(n) \} \) and two masking filters \( \{ h_{ma}(n) \} \), \( \{ h_{mc}(n) \} \) of length \( N_a \), \( N_{ma} \) and \( N_{mc} \), respectively. Let \( \{ h_{ua}(n) \} \) be the interpolated filter obtained from \( \{ h_a(n) \} \) with \( M - 1 \) zeros inserted between adjacent taps. As \( \{ h_{ua}(n) \} \) is normally assumed to be an even order linear phase filter, its complementary filter can be obtained easily with a simple delay line. The transition bandwidth of \( \{ h_{ua}(n) \} \) is \( 1/M \) of that of \( \{ h_a(n) \} \) and multiple spectrum replicas appear in \([0, 2\pi)\).

Using properly designed masking filters, some unwanted spectrum replicas are removed while the rest is integrated in passband synthesis. Denote \( N_m \) the maximum length of the masking filters, \( i.e. N_m = \max(N_{ma}, N_{mc}) \), and adjust length of both masking filters to \( N_m \) by appropriate zero-padding if necessary.

Although FRM filtering is well known to be computationally efficient, it was found that its implementation requires well organized memory to hold a large number of input samples in [LYZF03], where the discussion was focused on field
4.2. Analysis of the FRM Technique in Time Domain

programmable gate array implementation of digital filters synthesized using the FRM technique. Referring to Figure 4.1, assume that the input samples are divided into $M$ subgroups of $N_a$ samples each. When a new sample $x(k)$ is available, one subgroup is updated by replacing the oldest sample with $x(k)$. The subgroup is filtered by $\{h_a(n)\}$ as

$$v(k) = \sum_{n=0}^{N_a-1} h_a(n) x(k - nM), \quad (4.1)$$

which requires $N_a$ multiplications (for direct form implementation). The intermediate output is subsequently filtered by the masking filters, leading to

$$y_{ba}(k) = \sum_{n=0}^{N_m-1} h_{ma}(n) v(k - n), \quad (4.2a)$$

$$y_{bc}(k) = \sum_{n=0}^{N_m-1} h_{mc}(n) u(k - n), \quad (4.2b)$$

where $u(k) = x\left(k - \frac{(N_a-1)M}{2}\right) - v(k)$ and about $2N_m$ multiplications are required (or to be exact, $N_{ma} + N_{mc}$). Because $y_{ba}(k)$ and $y_{bc}(k)$ are functions of $N_m - 1$ previous outputs of $\{h_a(n)\}$, the final output, $y(k) = y_{ba}(k) + y_{bc}(k)$, is actually a linear combination of the following input samples,

$$x(k - (N_aM + N_m - M - 1)), \quad x(k - (N_aM + N_m - M - 2)), \quad \ldots, \quad x(k), \quad (4.3)$$

where $x(n) = 0$ for $n < 0$. As we can see, the computational complexity is $N_a + 2N_m$ multiplications per sample when $k$ is larger than $N_aM + N_m - M$. Meanwhile, memory holding $N_aM$ latest input samples and $2N_m$ intermediate outputs is constantly maintained throughout the filtering process.
4.2. Analysis of the FRM Technique in Time Domain

Assume that the condition $N_m \geq M$ holds ($N_m < M$ is of non-interest practically, as to be discussed shortly), then the effective filter length of the above FRM filter is equal to $N_aM + N_m - M$ [Lim86]. Let $F_s$ be the sampling frequency and $\Delta F$ be the transition bandwidth of a designed lowpass FRM filter. The normalized transition bandwidth of the filter is defined as $\Delta f = \Delta F / F_s$. 

The FRM technique is efficient when $\Delta f$ is very small, i.e., $\Delta f \ll 1/16$, which is assumed throughout the chapter (Note that $\Delta f$ is the same as $\beta$ introduced in [LL93]). To meet the same specifications, the length of an optimum (Remez) FIR filter, $N_o$, is slightly shorter than the effective filter length of the FRM filter [LL93].

Some key features related to filter implementation between the FRM filter and the corresponding optimum FIR filter are compared in Table 4.1, which are the effective filter length, the memory size and the computational complexity (in terms of number of multiplications). Despite the slightly longer effective filter length and the increased memory size for the case of $N_m \geq M$, the computational complexity of the FRM filter is considerably reduced compared to the optimum filter. The reduction in computational complexity is mainly attributed to the sparseness of the interpolated filter $\{h_{ua}(n)\}$.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Effective filter length</th>
<th>Memory size</th>
<th>Complexity</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_m \geq M$</td>
<td>$N_aM + N_m - M$</td>
<td>$N_aM + 2N_m$</td>
<td>$N_a + 2N_m$</td>
<td>$N_m \geq M$</td>
</tr>
<tr>
<td>$N_o$</td>
<td>$N_o$</td>
<td>$N_o$</td>
<td>$N_o$</td>
<td>$N_m \geq M$</td>
</tr>
</tbody>
</table>

In the FRM filter design, the factor $M$ affects the complexity of the subfilters. As to be seen in Section 4.3, the effective filter length of the FRM filter is actually the total number of sensors to be deployed in passive array beamformers. It is
therefore important to establish the relationship between $M$ and $N_m$.

For a given $M$, the transition bandwidths of the masking filters are

$$\Delta_{ma} = \frac{2\pi - w_{ap} - w_{as}}{M},$$
$$\Delta_{mc} = \frac{w_{ap} + w_{as}}{M},$$

where $w_{ap}$ and $w_{as}$ represent the passband and stopband cutoff frequencies of the prototype filter, respectively [Lim86]. Define the mean of $\Delta_{ma}$ and $\Delta_{mc}$

$$\Delta_{mid} = \frac{(\Delta_{ma} + \Delta_{mc})}{2} = \frac{\pi}{M},$$

which is a function of $M$.

Given a $\Delta f$, and magnitudes of the passband and stopband ripples, $\delta_1$ and $\delta_2$, the filter length can be approximated by the well known Kaiser’s equation [PM96],

$$N = \frac{-20\log_{10}\sqrt{\delta_1\delta_2} - 13}{14.6\Delta f} + 1.$$  

For the same magnitudes of the ripples, the filter lengths, $N_{ma}$, $N_{mc}$ and $N_{mid}$, corresponding to the transition bandwidth $\Delta_{ma}$, $\Delta_{mc}$ and $\Delta_{mid}$ respectively, satisfy the following inequality,

$$N_m = \max(N_{ma}, N_{mc}) \geq N_{mid} \geq \min(N_{ma}, N_{mc}).$$

It is assumed that the ripple magnitudes of the masking filters, $\delta_1$ and $\delta_2$, are 15% smaller than the allowed magnitudes of the designed filter, $\delta_p$ and $\delta_s$, respectively.
4.2. Analysis of the FRM Technique in Time Domain

[Lim86]. Using (4.5) and (4.6), it can be found that \( N_{mid} \geq M \) is equivalent to

\[
\delta_1 \delta_2 = 0.85^2 \delta_p \delta_s \leq 10^{0.73/M - 2.03}.
\]  

(4.8)

It follows from (4.7) that \( N_m \geq M \) when (4.8) is satisfied. As \( M \) is a positive integer,

\[
\delta_p \delta_s < 1.384 \times 10^{-2.03}
\]

(4.9)

gives a sufficient condition for \( N_m \geq M \) in terms of ripple magnitudes of the designed filter, as depicted in Figure 4.2. It is obvious from Figure 4.2 that for nontrivial FIR filters designed by the FRM technique, \( i.e. \), filters with maximum passband ripple less than 1 dB and minimum stopband attenuations larger than 13 dB, \( N_m \geq M \) is always satisfied. For example, in the design example given in [Lim86], the allowed maximum passband ripple \( \delta_p = 0.0115 \) (0.2 dB) and minimum stopband ripple \( \delta_s = 0.01 \) (40 dB) fall inside the region of \( N_m \geq M \) indicated in Figure 4.2. Hence, we should have \( N_m \geq M \). Indeed, it was shown in an example of [Lim86] that for each \( M \) \((M = 2, \ldots, 14)\), the corresponding \( N_m \) that resulted in minimum filter complexity was always greater than \( M \). For example, for \( M = 6 \) or 9, we have \( N_m = 33 \) or 41. Thus we conclude that it is generally true that \( N_m \) is no less than \( M \). This result has an important implication in the next section.
4.3 Passive Array Beamforming Using the FRM Technique

In this section we begin with a review on the relationship between narrowband array beamforming and FIR filtering [VB88], and then discuss the feasibility of the applications of the FRM technique in passive array beamforming.

4.3.1 Passive Array Beamforming

The frequency response of an FIR filter with an impulse response \( \{ h(n) \} \), \( 0 \leq n \leq N - 1 \), is given by

\[
H(\omega) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n},
\]

which represents the response of the filter to a complex sinusoid of frequency \( \omega \).

Similarly, beamformer response is defined as the amplitude and phase presented to a complex plane wave, \( e^{j\omega k} \), as a function of direction of arrival (DOA) \( \theta \) and frequency \( \omega \). For an \( N \)-sensor beamformer with an aperture function that is the same as the FIR filter \( \{ h(n) \} \), the beamformer response is [VB88]

\[
H(\theta, \omega) = \sum_{n=0}^{N-1} h(n)e^{-j\omega \tau_n(\theta)},
\]

where \( \tau_0(\theta) = 0 \) and \( \tau_n(\theta), 1 \leq n \leq N - 1 \), represents time delay due to propagation relative to the first sensor element.

The correspondence between FIR filtering and beamforming is closest when the beamformer operates at a single temporal frequency \( \omega = \omega_o \) and the array geometry is linear and equi-spaced, or the so-called ULA beamformer. Let the
inter-sensor spacing be \(d\), wavelength be \(\lambda\), and DOA relative to broadside be \(\theta\). Then the delay can be expressed as \(\tau_n(\theta) = 2\pi n \frac{d \sin \theta}{\lambda \omega_o}\) and the relationship between the temporal frequency \(\omega\) in FIR filtering and the direction \(\theta\) in array beamforming can be identified as \(\omega = 2\pi \frac{d \sin \theta}{\lambda}\). Thus, temporal frequency in FIR filtering corresponds to the sine of direction in narrowband ULA beamforming.

To avoid spatial ambiguity and increase the spatial resolution of a ULA beamformer, \(d = \lambda/2\) is usually used and in such a case, each \(\theta \in [-\pi/2, \pi/2]\) is uniquely mapped to one \(\omega \in [-\pi, \pi]\). The corresponding beamformer response becomes \(H(\theta, \omega_o) = H(\pi \sin \theta) = H(\phi)\).

The analogy between FIR filtering and ULA beamforming led to the applications of FIR filter design methods in array beamformer synthesis [Van02]. As the FRM technique is successful in FIR filter synthesis with considerable reduction in computational complexity, it is natural to ask whether we could apply the FRM subfilters as aperture functions to ULAs to achieve desirable beampatterns with fewer sensors compared with traditional beamformer design methods. The answer is negative for passive arrays, as to be discussed briefly in the following. To simplify the presentation, \(H(\phi)\) is adopted as the response of the beamformer with an aperture function \(\{h(n)\}\).

### 4.3.2 Interleaved Linear Array Beamformers

An intuitive idea in applying the FRM technique to array beamforming with a hope to reduce the number of sensors is to construct a linear sensor array consisting of several subarrays, each being associated with one FRM subfilter. A linear array comprising one \(N_m\)-sensor dense ULA with inter-sensor spacing \(d = \lambda/2\) and one \(N_a\)-sensor sparse ULA with inter-sensor spacing \(Md = M\lambda/2\)
4.3. Passive Array Beamforming Using the FRM Technique

Figure 4.3: An interleaved linear array beamformer using one sparse ULA in solid circles and one dense ULA in squares with \( M = 3 \).

is shown in Figure 4.3. This beamformer is called as interleaved linear array beamformer. Four sub-beamformers are formed when the subfilters \( \{h_{ma}(n)\} \) and \( \{h_{mc}(n)\} \) are applied as aperture functions for the dense ULA and \( \{h_a(n)\} \) and its complementary filter, \( \{h_c(n)\} \), for the sparse ULA. Without loss of generality, let the phase be zero at the first sensor, \( i.e. \ x_0(k) = e^{j\omega_0 k} \). The four
4.3. Passive Array Beamforming Using the FRM Technique

Sub-beamformer outputs responding to the signal from θ can be written as

\[ y_a(k) = x_0(k) \sum_{n=0}^{N_a-1} h_a(n) e^{-jnM\phi} = e^{j\omega_o k} H_a(M\phi), \]

\[ y_c(k) = x_0(k) \sum_{n=0}^{N_a-1} h_c(n) e^{-jnM\phi} = e^{j\omega_o k} H_c(M\phi), \]

\[ y_{ma}(k) = x_0(k) \sum_{n=0}^{N_m-1} h_{ma}(n) e^{-jn\phi} = e^{j\omega_o k} H_{ma}(\phi), \]

\[ y_{mc}(k) = x_0(k) \sum_{n=0}^{N_m-1} h_{mc}(n) e^{-jn\phi} = e^{j\omega_o k} H_{mc}(\phi). \]

However, if we linearly combine these sub-beamformer outputs, for example, by simply adding them together, we are unable to generate the desirable output in the form of \( e^{j\omega_o k} (H_a(M\phi)H_{ma}(\phi) + H_c(M\phi)H_{mc}(\phi)) \). Despite the similarity of the interleaved linear array beamformer to a temporal FRM filter, we fail to obtain the required array response \( H(\phi) = H_a(M\phi)H_{ma}(\phi) + H_c(M\phi)H_{mc}(\phi) \). The reason lies at the difference between temporal filtering and array beamforming. As analyzed in Section 4.2, in the FRM filtering process, though \( N_a + 2N_m \) multiplications are required for each instance, it involves \( N_a \) input samples from the memory for \( H_a(z^M) \) and \( 2N_m \) intermediate outputs from the memory for \( \{v(n)\} \) and \( \{u(n)\} \). Since these intermediate outputs in turn are linear combinations of past input samples, the FRM filter output at each instance is in fact a linear combination of past samples given in (4.3), most of which are available in the memory. Throughout the whole filtering process, memory for holding the signal samples is indispensable. However, in array beamforming, the spatial samples are linearly combined at each snapshot. There exists no mechanism to hold the spatial samples similar to that in the temporal filtering process. With the limited number of spatial samples provided by the proposed sensor array in
4.3. Passive Array Beamforming Using the FRM Technique

Figure 4.4: A typical multiple subarray beamformer with $N_m = M + 2$. The sensors shared by neighbouring subarrays are marked with a dot in a square.

Figure 4.3, beamforming equivalent to that with a ULA with $N_o$ sensors cannot be achieved.

4.3.3 Multiple Subarray Beamformers

Having shown the infeasibility of applying the FRM technique in the interleaved linear array beamformers, we consider another kind of beamformers, which is closely analogous to the FRM technique.

Figure 4.4 depicts a multiple subarray beamformer. It consists of $N_a$ subarrays each of which consists of $N_m$ sensors with inter-sensor spacing $d = \lambda/2$. Adjacent subarrays are displaced by $Md = M\lambda/2$. In addition, sensors at the common nodes are shared by neighbouring subarrays. The spatial samples in
each subarray are shared and weighted with the aperture functions \( \{ h_{ma}(n) \} \) and \( \{ h_{mc}(n) \} \) simultaneously. Subsequently, the subarray outputs are weighted with aperture functions \( \{ h_a(n) \} \) and \( \{ h_c(n) \} \). Assume that a narrowband plane wave impinges upon the linear array from angle \( \theta \) relative to broadside, and the spatial sample at the first sensor is \( x_0(k) = e^{j\omega_0 k} \). It can be easily shown that the beamformer output \( y(k) \) is

\[
y(k) = y_{ba}(k) + y_{bc}(k) = e^{j\omega_0 k} (H_a(M\phi)H_{ma}(\phi) + H_c(M\phi)H_{mc}(\phi)) \tag{4.12}
\]

and the beamformer response is

\[
H(\phi) = H_a(M\phi)H_{ma}(\phi) + H_c(M\phi)H_{mc}(\phi), \tag{4.13}
\]

which is identical to an FRM filter response.

From the array layout of the above multiple subarray beamformer, it is clear that, with the assumption of \( N_m \geq M \), the total number of sensors is equal to \( N_aM + N_m - M \), which is slightly larger than \( N_o \), the number of sensors for an equivalent ULA beamformer whose aperture function is designed using the Remez algorithm. For completeness, we mention that for the case of \( N_m < M \), the total number of sensors, \( N_aN_m \), could be smaller than \( N_o \). However, as we discussed in the previous section, the resultant FRM filter would be a poor one unable to meet the required frequency specifications. Again, we fail to reduce the number of sensors by applying the FRM to multiple subarray beamformers.

In summary, we have shown that infeasibility of applying the FRM technique in passive array beamforming through an analysis of the difference between temporal filtering and passive array beamforming. Nevertheless, for active array
4.4. Active Array Beamforming Using the FRM Technique

Active array beamforming is widely deployed in contemporary radar and sonar systems, ultrasonic diagnostic systems, etc. to remotely measure environmental parameters or detect objects of interest. An active array beamformer comprises a multitude of sensor elements. Subsets of these elements form the apertures that are used for transmission or reception. At each excitation, the transmitted waves are weighted before propagating. The wavefront is then reflected when it hits the object. The scattered wavefront is re-sampled and converted to electrical signals by the receiving elements.
4.4. Active Array Beamforming Using the FRM Technique

Figure 4.6: Aperture functions and effective aperture of a VSA with $p = 3$ and $\cos^2(\cdot)$ apodized functions.

Figure 4.7: An AILA beamformer using a sparse transmitting ULA (in solid circles) and a dense receiving ULA (in hollow squares).
4.4. Active Array Beamforming Using the FRM Technique

As the number of sensors directly affect the system cost, Von Ramm et al. were among the first to propose an approach to reduce the number of elements in a linear array while minimizing grating lobes caused by sparseness of the array layout [VST75]. They proposed different spacings for the transmitting and receiving elements so that the transmitting and receiving grating lobes could be moved to different positions in the two-way radiation pattern where their contributions would destructively interfere. The basic idea behind this approach is the concept of effective aperture of an active array, which will be briefly reviewed as follows.

Let \( \{h_t(n)\} \) and \( \{h_r(n)\} \) denote the aperture functions associated with the transmitting and receiving arrays, respectively. The effective aperture of an active array is the receiving aperture which would produce an identical two-way radiation pattern if the transmitting aperture were a point source [GA81]. Mathematically, the effective aperture function is simply the convolution of \( \{h_t(n)\} \) and \( \{h_r(n)\} \), i.e.,

\[
h_e(n) = h_r(n) * h_t(n),
\]

where \( * \) represents the convolution operation. Hence, unlike for passive array, one can come up with the same effective aperture function with different combinations of \( \{h_t(n)\} \) and \( \{h_r(n)\} \). This was investigated in detail in the context of designing sparse linear array suitable for imaging systems [LLOF96, LTB98, BL97].

We now briefly review a design method for sparse linear arrays [LLOF96]. Figure 4.5 shows an example of an effective aperture. In this example, the effective aperture is rectangular, with 16 elements and \( \lambda/2 \) element spacing. There
are many different ways of selecting sparse transmitting and receiving aperture functions to yield the same effective aperture and the corresponding two-way radiation pattern. For example, we can use a single-element transmitting array and a 16-element receiving array with $\lambda/2$ spacing as shown in Figure 4.5(b). Alternatively, we can use a four-element transmitting array with $\lambda/2$ spacing and a four-element receiving array with $2\lambda$ spacing in Figure 4.5(c). As it can be seen, the latter design uses only 8 elements instead of 16 elements in the former. In this example, the aperture functions for the transmitting and receiving arrays are rectangular. In practice, apodization can be used to control the shape of the effective aperture, offering more flexibility in designing an active array [LLOF96].

Among the various strategies for designing sparse arrays proposed in [LLOF96], the best one is the vernier sparse array (VSA), which is analogous with a linear vernier scale. In a VSA, by spacing the transmitting elements $pd$ apart and the receiving elements $(p - 1)d$ apart where $p$ is an integer, the convolution of the aperture functions will yield an effective aperture with elements spaced $d$ apart. Figure 4.6 shows the aperture functions and the associated effective aperture of a VSA consisting of a 10-element transmitting array and a 10-element receiving array.

Note all the existing methods for designing an active array using the concept of effective aperture [LLOF96, LTB98, BL97, AH02] tried to either avoid the grating lobes by eliminating the periodicity of the sparse arrays or attenuate the grating lobes by introducing nulls in the effective aperture functions. Using the FRM technique, it will be shown next that instead of suppressing all the grating lobes completely, some of them can be integrated in mainlobe synthesis.
by exploiting the complementary property of \( \{ h_a(n) \} \) and \( \{ h_c(n) \} \) and through properly designing the masking filters.

### 4.4.1 Active Interleaved Linear Array Beamformers and Effective Aperture

Motivated by our discussion on passive array beamforming, we propose an active interleaved linear array (AILA) beamformer comprising one transmitting array and one receiving array, as presented in Figure 4.7. The receiving array is an \( N_r \)-sensor ULA with inter-sensor spacing \( d = \lambda / 2 \) while the transmitting array is an \( N_t \)-sensor uniform sparse array with inter-sensor spacing \( M d = M\lambda / 2 \). As before, denote \( \{ h_t(n) \} \) and \( \{ h_r(n) \} \) the aperture functions associated with the transmitting and receiving arrays, respectively.

Take the first sensor on the left of the AILA as a reference and assume an object is located at \( \theta \) relative to broadside in the far field. The narrowband signal (of center frequency \( \omega_o \)) is weighted before transmitting. Without loss of generality, assume the signal from the first sensor arrives at the object at phase 0. The resultant signal scattered from the object can be expressed as a summation of the signals from all transmitting elements,

\[
x_s(k) = \sum_{n=0}^{N_t-1} h_t(n)e^{j\omega_0 k}e^{-jnM\phi} = e^{j\omega_0 k}H_t(M\phi).
\] (4.15)

The scattered plane wave is then spatially sampled by the receiving array and weighted with the receiving aperture function \( \{ h_r(n) \} \). The output signal be-
4.4. Active Array Beamforming Using the FRM Technique

comes

\[ y(k) = \sum_{n=0}^{N_r-1} h_r(n)x_s(k)e^{-jn\phi} = e^{j\omega_o k}H_t(M\phi)H_r(\phi). \quad (4.16) \]

By (4.14), the effective aperture of the AILA is

\[ h_e(n) = h_r(n) * h_{ut}(n), \quad (4.17) \]

where

\[ h_{ut}(m) = \begin{cases} 
  h_t(m/M), & \text{for } m = 0, M, \ldots, (N_t - 1)M, \\
  0, & \text{otherwise}. 
\end{cases} \]

4.4.2 Aperture Functions by the FRM Technique

If \( \{h_a(n)\} \) and \( \{h_{ma}(n)\} \) are taken as the transmitting and receiving aperture functions, respectively, the output signal (4.16) becomes

\[ y_{ba}(k) = e^{j\omega_o k}H_a(M\phi)H_{ma}(\phi). \quad (4.18) \]

When the above process is repeated for the second excitation with the same transmitting signal, and the new transmitting and receiving aperture functions are \( \{h_e(n)\} \) and \( \{h_{me}(n)\} \), respectively, the corresponding output signal becomes

\[ y_{be}(k) = e^{j\omega_o k}H_e(M\phi)H_{me}(\phi). \quad (4.19) \]

The AILA beamformer output is the sum of the output signals (4.18) and
(4.19) for the two excitations,

$$y(k) = e^{j\omega_k} (H_a(M\phi)H_{ma}(\phi) + H_c(M\phi)H_{mc}(\phi)). \quad (4.20)$$

As seen from (4.18) and (4.19), the AILA beamformer actually makes use of two effective aperture functions

$$h_{ea}(n) = h_{ua}(n) \ast h_{ma}(n),$$
$$h_{ec}(n) = h_{uc}(n) \ast h_{mc}(n),$$

in which \(\{h_{ua}(n)\}\) and \(\{h_{uc}(n)\}\) are the interpolated filters of \(\{h_a(n)\}\) and \(\{h_c(n)\}\), respectively.

The novel combination of the concept of effective aperture and the FRM technique leads to the synthesis of desirable beamformers. Specifically, with the properly designed complementary filter pair and the masking filters, some of the grating lobes are integrated into the mainlobe synthesis instead of being suppressed completely. As a result, these beamformers have effective beampatterns with sharp transition bands and low sidelobes, and can be implemented with only \(N_a + N_m\) sensors, much less than \(N_aM + N_m - M\) (or \(N_o\)) sensors using conventional beamformer design techniques.

### 4.4.3 Computer Simulations

A lowpass filter with a sharp transition band at \((0.174\pi, 0.191\pi)\) (corresponding to \((10^\circ, 11^\circ)\) in azimuth domain) and minimum 40 dB stopband attenuation is synthesized. The frequency specifications of the FRM subfilters are presented...
Table 4.2: Frequency specifications of the FRM subfilters with $M = 8$.

<table>
<thead>
<tr>
<th>Subfilter</th>
<th>Filter length</th>
<th>Passband cutoff</th>
<th>Stopband cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>${h_a(n)}$</td>
<td>35</td>
<td>$0.472\pi$</td>
<td>$0.608\pi$</td>
</tr>
<tr>
<td>${h_{ma}(n)}$</td>
<td>39</td>
<td>$0.076\pi$</td>
<td>$0.191\pi$</td>
</tr>
<tr>
<td>${h_{mc}(n)}$</td>
<td>31</td>
<td>$0.174\pi$</td>
<td>$0.309\pi$</td>
</tr>
</tbody>
</table>

Figure 4.8: Spatial filtering simulation of the AILA beamformer. (a) The beamformer output signal when a single object is present at $10^\circ$ or $11^\circ$ relative to broadside. (b) Scattered signals from two objects present simultaneously at the first excitation. (c) Scattered signals from two objects present simultaneously at the second excitation. (d) The AILA beamformer output when two objects present simultaneously.
4.4. Active Array Beamforming Using the FRM Technique

Figure 4.9: The effective beampattern of the AILA beamformer.

Table 4.3: Frequency specifications of the aperture functions in a narrow-beamwidth beamformer synthesis.

<table>
<thead>
<tr>
<th>Aperture functions</th>
<th>Type</th>
<th>Passband cutoff</th>
<th>Stopband cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>{h_a(n)}</td>
<td>Lowpass</td>
<td>(M\omega_p)</td>
<td>(M\omega_s)</td>
</tr>
<tr>
<td>{h_{ma}(n)}</td>
<td>Lowpass</td>
<td>(\omega_p)</td>
<td>((2\pi - M\omega_p)/M)</td>
</tr>
</tbody>
</table>

Table 4.4: Frequency specifications of the FRM subfilters with \(M = 7\).

<table>
<thead>
<tr>
<th>Subfilter</th>
<th>Filter length</th>
<th>Passband cutoff</th>
<th>Stopband cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>{h_a(n)}</td>
<td>35</td>
<td>0.394(\pi)</td>
<td>0.509(\pi)</td>
</tr>
<tr>
<td>{h_{ma}(n)}</td>
<td>33</td>
<td>0.342(\pi)</td>
<td>0.499(\pi)</td>
</tr>
<tr>
<td>{h_{mc}(n)}</td>
<td>33</td>
<td>0.229(\pi)</td>
<td>0.358(\pi)</td>
</tr>
</tbody>
</table>
Figure 4.10: Beampattern comparison between the AILA beamformer ($N_t = 10$, $N_r = 30$, $M = 10$) and the VSA beamformer with $p = 3$ and $\cos^2(\cdot)$ aperture function (24 sensors in the transmitting array and 24 sensors in the receiving array).

in Table 4.2.

The subfilters are applied as the aperture functions of the AILA beamformer and the simulation results are plotted in Figure 4.8. Assume a sinusoidal signal with frequency of 1 Hz is weighted with $\{h_a(n)\}$ and transmitted to the object. The transmitted waves reach the object at different phases. The scattered signal is re-sampled by the receiving array and weighted with the aperture function $\{h_{ma}(n)\}$. For the second excitation, the aperture function pair is changed to $\{h_e(n)\}$ and $\{h_{me}(n)\}$. Two cases are simulated. Case 1: When a single object is located at $10^\circ$ or $11^\circ$, the beamformer outputs of the object at $11^\circ$ is severely attenuated, as shown in Figure 4.8(a). Case 2: When the two objects are present simultaneously, the scattered signals for the two excitations are shown in Figure 4.8(b) and 4.8(c). Due to the difference in the distance, it
is assumed, without loss of generality, the scattered signal from the object at 11° experiences an extra $\pi/2$ phase difference. The scattered signals are then re-sampled and processed, leading to the AILA beamformer output in Figure 4.8(d). By comparing Figure 4.8(d) with 4.8(a), it can be seen that the signal from the object at 10° is dominant in the beamformer output.

We next generate the effective beampattern of the AILA beamformer. Assume an object is located in the direction of $\theta$ in the far field. The beamformer output power is recorded when $\theta$ varies from 0° to 40° at an increment of 0.1°, as shown in Figure 4.9. The transition from 10° to 11° suggests the excellent spatial discrimination capability of this beamformer. With the aperture functions obtained using the FRM technique, the AILA beamformer is capable of separating two closely spaced objects with only 74 sensors ($N_a = 35, N_m = 39$). However, to obtain a similar beampattern using a ULA beamformer whose aperture function is designed using the Remez algorithm, at least 215 sensors are required.

The only drawback of the AILA beamformer is that two excitations and two alternating aperture function pairs are required to complete one cycle. However, in narrow-beamwidth beamformer synthesis, only one excitation and one pair of aperture functions are needed. It can be regarded as a special case of broad-beamwidth beamformer synthesis described above. In this case, the grating lobes caused by the sparseness of the transmitting array are attenuated by the properly designed receiving aperture function. For example, to have a desired beampattern with mainlobe cutoffs at $\omega_p$ and $\omega_s$, the frequency specifications of the prototype filter and the masking filter are given in Table 4.3.

The synthesis of a narrow-beamwidth beamformer is compared between an
4.5. Active Two-Dimensional Array Beamforming

Figure 4.11: A typical active 2D array comprises one $11 \times 7$ sparse transmitting array in solid circles and one $17 \times 9$ dense receiving array in blank squares with $M_1 = 3$ and $M_2 = 2$.

AILA beamformer using the proposed method and a VSA beamformer using the method of [LLOF96]. Both the transmitting and receiving arrays of the VSA consist of 24 sensors and a $\cos^2(\cdot)$ is applied as the individual aperture function with $d = \lambda/2$ and $p = 3$. By using the FRM technique, the transmitting and receiving aperture functions are redesigned for an AILA beamformer with a 10-sensor transmitting array and a 30-sensor receiving array. The effective beampatterns of the VSA and AILA beamformers are very similar in terms of transition bandwidth and sidelobe level, as shown in Figure 4.10. Note that the VSA beamformer needs 48 sensors while the AILA beamformer employs only 40 sensors, a saving of 8 sensors or 20 percent.
4.5. Active Two-Dimensional Array Beamforming

Figure 4.12: Block diagram of 2D FRM filter synthesis.

Figure 4.13: Effective beampattern of the 2D array.
4.5 Active Two-Dimensional Array Beamforming

Two-dimensional (2D) FIR filter design using the FRM technique were explored in [LL98, LL99], in which the 2D frequency plane is divided into a number of complementary regions. By properly designing the masking filters, 2D FIR filters with sharp transitions can be synthesized. In this section, we generalize the beamforming method described in Section 4.4 to 2D active array beamformers with separable aperture functions.

A layout of an active 2D array is shown in Figure 4.11 with a reference sensor located at the left bottom corner. The receiving array is an \((N_{m_1} \times N_{m_2})\)-sensor array with inter-sensor spacing \(d = \lambda/2\) in both \(n_1\) and \(n_2\) directions while the transmitting array is an \((N_{a_1} \times N_{a_2})\)-sensor array with inter-sensor spacings \(M_1d = M_1\lambda/2\) and \(M_2d = M_2\lambda/2\) in the \(n_1\) and \(n_2\) directions, respectively.

Assume that an object is located in the direction of \((\theta_a, \theta_e)\) in the far field, where \(\theta_a\) and \(\theta_e\) are the azimuth angle and the elevation angle, respectively. The transmitted narrowband signal with center frequency \(\omega_o\), weighted with the transmitting aperture function \(\{h_t(n_1, n_2)\}, 0 \leq n_1 \leq N_{a_1} - 1, 0 \leq n_2 \leq N_{a_2} - 1\), hits at the object in the far field at different phases, leading to the scattered signal

\[
x_s(k) = \sum_{n_1=0}^{N_{a_1}-1} \sum_{n_2=0}^{N_{a_2}-1} e^{j\omega_0 k} h_t(n_1, n_2) e^{-2\pi k^T r/\lambda},
\]

where \(k\) is a unit direction vector pointing to the reference sensor of the transmitting array from the object and \(r\) is a sensor location vector with respect to
4.5. Active Two-Dimensional Array Beamforming

the reference sensor. Both vectors can be expressed in Cartesian coordinates

\[ \mathbf{k} = [k_x \ k_y \ k_z]^T = [\cos \theta_a \sin \theta_e \ \sin \theta_a \sin \theta_e \ \cos \theta_e]^T, \]

\[ \mathbf{r} = [r_x \ r_y \ 0]^T = [M_1 \ n_1 \ d \ M_2 \ n_2 \ d \ 0]^T, \]

where superscript \([\cdot]^T\) represents the transpose operation. Assume the transmitting aperture function is separable,

\[ h_t(n_1, n_2) = h_{a_1}(n_1)h_{a_2}(n_2). \quad (4.22) \]

The scattered signal can be simplified as

\[ x_s(k) = \sum_{n_1=0}^{N_{a_1}-1} \sum_{n_2=0}^{N_{a_2}-1} e^{j\omega k} h_{a_1}(n_1)h_{a_2}(n_2)e^{-2\pi M_1 \lambda k_x n_1 d}e^{-2\pi M_2 \lambda k_y n_2 d} \]

\[ = e^{j\omega k} H_{a_1}(M_1 \phi_1)H_{a_2}(M_2 \phi_2), \quad (4.23) \]

where \(\phi_1 = \pi k_x\) and \(\phi_2 = \pi k_y\). The scattered signal is then re-sampled by the receiving array and weighted with the aperture function \(\{h_r(n_1, n_2)\}\). Assume \(h_r(n_1, n_2)\) is also separable as \(h_r(n_1, n_2) = h_{ma_1}(n_1)h_{ma_2}(n_2)\). The effective response can be expressed as

\[ H_{aa}(\phi_1, \phi_2) = H_{a_1}(M_1 \phi_1)H_{a_2}(M_2 \phi_2)H_{ma_1}(\phi_1)H_{ma_2}(\phi_2). \quad (4.24) \]

As both transmitting and receiving aperture functions are separable, the whole spectrum \((-1 \leq k_x, k_y \leq 1)\) is divided into four regions. These four regions form a complementary set. Thus in 2D active array beamforming, four excitations are needed as shown in Figure 4.12. In addition to the effective response shown
in (4.24), the other three effective responses are

\[ H_{ac}(\phi_1, \phi_2) = H_{a1}(M_1\phi_1)H_{c2}(M_2\phi_2)H_{ma1}(\phi_1)H_{mc2}(\phi_2), \]  
(4.25a)

\[ H_{ca}(\phi_1, \phi_2) = H_{c1}(M_1\phi_1)H_{a2}(M_2\phi_2)H_{ma2}(\phi_1)H_{mc1}(\phi_2), \]  
(4.25b)

\[ H_{cc}(\phi_1, \phi_2) = H_{c1}(M_1\phi_1)H_{c2}(M_2\phi_2)H_{mc1}(\phi_1)H_{mc2}(\phi_2). \]  
(4.25c)

The effective active array beamformer response, as a summation of (4.24) and (4.25), can be written as

\[
H(\phi_1, \phi_2) = H_{aa}(\phi_1, \phi_2) + H_{ac}(\phi_1, \phi_2) + H_{ca}(\phi_1, \phi_2) + H_{cc}(\phi_1, \phi_2) \\
= (H_{a1}(M_1\phi_1)H_{ma1}(\phi_1) + H_{c1}(M_1\phi_1)H_{mc1}(\phi_1)) \\
  \times (H_{a2}(M_2\phi_2)H_{ma2}(\phi_2) + H_{c2}(M_2\phi_2)H_{mc2}(\phi_2)).
\]  
(4.26)

Hence, \( H(\phi_1, \phi_2) \) is the product of two 1D active array beamformer responses, which are of the same form as two 1D FRM filters. Similar to our discussion in the previous section, we can then synthesize 2D active beamformers with desirable beampatterns with fewer sensors than using other design techniques, as shown in the following example.

In this example, we simulate the effective beampattern of a 2D array using two 1D FRM filters. In addition to the 1D FRM filter presented in Table 4.2, another 1D FRM filter is designed to meet the same specification as described in Section 4.4.3, whose subfilters are shown in Table 4.4. 2D aperture functions are formed and applied to the 2D array with a \((35 \times 35)\)-sensor transmitting array and a \((39 \times 33)\)-sensor receiving array. Assume an object is located in the direction of \((\theta_a, \theta_e)\) in the far field, \((0 \leq \theta_a < 2\pi, 0 \leq \theta_e \leq \pi/2)\). The transmitted signals weighted with the 2D transmitted aperture function hit the object
4.6 Conclusions

The feasibility of the applications of the frequency-response masking (FRM) technique in digital array beamforming has been investigated in detail in this paper. Despite the reduced computational complexity in temporal FRM filtering, a large memory to hold the input samples is required throughout the filtering process. Because there is no mechanism similar to that of temporal filtering in array beamforming, a large number of sensor elements are required to provide enough spatial samples for processing in the passive array beamforming. Therefore, it is infeasible to apply the FRM technique in passive array beamforming to reduce the number of sensors while maintaining the same beampattern. However, the FRM technique does find its applications in active array beamforming by a novel combination with the concept of effective aperture. With fewer sensor elements, beampattern with sharp transition and low sidelobes can be achieved. The proposed active array beamforming method is also flexible in meeting a specific mainlobe width. The active array beam-
4.6. Conclusions

forming method has also been generalized to 2D active array beamforming, and illustrated by simulations.
Chapter 5

General Sidelobe Cancellers with Leakage Constraints
5.1 Introduction

In the past decades, adaptive beamforming has attracted a lot of interest in the fields of wireless communications, seismology and speech enhancement, etc. for signal detection and estimation. An adaptive beamformer is able to adjust its beampattern in real time to maintain the prescribed frequency responses in the desired directions while introducing nulls in the interference directions.

A typical adaptive array beamformer is the linear constrained minimum variance (LCMV) beamformer. In narrowband applications, a famous representative of the LCMV beamformer is the Capon beamformer [Cap69], while in broadband applications, the well studied LCMV beamformer is the Frost beamformer [Fro72]. Using linear constraints is a common approach that permits extensive control over the adapted response of the beamformer [VB88]. The

![Diagram of a broadband beamformer](image)

Figure 5.1: A broadband beamformer with $K$ sensors and $L$-tap delay lines, where $\tau_i$ ($i = 1, \ldots, K$) are presteering delays and $T_s$ is the sampling interval.
constraints would require the weight to either accentuate signals propagating from some direction or suppress the interferences [JD93].

In practice, the performance of an adaptive beamformers severely degrades when the ideal assumptions do not exist, such as mismatch between the direction of arrival (DOA) of the signal and the looking direction of the array, (in short, DOA mismatch), imperfect array calibration and distorted antenna shape, limited number of training snapshots, and imperfect knowledge of the statistical information of the signal and interference, etc. In this chapter, we restrain our discussion on robust beamforming against DOA mismatch. To suppress the signal cancellation caused by DOA mismatch, some constraints were proposed such as multipoint linear constraint [Nun83], soft quadratic response constraint [EC85], and maximally flat spatial power response derivative constraints, in short, derivative constraints [EC83]. With these constraints, an adaptive beamformer is robust in the vicinity of the assumed direction. The widened beamwidth of its beampattern is achieved at the cost of the decreased capability in interference suppression due to the reduction in the degree of freedom.

An alternative structure for implementing the LCMV beamformer is the general sidelobe canceller (GSC) [BS02, WA03, GJ82, FL96, HSH99], which consisting of a presteering front end, a fixed beamformer in the main channel, a blocking matrix and an adaptive canceller in the auxiliary channel. The presteering front end is composed of variable time delays, with which the main lobe of the GSC can be steered to the desired direction. The fixed beamformer is used to enhance the desired signal from the look direction. The blocking matrix is to prevent any signal from the look direction from passing through, so
that the output from the blocking matrix contains interferences and noise in the auxiliary channel. The adaptive canceller is able to adjust its weights so that the interferences and noise can be subtracted from the main channel output in an optimal way. As the interferences and noise are assumed to be uncorrelated with the target signal, the adaptation of the canceller is performed to minimize the output power.

Due to DOA mismatch, the blocking matrix cannot perfectly block the target signal, hence leakages will lead to target signal cancellation at the beamformer output. A modified GSC with two adaptive modules was proposed in [HSH99]. The method proposed there was robust to steering vector errors but control of two adaptive modules was difficult. The derivative constraints [EC83] were extended to GSCs in [FL96]. However, the adaptive beamformers with derivative constraints were too conservative as a large number of linear constraints were required, resulting in the degradation of the interferences and noise cancellation capability.

In this chapter, a new class of linear constraints, called leakage constraints, is proposed for GSCs. By minimizing the leakage influence on the adaptive canceller in a GSC, a set of leakage constraints is obtained. The number of leakage constraints is much smaller than that of derivative constraints and hence the proposed GSC with leakage constraints is robust against DOA mismatch but less conservative. The proposed modified GSC may be considered as a compromise of the conventional GSC and the GSC with derivative constraints, as to be discussed later.

The rest of the chapter is organized as follows. Derivative constraints for broadband beamformers are briefly reviewed in Section 5.2 where some notations
are introduced that are to be used in the rest of the chapter. The class of leakage constraint is derived in Section 5.3. The implementation issues are discussed in 5.4 and computer simulations are presented in Section 5.5, followed by conclusions in Section 5.6.

5.2 Derivative Constraints for Broadband Beamformers

There are two kinds of beamformers, narrowband and broadband beamformers. Narrowband beamformers sample the propagating wave field in space and linearly combine the spatial samples, while broadband beamformers sample the propagating wave field in both space and time and are often used when signals of significant frequency extent are of interest. Throughout the chapter, broadband beamformers are discussed.

A typical broadband array beamformer with a $K$-sensor ULA and $L$-tap delay lines is shown in Figure 5.1 and its output can be expressed as

$$y(n) = \sum_{k=1}^{K} \sum_{l=0}^{L-1} w_{k,l}^* x_k(n-l+1)$$

(5.1)

where $w_{k,l}$ and $x_k(n-l+1)$ are the weight coefficient and spatio-temporal sample at the $j$-th tap of the $k$-th delay line. Without loss of generality, let the phase be zero at the first sensor. We have $x_1(n) = e^{-j\omega n}$ and $x_k(n-l+1) = e^{-j\omega(n-\Delta_{k,l}(\theta))}$, where $\theta$, $\omega$ represent the DOA and the frequency of the signal, and $\Delta_{k,l}(\theta)$ represents the time delay associated with the $(k,l)$-th sample with reference to
5.2. Derivative Constraints for Broadband Beamformers

The beamformer output becomes

\[ y(n) = e^{j\omega n} \sum_{k=1}^{K} \sum_{l=0}^{L-1} w^*_k, l e^{-j\omega(k-\Delta_{k,l}(\theta))} \]  

(5.2)

and \( \Delta_{1,0}(\theta) = 0 \). As we can see that the beamformer output is a function of both temporal and spatial variables. Both (5.1) and (5.2) can be simplified as an inner product of two vectors \( y(n) = w^H x(n) \), where \( w \) and \( x(n) \in \mathbb{R}^{KL \times 1} \) and

\[ x(n) = [x_1(n) \ldots x_K(n) x_1(n-1) \ldots x_K(n-1) \\
\ldots x_1(n-L+1) \ldots x_K(n-L+1)]^T, \]

\[ w = [w_{1,0} \ldots w_{K,0} w_{1,1} \ldots w_{K,1} \ldots w_{1,L-1} \ldots w_{K,L-1}]^T. \]

The frequency response of the beamformer to the complex plane wave is then

\[ H(\theta, \omega) = w^H d(\theta, \omega) \] where \( d(\theta, \omega) \) is the array manifold vector

\[ d(\theta, \omega) = [1 \ e^{-j\omega \Delta_{1,0}(\theta)} \ldots e^{-j\omega \Delta_{K,0}(\theta)} \ldots e^{-j\omega \Delta_{K,L}(\theta)}]^T. \]

Assume the signal received at the sensors is wide sense stationary with zero mean. The expected power of the beamformer is given by

\[ P = \mathbb{E}\{|y(n)|^2\} = w^H \mathbb{E}\{xx^H\} w = w^H \mathbb{R}_{xx} w \]

where \( \mathbb{E}\{\ast\} \) denotes the expectation operation and \( \mathbb{R}_{xx} \in \mathbb{R}^{KL \times KL} \) is the correlation matrix of the received signal.

The basic idea of using the linear constraints is to constrain the response
5.3. Leakage Constraints for GSCs

of the beamformer so that the signal from the desired signal is passed with specified gain and phase. Thus, the LCMV problem can be written as

$$
\min_w \mathbf{P} = \min_w \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w}, \quad \text{subject to } \mathbf{w}^H \mathbf{d}(\theta, \omega) = g,
$$

(5.3)

where the constraint in (5.3) represents general linear constraints.

It is well known that statistically optimum beamformers suffer significant performance degradation when ideal assumptions do not exist in practice, typically, DOA mismatch between signal DOA and pointing direction. A well-known class of derivative constraints was proposed in [EC83]. Denote the power response of the beamformer \( \rho(\theta, \omega) = H^*(\theta, \omega)H(\theta, \omega) \). It is constrained to be approximately equal to \( \rho(\theta_0, \omega) \) at the vicinity of \( \theta_0 \) in a maximally flat sense by setting the first-order up to \( n \)-th-order partial derivatives of \( \rho(\theta, \omega) \) with respect to \( \theta \) to be zeros, i.e.,

$$
\left. \frac{\partial^p \rho(\theta, \omega)}{\partial \theta^p} \right|_{\theta_0} = 0, \quad p = 1, 2, \ldots, n.
$$

(5.4)

With these constraints, the beamformer is robust in the vicinity of the assumed direction. The widen beamwidth is achieved at the cost of the reduction capability in interference suppression due to reduce in the degree of freedom. Moreover, the number of constraints in this class is proportional to the order of the constraints. To be precise, the total number of derivative constraints, including the zeroth, first- and second-order constraints, is \( 3L \).
5.3 Leakage Constraints for GSCs

A GSC with $K$-sensor ULA and $L$-tap delay lines is shown in Figure 5.2. For simplicity of presentation, that the look direction of the GSC is steered to $0^\circ$, i.e., the direction perpendicular to the array axis.

Denote $w_F \in \mathcal{R}^{KL \times 1}$ and $w_M \in \mathcal{R}^{(K-1)L \times 1}$ the weight vectors of the fixed beamformer and the adaptive canceller respectively, the optimization problem of the GSC is formulated as

$$
\min_{w_M} P = \min_{w_M} (w_F - B_0^T w_M)^T R_{xx} (w_F - B_0^T w_M),
$$

where $B_0 \in \mathcal{R}^{(K-1)L \times KL}$ is the blocking matrix, acting as a simple spatial filter with a null at $0^\circ$.

Although the look direction of the GSC is assumed to be $0^\circ$, the actual DOA of the target signal, $\theta_s$, may not be $0^\circ$ exactly. It is reasonable to assume that the absolute value of $\theta_s$ is small. Assuming that the target signal is uncorrelated
with the interferences and noise, the correlation matrix of array signal can be written as

\[ R_{xx} = R_{nn} + R_{ss}(\theta_s), \]

where \( R_{nn} \) represents the overall array correlation matrix of interferences and noise, and \( R_{ss}(\theta_s) \) represents the array correlation matrix of the target signal. As \( \theta_s \) is small, the Taylor series expansion of \( R_{ss}(\theta_s) \) at the vicinity of \( 0^\circ \) is given by

\[
R_{ss}(\theta_s) = R_{ss}(0) + \frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \bigg|_{\theta_s=0} \theta_s + \frac{\partial^2 R_{ss}(\theta_s)}{\partial \theta_s^2} \bigg|_{\theta_s=0} \frac{\theta_s^2}{2!} + \ldots \quad (5.6)
\]

Since \( B_0 \) is able to block the signal from \( 0^\circ \), i.e.,

\[ B_0 R_{ss}(0) = 0, \]

the leakages at the blocking matrix in the presence of DOA mismatch (\( \theta_s \neq 0 \)) are attributed to those non-zero derivative terms in (5.6). We call

\[
\frac{\partial^p R_{ss}(\theta_s)}{\partial \theta_s^p} \bigg|_{\theta_s=0} \frac{\theta_s^p}{p!}, \quad p = 1, 2, \ldots
\]

the \( p \)-th order leakage. We can select the order of approximation in (5.6) based on accuracy requirement. For small \( \theta_s \), the first and second order leakages are more significant than higher order leakages. In the following, we derive the linear constraints based on the first and second order leakages. Higher order leakage constraints can be derived in a similar way.
5.3. Leakage Constraints for GSCs

5.3.1 First Order Leakage Constraint

Taking the first order leakage only, the objective function in (5.5) becomes

\[ P \approx (w_F - B_0^T w_M)^T \left( R_{nn} + R_{ss}(0) + \frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \bigg|_{\theta_s=0} \theta_s \right) (w_F - B_0^T w_M). \]  

(5.7)

With careful observation, the influence of the first order leakage in (5.7) on the adaptation of \( w_M \) is mitigated if

\[ w_F^T \frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \bigg|_{\theta_s=0} B_0^T w_M = 0, \]  

(5.8)

and

\[ w_M^T B_0 \frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \bigg|_{\theta_s=0} B_0^T w_M = 0. \]  

(5.9)

Next we are going to show that (5.9) is always true.

Note that

\[ B_0 R_{ss}(0) = 0, \text{ and hence } B_0 R_{ss}(0) B_0^T = 0. \]  

(5.10)

Moreover, as \( B_0 R_{ss}(\theta_s) B_0^T \) is always positive semidefinite for any \( \theta_s \), we have

\[ y^T \left( B_0 R_{ss}(\theta_s) B_0^T \right) y \geq 0. \]  

(5.11)

for any \( y \in \mathcal{R}^{(K-1)L \times 1} \). From (5.10), it is obvious that

\[ y^T \left( B_0 R_{ss}(0) B_0^T \right) y = 0. \]  

(5.12)
5.3 Leakage Constraints for GSCs

Eqns. (5.11) and (5.12) imply
\[
\frac{\partial}{\partial \theta_s} \left( y^T \left( B_0 R_{ss}(\theta_s) B_0^T \right) y \right) \bigg|_{\theta_s=0} = y^T \left( B_0 \frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \bigg|_{\theta_s=0} B_0^T \right) y = 0. \tag{5.13}
\]

As \( y \) is an arbitrary vector, from (5.13) we have
\[
B_0 \frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \bigg|_{\theta_s=0} B_0^T = 0.
\]

So (5.9) is always true, and only (5.8) gives an effective constraint.

5.3.2 Second Order Leakage Constraints

Taking the second order leakage only, the objective function in (5.5) becomes
\[
P \approx (w_F - B_0^T w_M)^T \left( R_{nn} + R_{ss}(0) + \frac{\partial^2 R_{ss}(\theta_s^2)}{\partial \theta_s^2} \bigg|_{\theta_s=0} \right) \left( \frac{\theta_s^2}{2!} \right) (w_F - B_0^T w_M).
\]

The constraints for the second order leakage are given by
\[
w_F^T \frac{\partial^2 R_{ss}(\theta_s^2)}{\partial \theta_s^2} \bigg|_{\theta_s=0} B_0^T w_M = 0, \tag{5.14}
\]
\[
w_M^T B_0 \frac{\partial^2 R_{ss}(\theta_s^2)}{\partial \theta_s^2} \bigg|_{\theta_s=0} B_0^T w_M = 0. \tag{5.15}
\]

Eqn. (5.14) is a linear constraint on \( w_M \), but (5.15) is not. We can convert (5.15) into linear constraints based on singular value decomposition (SVD),
\[
B_0 \frac{\partial^2 R_{ss}(\theta_s^2)}{\partial \theta_s^2} \bigg|_{\theta_s=0} B_0^T = U D U^T, \tag{5.16}
\]
where $D$ is a diagonal matrix and $U$ is an orthogonal matrix. The linear constraints derived from (5.15) can then be written as

$$U_s^T w_M = 0,$$

(5.17)

where $U_s$ is the matrix whose columns are those of $U$ that correspond to more significant eigenvalues. The columns in $U$ that correspond to less significant eigenvalues can be omitted without dramatically degrading the performance.

From the above derivation, it can be seen that the first order leakage (5.8) requires only one constraint while the number of constraints for the second order leakage depends on the SVD in (5.16). In general, the total number of leakage constraints based on the first and second order approximation is less than that of the derivative constraints [EC83] (which requires $3L$ constraints).

Representing all selected leakage constraints by matrix $C_0$, the optimization problem becomes

$$\min_{w_M} P = \min_{w_M} (w_F - B_0^T w_M)^T R_{xx} (w_F - B_0^T w_M),$$

subject to $C_0^T w_M = 0$,  

(5.18)

where

$$C_0 = [c_{1,1} \ c_{2,1} \ U_s],$$

$$c_{1,1} = B_0 \left. \frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \right|_{\theta_s=0} w_F,$$

$$c_{2,1} = B_0 \left. \frac{\partial^2 R_{ss}(\theta_s)}{\partial \theta^2_s} \right|_{\theta_s=0} w_F,$$

and $U_s$ is given in (5.17). Denote $N_L$ the total number of leakage constraints.
5.4 Implementations

In this section, some implementation issues of the GSC with the leakage constraints are discussed.

5.4.1 On Estimation of $\frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \bigg|_{\theta_s=0}$ and $\frac{\partial^2 R_{ss}(\theta_s)}{\partial \theta_s^2} \bigg|_{\theta_s=0}$

The correlation matrix of the target signal, $R_{ss}(\theta_s)$, is defined as

$$R_{ss}(\theta_s) = E[s(n)s(n)^T],$$

where the signal vector $s(n)$ is

$$s(n) = [s_1(n) \ldots s_K(n) s_1(n-1) \ldots s_K(n-1)$$
$$\ldots s_1(n-L+1) \ldots s_K(n-L+1)]^T.$$

Since the target signal received by each sensor is just a duplicate of the original desired signal when the sensor gain is normalized, the elements in matrix $R_{ss}(\theta_s)$ are given by

$$r_{ss}(t_2 - t_1) = E[s_{k_1}(n-l_1)s_{k_2}(n-l_2)], \quad (5.19)$$

where $k_1$, $l_1$, $k_2$ and $l_2$ are all integers, $r_{ss}(t)$ is the autocorrelation function of the original continuous desired signal $s(t)$, and $t_1$, $t_2$ are the time delays of $s_{k_1}(n-l_1)$ and $s_{k_2}(n-l_2)$ relative to $s_1(n)$ respectively. They could be expressed
as

\[ t_n = \frac{(k_n - 1)d \sin \theta_s}{c} + l_n T_s, \quad n = 1, 2, \ldots \]

where \( d \) is the inter-element spacing, \( c \) is the wave speed, and \( T_s \) is the sampling interval.

As the output from the fixed beamformer is an enhanced estimate of the target signal, we can estimate \( r_{ss}(t) \) from the fixed beamformer output. When \( t \) is an integer multiple of \( T_s \),

\[ r_{ss}(mT_s) \approx E[w_T^T x(n + m)x(n)^T w_F] \approx \frac{1}{N} \sum_{n=0}^{N-1} w_T^T x(n + m)x(n)^T w_F, \]

where \( m \) is an integer and \( N \) is the length of observation. When \( t \) is not an integer multiple of \( T_s \), the interpolation formula is used to obtain \( r_{ss}(t) \)

\[ r_{ss}(t) = \sum_{m=\infty}^{\infty} \frac{\sin \left[ \pi \left( m - \frac{t}{T_s} \right) \right]}{\pi \left( m - \frac{t}{T_s} \right)} r_{ss}(mT_s). \tag{5.20} \]

Therefore, combining (5.19) and (5.20), the \( p \)-th order derivatives of the elements of \( R_{ss}(\theta_s) \) are given by

\[ \frac{\partial^p r_{ss}(t_2 - t_1)}{\partial \theta_s^p} = \sum_{m=-\infty}^{\infty} \left( \frac{\partial^p \sin \left[ \pi \left( m - \frac{t_2 - t_1}{T_s} \right) \right]}{\pi \left( m - \frac{t_2 - t_1}{T_s} \right)} \right) r_{ss}(mT_s). \]

Here, the infinite sum about \( m \) can be approximated by a finite sum because in general the magnitude of \( r_{ss}(mT_s) \) and the derivatives of \( \frac{\sin \left[ \pi \left( m - \frac{t}{T_s} \right) \right]}{\pi \left( m - \frac{t}{T_s} \right)} \) are negligibly small with the increase of the magnitudes of \( m \) and \( m - \frac{t}{T_s} \) respectively.
5.4.2 Formulation and Implementation in SOCP

As the robust adaptive beamforming problem can be formulated as a second-order cone program (SOCP) and solved via the well-established interior point method. This approach was adopted recently to combat various steering vector errors [VGL03, VGLM04, YM05, LB05]. The problem in (5.18) can also be cast as an SOCP problem in the following.

With Cholesky factorization, $R_{xx} = QQ^T$, the optimization problem in (5.18) is equivalent to

$$\min_{w_M} \delta \quad (5.21a)$$

subject to:

$$\|Q^T(w_F - B_0^T w_M)\| \leq \delta \quad (5.21b)$$

$$C_0^T w_M = 0. \quad (5.21c)$$

If we treat $\delta$ as an additional design variable and define an augmented vector as $y = [\delta \ w_M^T]^T$, the objective function can be expressed as $\delta = f^T y$ with $f = [1 \ 0 \ldots \ 0]^T$. Note that the constraints in (5.21b)-(5.21c) are actually quadratic cones and can be expressed in unit second-order cones under affine mapping. The constrained optimization problem defined by (5.21) can then be formulated into an SOCP framework as

$$\min_y f^T y \quad (5.22a)$$

subject to:

$$\begin{bmatrix} 1 & 0^T \\ 0 & Q^T B_0^T \end{bmatrix} y + \begin{bmatrix} 0 \\ Q^T w_F \end{bmatrix} \in C_1 \quad (5.22b)$$

$$\begin{bmatrix} 0 \\ C_0^T \end{bmatrix} y \in C_2 \quad (5.22c)$$
5.4. Implementations

where $C_1, C_2$ are second-order cones in $R^{KL+1}$ and $R^{NL}$, respectively.

The adaptive beamforming can be implemented using the following steps.

1. When the new array signal $x(n)$ is received, the correlation matrix is updated by

$$R_{xx}^{(n)} = \alpha R_{xx}^{(n-1)} + (1 - \alpha)x(n)x^T(n).$$

where $\alpha$ is the update coefficient that is close to 1.

2. Obtain $Q$ by Cholesky decomposition of $R_{xx}$.

3. Apply the SeDuMi toolbox to solve (5.22) and go to Step 1.

5.4.3 On Adaptive Algorithm

Besides the formulation in SOCP, the adaptive algorithm for (5.18) can be derived in a similar way to [Fro72].

With a zero initial weight vector, i.e., $w_M^{(0)} = 0$, the weight vector at the $(n + 1)$-th iteration is given by

$$w_M^{(n+1)} = w_M^{(n)} + u \left( I - C_0(C_0^TC_0)^{-1}C_0^T \right) B_0x(n) \left( x(n)^T w_F - x(n)^T B_0w_M^{(n)} \right)$$

(5.23)

where $u$ is the stepsize, which is a small positive number.

Regarding the two implementation methods in Sections 5.4.2 and 5.4.3, the simulation results are similar. However, the latter is more efficient. For the former, heavier computation is involved in Cholesky decomposition and nonlinear convex optimization at each time instance, while a linear update is involved.
5.4. Implementations

for the latter. Simulation result using the latter method is presented in the following simulation.

5.4.4 On Norm Constrained Adaptive Filtering

The constraints imposed by \( C_0 \) aim to mitigate the leakage signal as much as possible from the adaptive process so that the adaptation of \( w_M \) is conducted only towards the suppression of interferences and noise. However, since \( C_0 \) is obtained by neglecting the higher order leakages, minor leakages exist which might cause certain signal cancellation. To overcome this problem, a norm constraint can be imposed on the weight vector in the adaptive process as [CZO87]

\[
\| w_M^{(n+1)} \| \leq \varepsilon.
\]

where \( \| \cdot \| \) represents the \( l_2 \)-norm and \( \varepsilon \) is a positive constant, which could be approximately given by [ZYL03]

\[
\varepsilon \approx \left( \frac{2}{\theta_s} \right) G + H
\]

(5.24)

where \( G \) and \( H \) are defined as

\[
G = \left( B_0 \frac{\partial^2 R_{ss}(\theta_s)}{\partial \theta_s^2} \bigg|_{\theta_s=0} \right) B_0^T \left( B_0 \frac{\partial R_{ss}(\theta_s)}{\partial \theta_s} \bigg|_{\theta_s=0} \right) w_F,
\]

(5.25a)

\[
H = \left( B_0 \frac{\partial^2 R_{ss}(\theta_s)}{\partial \theta_s^2} \bigg|_{\theta_s=0} \right) B_0^T \left( B_0 \frac{\partial^2 R_{ss}(\theta_s)}{\partial \theta_s^2} \bigg|_{\theta_s=0} \right) w_F.
\]

(5.25b)

Note that \( \theta_s, w_F, w_M, B_0 \) and \( R_{ss}(\theta_s) \) are given in Section 5.3. For more details, please see [ZYL03].
5.5 Computer Simulations

In this section, the performance of a GSC with the first and second order leakage constraints is evaluated and compared with the conventional GSC and an adaptive beamformer with derivative constraints by computer simulations.

A six-element ($K = 6$) uniform linear microphone array with a sensor spacing of 4 cm is used. The length of tap line for each sensor is $L = 33$. The sampling frequency is 8 kHz. For the proposed GSC, a group of 23 leakage constraints is formed based on the first and second order leakage approximation. An additional norm constraint with $\varepsilon = 2.14$ is also added for the modified GSC. For the adaptive algorithm in (5.23), the stepsize is set at $u = 0.05$. For the adaptive beamformer with derivative constraints, up to the second order derivative constraints are used with 99 linear constraints in total.

Firstly, to demonstrate the capability of the GSC with leakage constraints in broadening the acceptance angle, a broadband signal source (0.1-3.4 kHz) is assumed. The look direction is $0^\circ$. $w_F$ is a simple delay-and-sum beamformer and blocking matrix $B_0$ is the same as the one used in [GJ82]. The signal-to-noise ratio (SNR) is assumed to be 30 dB. The spatial responses of the beamformers in Gaussian noisy environment are plotted in Figure 5.3. It is seen that the noise cancellation capability of the GSC with leakage constraints is slightly reduced compared with the conventional GSC due to the reduction in the degree of freedom in the weight space. But the acceptance angle is broadened considerably using the proposed method. Compared with the adaptive beamformer with derivative constrains, our GSC performs much better in terms of its capability of noise suppression, though the acceptance angle is not as broad as that of the adaptive beamformer with derivative constrains.
Secondly, to evaluate the performance of the proposed method, as compared with that of the conventional GSC and the adaptive beamformer with derivative constraints, in the environment of a strong interference, it is assumed that a desired broadband signal (0.1-3.4 kHz) and an interference with the same frequency bandwidth are received by the modified GSC whose look direction is 0°. The interference is from 20°. The input SNR is 30 dB and the input signal-to-interference ratio (SIR) is 0 dB. The signal DOA is from −40° to 40°. The SIR of the beamformer outputs are measured and presented in Figure 5.4. It can be seen that when the signal is from 0°, the interference suppression capability of the GSC with leakage constraints is about 2 dB weaker than that of the conventional GSC. However, when the signal DOA moves away from 0°, the SIR of the conventional GSC decreases faster than that of the GSC with leakage constraints. In fact, it can be seen from Figure 5.4 that the proposed GSC with leakage constraints outperforms the conventional GSC in the region of [1°, 10°] for the signal DOA. For example, when the signal DOA is 5°, the GSC with leakage constraints outperforms the conventional GSC by 3 dB. It suggests that the signal cancellation due to DOA mismatch is alleviated using the proposed method. The adaptive beamformer with derivative constraints has poorer performance in terms of interference suppression due to its more conservative constraints though it is more robust to DOA mismatch.

Thirdly, the convergence rate of the three adaptive beamformers are compared when a desired broadband signal from 0° and an interference from 20° are received in a Gaussian noisy environment. The input SNR and SIR are -30 dB and 0 dB, respectively and the beamformer output power versus the samples is measured and presented in Figure 5.5. It can be seen that with the same stepsize, the convergence rate of the modified GSC is comparable to that of
5.6 Conclusions

In this chapter, a class of leakage constraints is proposed for GSCs in the presence of DOA mismatch by exploiting the statistical property of the leakage signals. From the simulation results, it can be seen that the proposed method is more robust compared with the conventional GSC as target signal cancellation due to DOA mismatch is alleviated by the imposed leakage constraints. With fewer linear constraints compared with the adaptive beamformers with derivative constraints, the proposed GSC is less conservative as it can broaden the acceptance angle with slight degradation of its noise cancellation and interference suppression capability.

Figure 5.3: Spatial responses of various adaptive beamformers in Gaussian noisy environment.
5.6. Conclusions

Figure 5.4: Signal-to-interference ratio comparison among various adaptive beamformers.

Figure 5.5: Convergence of the proposed algorithm. The line on the top is for the GSC with derivative constraints, while the lines for the conventional GSC and the modified GSC almost overlap.
Chapter 6

Conclusions and
Recommendations for Future Work

6.1 Conclusions

In this thesis, we have presented the research work on two related subjects: finite impulse response (FIR) filter design and array beamforming. The focus is on the design of FIR filters with reduced group delay errors and robust/efficient array beamforming. The achievements include:

1. A new design method for general FIR filters with reduced passband group delay errors is proposed.

2. An improved design method for frequency response masking (FRM) FIR filters with reduced passband group delay errors is proposed.
6.2 Recommendations

3. The feasibility of the application of the FRM technique in array beam-forming shows that the computational complexity of an active array beam-former can be reduced with a novel combination of the FRM technique and the concept of effective aperture.

4. A new class of linear constraints to alleviate the signal cancellation due to DOA mismatch for broadband general sidelobe cancellers (GSCs) is proposed.

6.2 Recommendations

Based on the research work presented in this thesis, some recommendations for future work are given as follows:

1. In Chapter 2, we have presented the design method for FIR filters in SDP framework. The proposed method may be further explored to carry out the FIR filter design in SOCP framework.

2. In Chapter 3, we have presented the design method for real FRM filters. The proposed design method can be extended in principle to complex FRM filters. As more coefficients (due to the imaginary parts) are involved, the expressions for the group delay and its gradients will be more complicated.

3. In subband coding, subband adaptive filtering, etc., there is a growing interest in designing perfect reconstruction filter banks with reduced group delays. Based on the proposed methods in Chapters 2 or 3, a new design
6.2. Recommendations

method may be investigated to meet the specifications in both magnitude response and group delay.

4. The proposed filter design algorithms in this thesis focus on the design methodology and assume that the filters can be implemented with very long word lengths. However, digital filters with finite word lengths are of practical use, typically 8- or 16-bit, and direct truncation of the word lengths of the filter coefficients leads to distortion. Therefore, a new design method for FRM filters with reduced group delay errors and finite word lengths is worth exploring.

5. For robust array beamforming, more practical errors inherent in sensors, such as phase and magnitude errors, and position error, etc., should be considered in the future development.
Author’s Publications

Journal Papers


Conference Papers


**Journal Paper Submitted**

Bibliography


Bibliography


