Filter Bank Design for Subband Adaptive Filtering

Methods and Applications

Jan Mark de Haan

Ronneby, May 2001
Department of Telecommunications and Signal Processing
Blekinge Institute of Technology
S-372 25 Ronneby, Sweden

Licenciate Dissertation Series No. 03/01
ISSN 1650-2140
ISBN 91-7295-003-X

Published 2001
Printed by Kaserntryckeriet AB
Karlskrona 2001
Sweden

Abstract

Adaptive filtering is an important subject in the field of signal processing and has numerous applications in fields such as speech processing and communications. Examples in speech processing include speech enhancement, echo- and interference-cancellation, and speech coding.

Subband filter banks have been introduced in the area of adaptive filtering in order to improve the performance of time domain adaptive filters. The main improvements are faster convergence speed and the reduction of computational complexity due to shorter adaptive filters in the filter bank subbands.

Subband filter banks, however, often introduce signal degradations. Some of these degradations are inherent in the structure and some are inflicted by filter bank parameters, such as analysis and synthesis filter coefficients. Filter banks need to be designed so that the application performance degradation is minimized. The presented design methods in this thesis aim to address two major filter bank properties, transmission delay in the subband decomposition and reconstruction as well as the total processing delay of the whole system, and distortion caused by decimation and interpolation operations. These distortions appear in the subband signals and in the reconstructed output signal.

The thesis deals with different methods for filter bank design, evaluated on speech signal processing applications with filtering in subbands.

Design methods are developed for uniform modulated filter banks used in adaptive filtering applications. The proposed methods are compared with conventional methods. The performances of different filter bank designs in different speech processing applications are compared. These applications are acoustic echo cancellation, speech enhancement including spectral estimation, subband beamforming, and subband system identification. Real speech signals are used in the simulations and results show that filter bank design is of major importance.

Preface

This Licenciate thesis summarizes my work in the field of filter bank design for subband adaptive filtering applications. The work has been carried out at the Department of Telecommunications and Signal Processing at Blekinge Institute of Technology in collaboration with Ericsson Mobile Communications.

The thesis consists of three parts:

Part	
I	Filter Bank Design for Subband Adaptive Filtering Applications
II	Design of Oversampled Uniform DFT Filter Banks with Delay Specification using Quadratic Optimization
III	Design of Oversampled Uniform DFT Filter Banks with Reduced Inband Aliasing and Delay Constraints

Acknowledgments

I wish to express my outmost gratitude to Professor Ingvar Claesson for supervision during my studies as a Ph. D. candidate. Special thanks go to both Professor Ingvar Claesson and Professor Sven Nordholm for offering me the possibility and the challenge of being a Ph. D. student. I wish to express my gratitude for all the support and encouragement they have given in my every day work.

I am indebted to Tech. Lic. Nedelko Grbić for all inspiring, interesting and constructive discussions in the course of this work. It has been a great pleasure to work together with him.

I also would like to thank my other colleagues, both present and former, at the Department of Telecommunications and Signal Processing for the professional and the non-professional encounters, at work and on the sports ground. I would especially like to thank M. Sc. Timothy Samuels and Dr. Abbas Mohammed proofreading the manuscript.

I am grateful to all my friends for being good friends, and to my family. I especially thank my father, Jan Thijs, and my mother, Heleen, for their encouragement and support during my life.

Finally, I thank my fiancé Lisa for her everlasting love and support.

Jan Mark de Haan
Ronneby, May 2001

Contents

Publication list

Part I is published as:

J. M. de Haan, N. Grbić and I. Claesson, "Filter Bank Design for Subband Adaptive Filtering Applications," Research Report, ISSN 1103-1581, May 2001.

Parts of this research report have been submitted for publication as:

J. M. de Haan, N. Grbić and I. Claesson, "Filter Bank Design for Subband Adaptive Filtering Applications," submitted to IEEE transactions on Speech and Audio Processing, May 2001.

Part II is published as:

J. M. de Haan, N. Grbić, I. Claesson and S. Nordholm, "Design of Oversampled Uniform DFT Filter Banks with Delay Specification using Quadratic Optimization," accepted for publication at the International Conference on Acoustics, Speech, and Signal Processing, ICASSP, Salt Lake City, USA, May 2001.

Part III is submitted for publication as:

N. Grbić, J. M. de Haan, I. Claesson and S. Nordholm "Design of Oversampled Uniform DFT Filter Banks with Reduced Inband Aliasing and Delay Constraints," submitted for presentation at the Sixth International Symposium on Signal Processing and its Applications, ISSPA, Kuala-Lumpur, Malaysia, Aug. 2001.

Introduction

Adaptive Filtering is an important concept in the field of signal processing and has numerous applications in fields such as speech processing and communications. Examples in speech processing include speech enhancement, echo- and interference-cancellation and speech coding. Filter banks have been introduced in order to improve the performance of time domain adaptive filters. The main improvements are faster convergence and the reduction of computational complexity due to shorter adaptive filters in the subbands, operating at a reduced sample rate.

Filter banks, however, introduce signal degradations. Some degradations are related to the structure; others due to the design of the filter banks. Filter banks come in many structures and offer many different decompositions; for example the tree structure and parallel structure can both provide a uniform decomposition of the signal spectrum. Filter banks need to be designed so that the performance degradation, due to filter bank properties, is minimized. Here, the word *performance* refers to the application in which the filter bank is used. The design methods which are presented in this thesis aim to address the two main filter bank features:

- the transmission delay in the subband decomposition and reconstruction operations,
- the distortion caused by the decimation and interpolation operations, aliasing and imaging distortions, respectively, which appear in the subband signals and the reconstructed output signal.

The thesis comprises different methods which are evaluated on speech signal processing applications with filtering operations in subband signals.

PART I - Filter Bank Design for Subband Adaptive Filtering Applications

A design method is developed for uniform modulated filter banks for use in adaptive filtering applications. Uniform modulated filter banks are filter banks with a parallel structure and subband signals with bandwidth uniformly distributed across the full frequency band. Single low pass prototype filters describe the analysis and synthesis filter banks. Therefore the design problem is reduced into a two-step design of single low pass filters.

Pre-specified parameters of the filter bank design method are the number of subbands, the decimation factor and the lengths of the analysis and synthesis filters. The transmission delays are also pre-specified parameters. In the design methods, the aliasing distortions are minimized as well as the amplitude and phase distortions. The two design steps involve quadratic form expressions, which enable the use of simple quadratic optimization techniques.

The filter banks obtained from the proposed method are compared with conventional filter banks. Eight filter bank cases are studied extensively, of which four are designed with the proposed method. In the first evaluation section, properties of the filter banks themselves are compared. In the next sections, the performance of the filter bank in adaptive filtering applications is compared. The first evaluation is done in subband acoustic echo cancellation. The second evaluation is done in the applications of subband speech enhancement, and the third evaluation is done in techniques of spectral estimation. A separate section is devoted to an efficient implementation of uniform modulated filter banks and issues of computational complexity.

Part II - Design of Oversampled Uniform DFT Filter Banks with Delay Specification using Quadratic Optimization

The two-step design method for uniform modulated filter banks, using quadratic optimization, is described.

The filter banks, obtained with the proposed design method, are applied in subband beamforming. Beamforming is a technique which is used in speech signal processing for extraction of a single speaker in background noise. Broadband beamforming techniques exploit the spatial as well as the temporal distribution of the input signal.

In the evaluation, beamforming is applied in a handsfree situation in an automobile. The handsfree situation comprises background noise caused for instance by wind and tire friction, and interference caused by the far-end speech entering the return path. Real speech signals are used to evaluate the performance of the beamformer, such as noise suppression, interference suppression and speech distortion.

Part III - Design of Oversampled Uniform DFT Filter Banks with Reduced Inband Aliasing and Delay Constraints

In Part III, a simple design method for uniform modulated filter banks is proposed. The core of the design method is filter design using complex approximation with least squares error. A single low pass filter is used in both the analysis filter bank and the synthesis filter bank, of which the group delay may be specified. Design parameters are optimized using a line search method with discretized gradient approximations, in order to minimize the magnitude- and phase- errors and the

inband- and residual- aliasing. The filter banks obtained by the proposed method are evaluated in optimal subband system identification. A real conference room impulse response is the unknown system in the evaluation.

Part I

Filter Bank Design for Subband Adaptive Filtering Applications

Part I is published as:

J. M. de Haan, N. Grbić and I. Claesson, "Filter Bank Design for Subband Adaptive Filtering Applications," Research Report, ISSN 1103-1581, May 2001.

Parts of this research report have been submitted for publication as:

J. M. de Haan, N. Grbić and I. Claesson, "Filter Bank Design for Subband Adaptive Filtering Applications," submitted to IEEE transactions on Speech and Audio Processing, May 2001.

Filter Bank Design for Subband Adaptive Filtering Applications

J. M. de Haan, N. Grbić, I. Claesson

Department of Telecommunications and Signal Processing
Blekinge Institute of Technology
Ronneby, Sweden

Abstract

Subband adaptive filtering has been proposed to improve the convergence rate while reducing the computational complexity associated with time domain adaptive filters. Subband processing introduces transmission delays caused by the filters in the filter bank and signal degradations due to aliasing effects. One way to reduce the aliasing effects is to impose oversampling in subbands rather than critical sampling and thus reduce signal degradation. By doing so, additional degrees of freedom are introduced for the design of filter banks which may be optimally exploited. In part I, a design method for modulated uniform filter banks with any oversampling is suggested, where the analysis filter bank delay and the total filter bank delay may be specified, and where the aliasing and magnitude/phase distortions are minimized.

Uniformly modulated filter banks are designed and performance evaluations are coducted for the applications of subband system identification and subband speech enhancement. Filter banks from the suggested method are compared with conventional filter banks using real speech signals.

1 Introduction

Filter banks have been of great interest in a number of signal processing applications. A large group of these applications comprise those utilizing subband adaptive filtering. Examples of applications where subband adaptive filtering succesfully is applied are acoustic echo cancellation [1, 2, 3, 4, 5], speech enhancement [6], signal separation [7], and beamforming [8].

To decrease the complexity of filter bank structures, the sampling rate can be reduced in the subbands. These filter banks are refered to as decimated filter banks, and are afflicted with three major types of distortions: amplitude-, phase- and

aliasing- distortion. These distortions degrade the performance of the application in which the filter bank is used.

Several design methods have been proposed and evaluated on subband adaptive filtering applications. Prototype filters for modulated filter banks are designed by interpolation of a two-channel Quadrature Mirror Filter (QMF), and evaluated in a real-time acoustic echo cancellation application in [9]. A design method is proposed using an iterative least-squares algorithm where a reconstruction error and the stopband energy are simultaneously minimized in [10]. Other methods have been proposed for the design of perfect-reconstruction or paraunitary filter banks, which is a class of perfect-reconstruction filter banks. A method using non-constrained optimization is described in [11], for the design of perfect reconstruction polyphase filter banks with arbitrary delay, aimed at applications in audio coding. Kliewer proposed a method for linear-phase prototype FIR filters with power complementary constraints for cosine modulated filter banks, based on an improved frequency-sampling design [12]. In [13], the design of perfect reconstruction cosine-modulated paraunitary filter banks is discussed. Heller presented a design method for prototype filters for perfect reconstruction cosine-modulated filter banks with arbitrary prototype filter length and delay in [14].

Uniformly modulated filter banks have been of special interest because of their simplicity and efficient implementation [15, 16]. In part I a design method for analysis and synthesis uniformly modulated filter banks is presented, using unconstrained quadratic optimization. The goal of the design method is to minimize magnitude-, phase-, and aliasing- distortion in the reconstructed output signal, caused by the filter bank, as well as to minimize aliasing in the subband signals. Aliasing affects the performance of subband adaptive filters. The goal is to design filter banks, taking into consideration that adaptive filtering in the subbands should cause minimal degradation.

The proposed filter bank design method uses pre-specified design parameters which control the filter bank properties, such as:

- Number of subbands
- Decimation factor
- Analysis and synthesis filter lengths
- Analysis filter bank delay
- Total delay of the analysis-synthesis structure

Design issues of M-channel filter banks are discussed in Section 2. The uniformly modulated filter bank is presented in Section 3. In Sections 4 and 5 the proposed filter bank design method is described. In Section 6, the influence of the design parameters on the performance of the filter bank itself is conducted.

Section 7 deals with the implementation and the computational complexity of 2-times oversampled uniformly modulated filter banks compared to critically sampled uniformly modulated filter banks. In Section 8, illustrative filter bank examples are presented and the performance of these filter banks is compared with conventional filter banks. In Section 9, 10, and 11, evaluations are conducted in acoustic echo cancellation, subband speech enhancement and spectral estimation. Finally, Section 12 concludes part I.

2 Design of M-channel Filter Banks

2.1 Decimated Filterbanks with Parallel Structure

An M-channel analysis filter bank is a structure transforming an input signal to a set of M subband signals. A corresponding M-channel synthesis filter bank transforms M subband channels to a full band signal, in orther words the inverse operation. Different kinds of structures of filter banks exist, for example the *tree structure* and the *parallel structure* [17]. The parallel structure is more general and favorable because of its simplicity. In Fig. 1, the analysis and synthesis filter

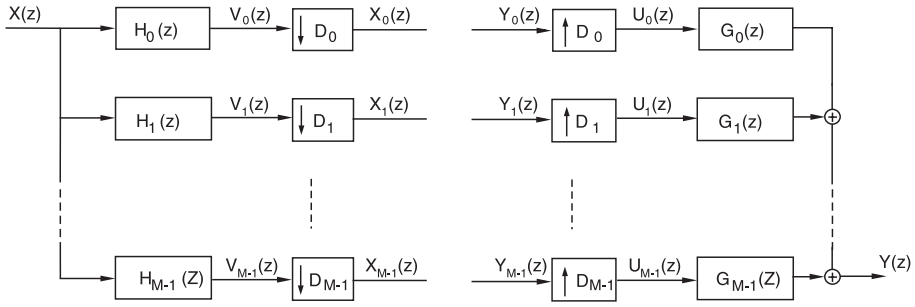

Figure 1: *Parallel structure of general M-channel analysis and synthesis filter banks.*

banks with parallel stucture are illustrated. The analysis filter bank consists of M analysis filters, $H_m(z)$, with non-overlapping bandwidth. The first analysis filter, $H_0(z)$, is usually of lowpass type and the other filters are of bandpass type. All filters may generally have arbitrary bandwidth. A special case is the *uniform* filter bank, where all filters have the same bandwidth.

The input signal, $x(n)$, is filtered by the analysis filters to form the signals $v_m(n)$, $m = 0, \ldots, M-1$,

$$v_m(n) = h_m(n) * x(n) \quad \longleftrightarrow \quad V_m(z) = H_m(z)X(z). \tag{1}$$

The reduced bandwith of $v_m(n)$ allows for a reduction of the sample rate. Reducing the sample rate by decimation gives rise to efficient implementation of the filter

bank. The decimation factor, D_m, generally depends on the bandwidth of the corresponding signal $v_m(n)$,

$$x_m(l) = v(lD_m) \quad \longleftrightarrow \quad X_m(z) = \frac{1}{D_m} \sum_{d=0}^{D_m-1} V_m(z^{\frac{1}{D_m}} e^{-j2\pi d/D_m}), \tag{2}$$

where a new discrete time variable l is introduced for the signals with lower sample rate. Note that $x_m(l)$ consists of D_m terms, which will be referred to as *aliasing terms*. In subband adaptive filtering, filters are applied on the subband signals $x_m(l)$,

$$y_m(l) = \xi_m(l) * x_m(l) \quad \longleftrightarrow \quad Y_m(z) = \xi_m(z) X_m(z). \tag{3}$$

The output signals, $y_m(l)$ of the subband filters $\xi_m(l)$, are then transformed to the full-band output signal $y(n)$ by a synthesis filter bank. The synthesis filter bank consists of interpolators and synthesis filters $G_m(z)$. The input subband signals, $y_m(l)$, are interpolated,

$$u_m(n) = \begin{cases} y_m(n/D_m), & n = 0, \pm D_m, \pm 2D_m, \dots \\ 0, & \text{otherwise} \end{cases} \quad \longleftrightarrow \quad U_m(z) = Y_m(z^{D_m}), \tag{4}$$

and then filtered by the synthesis filters. Finally, the signals are summed together to form the output signal, $y(n)$,

$$y(n) = \sum_{m=0}^{M-1} g_m(n) * u_m(n) \quad \longleftrightarrow \quad Y(z) = \sum_{m=0}^{M-1} G_m(z) U_m(z). \tag{5}$$

2.2 Design Issues of Decimated Filter Banks

Analysis-synthesis filter bank structures are inflicted with different types of distortions. These distortions are, among other important filter bank properties, illustrated in Fig. 2. Without any filtering in the subbands, i.e. $\xi_m(z) = 1$, the desired response of the total analysis-synthesis structure has unit amplitude and linear phase. Henceforth this response will be refered to as the *total response*. Deviations from these requirements are refered to as amplitude distortion and phase distortion.

The decimation and interpolation operations cause aliasing and imaging distortion. The stopband of the analysis filters attenuates the spectral portions of the input signal, which appear in the subband signals, $x_m(l)$, as undesired aliasing components. The stopband of the synthesis filters attenuates the multiple images of the subband spectrum, which appear in the signals, $u_m(n)$, after interpolation. Depending on how the analysis and synthesis filters are designed, and on the subband filtering, *residual aliasing* might be present in the output signal, $y(n)$.

Filter design procedures can be regarded as optimization procedures with error functions, which are minimized in a certain way. One of the ways in which error

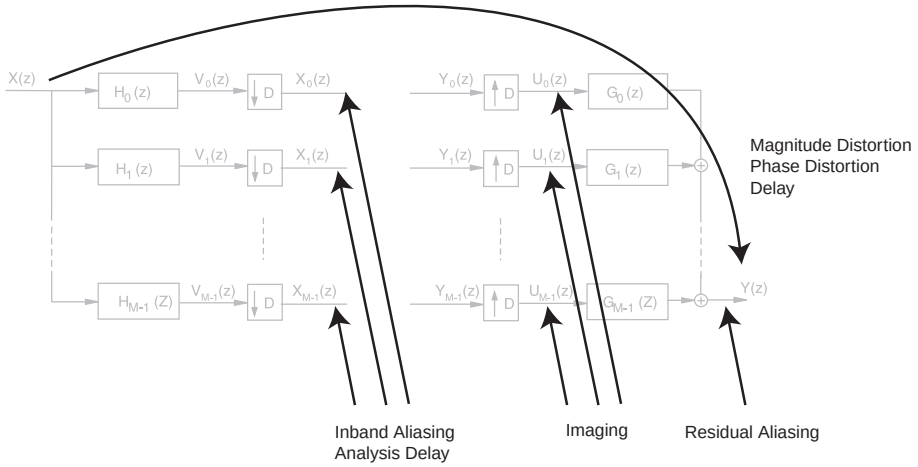

Figure 2: *Important properties in filter bank design.*

functions may be minimized, is by using the least squares criterion. In the design of M-channel filter banks, the least squares criterion can be stated as

$$\min_{\mathbf{h}_m, \mathbf{g}_m} \int_{-\pi}^{\pi} |E_{\mathbf{h}_m, \mathbf{g}_m}(\omega)|^2 d\omega, \tag{6}$$

where $E_{\mathbf{h}_m, \mathbf{g}_m}(z)$ is the error function which depends on the impulse responses of the analysis and synthesis filters $\mathbf{h}_m$ and $\mathbf{g}_m$, and contains the desired properties of the filter bank. The least squares error in Eq. (6) needs generally to be solved using non-linear optimization procedures. However, when the analysis and synthesis filters are defined from single modulated prototype filters, in uniformly modulated filter banks, and when the analysis filters are designed *prior* to the synthesis filters, the design problem can be divided into two sequential quadratic optimization problems. This design procedure will be described in detail in sections 4 and 5.

3 Uniformly Modulated Filter Banks

3.1 Definitions

An M-channel modulated filter bank consists of a set of M branches. Each branch consists of an analysis filter $H_m(z)$, a decimator with decimation factor D, an interpolator with interpolation factor D, and a synthesis filter $G_m(z)$. An illustration of such a filter banks is given in Fig. 3, where subband filtering is applied in the subband signals. All subbands have equal bandwidth and the same decimation and interpolation factors. The analysis and synthesis prototype filters are FIR filters of length $L_{\mathbf{h}}$ and $L_{\mathbf{g}}$, respectively. In order to construct a uniform filter bank, which means that all subbands have the same bandwidth, low pass analysis and synthesis prototype filters, $H(z)$ and $G(z)$, are defined. All analysis filters in the filter bank

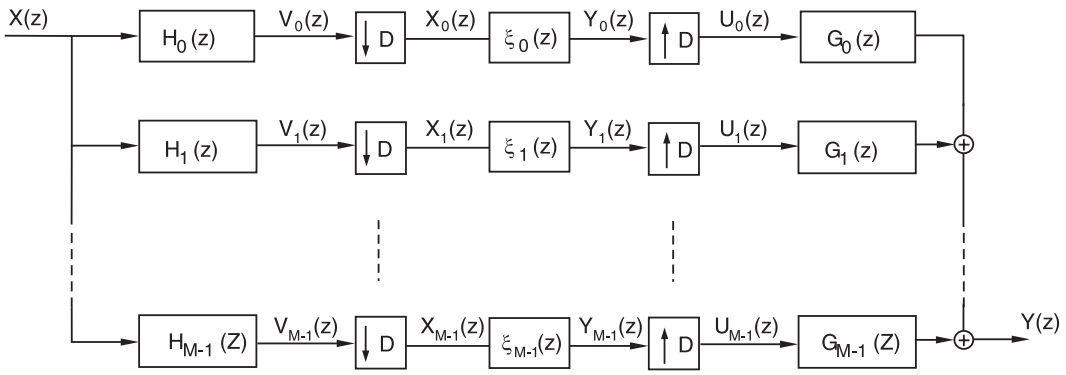

Figure 3: *Analysis and Synthesis Filter Banks with Subband Filtering.*

are modulated versions of the prototype analysis filter according to

$$h_m(n) = h(n)W_M^{-mn} = h(n)e^{j2\pi mn/M} \quad \longleftrightarrow \quad H_m(z) = H(zW_M^m). \tag{7}$$

Similarly, all synthesis filters are modulated versions of the prototype synthesis filter according to

$$g_m(n) = g(n)W_M^{-mn} = g(n)e^{j2\pi mn/M} \quad \longleftrightarrow \quad G_m(z) = G(zW_M^m), \tag{8}$$

where $W_M = e^{-j2\pi/M}$. Observe that the analysis and synthesis filters in the first branch ($m = 0$) are the same as the prototype filters, $H_0(z) = H(z)$ and $G_0(z) = G(z)$. An example of a modulated prototype filter is given in Fig. 4.

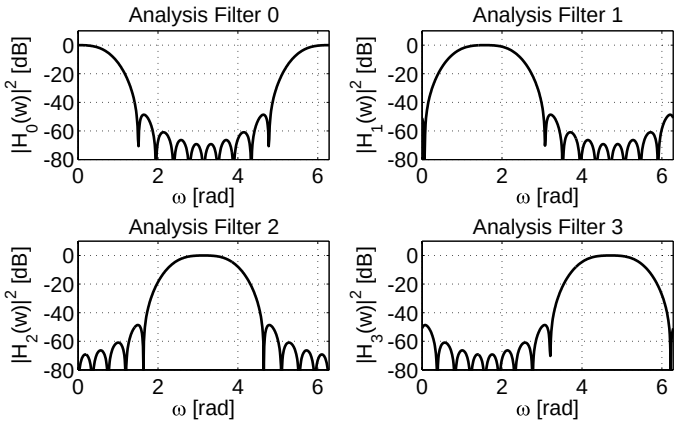

Figure 4: *Analysis filters $H_0(z), \ldots, H_3(z)$ for a modulated filter bank with $M = 4$ subbands.*

3.2 Filter Bank Response

In order to analyze filter bank properties, the input-output relation is derived. Each branch signal, $V_m(z)$, will simply be a filtered version of the input signal as

$$V_m(z) = H_m(z)X(z) = H(zW_M^m)X(z). \tag{9}$$

The decimators will expand the spectra of the branch signals according to

$$X_m(z) = \frac{1}{D}\sum_{d=0}^{D-1} V_m(z^{\frac{1}{D}}W_D^d) = \frac{1}{D}\sum_{d=0}^{D-1} H(z^{\frac{1}{D}}W_M^m W_D^d)X(z^{\frac{1}{D}}W_D^d), \tag{10}$$

where $W_D = e^{-j2\pi/D}$. The summation in Eq. (10) shows that the subband signals consist of D aliasing terms. Depending on the subband index and the decimation factor, the desired spectral content is present in one or more aliasing terms. This is shown in Fig. 5 for a critically sampled case and in Fig. 6 for an oversampled case.

In many filter bank applications, the subband signals $X_m(z)$ are processed (filtered). The filters in the subbands are denoted by $\xi_m(z)$. The processed subband signals, $Y_m(z)$, are filtered versions of the input subband signals, $X_m(z)$, according to

$$Y_m(z) = \xi_m(z)X_m(z) \tag{11}$$

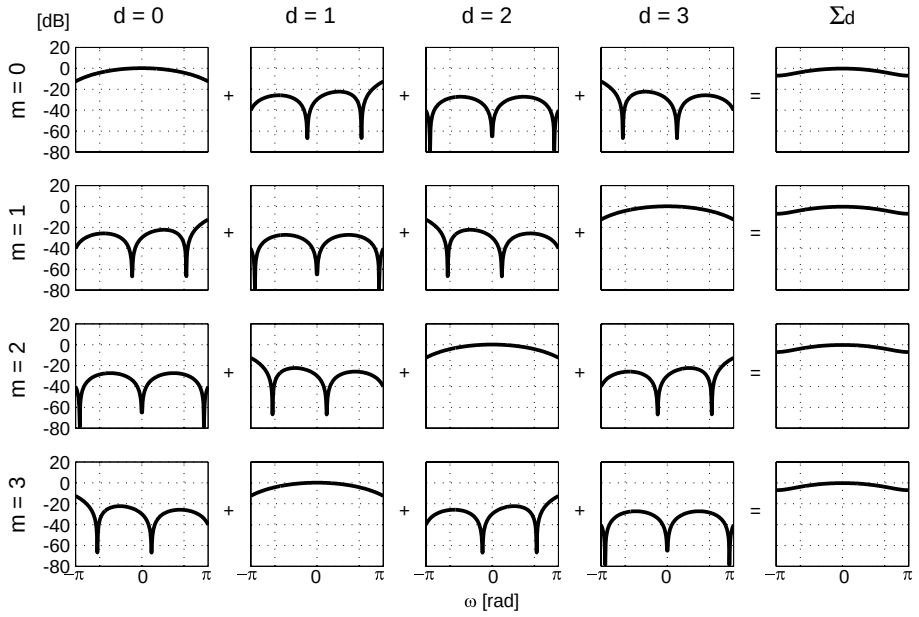

Figure 5: *The influence of $H_m(z)$ on the inband-aliasing terms for a case with $M = D = 4$. The plots represent $|H(e^{j\omega/D}W_M^m W_D^d)|^2$ in dB for different m and d. The sum over d of the terms is plotted in the right column.*

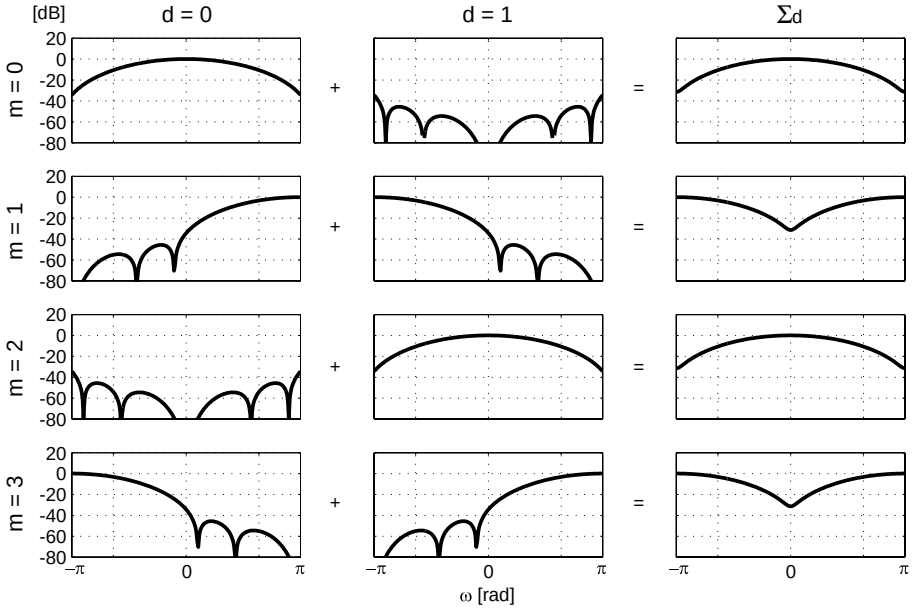

Figure 6: *The influence of $H_m(z)$ on the inband-aliasing terms for a case with $M = 4$ and $D = 2$. The plots represent $|H(e^{j\omega/D}W_M^m W_D^d)|^2$ in dB for different m and d. The sum over d of the terms is plotted in the right column.*

The interpolators compress the spectra of the processed subband signals according to

$$U_m(z) = Y_m(z^D) = \frac{1}{D}\xi_m(z^D)\sum_{d=0}^{D-1} H(zW_M^m W_D^d)X(zW_D^d), \tag{12}$$

After interpolation, multiple images of the signal spectrum will be present in $U_m(z)$, due to the repetitive character of the discrete signal spectrum and the compressing effect of the interpolator. The undesired images are attenuated by the recontruction filters, $G_m(z)$, whereafter the signals are added to form the output signal, $Y(z)$.

$$Y(z) = \sum_{m=0}^{M-1} U_m(z)G_m(z). \tag{13}$$

Inserting Eq. (12) into Eq. (13), the relation between the input signal, $X(z)$, and the output signal, $Y(z)$, becomes

$$Y(z) = \frac{1}{D}\sum_{d=0}^{D-1} X(zW_D^d)\sum_{m=0}^{M-1} \xi_m(z^D)H(zW_M^m W_D^d)G(zW_M^m), \tag{14}$$

which is expressed in terms of the input signal, $X(z)$, the prototype filters, $H(z)$ and $G(z)$, and the subband filters, $\xi_m(z)$. For simplicity, the product of filters in Eq. (14) is defined as $A_{m,d}(z)$, according to

$$A_{m,d}(z) = \frac{1}{D}\xi_m(z^D)H(zW_M^m W_D^d)G(zW_M^m). \tag{15}$$

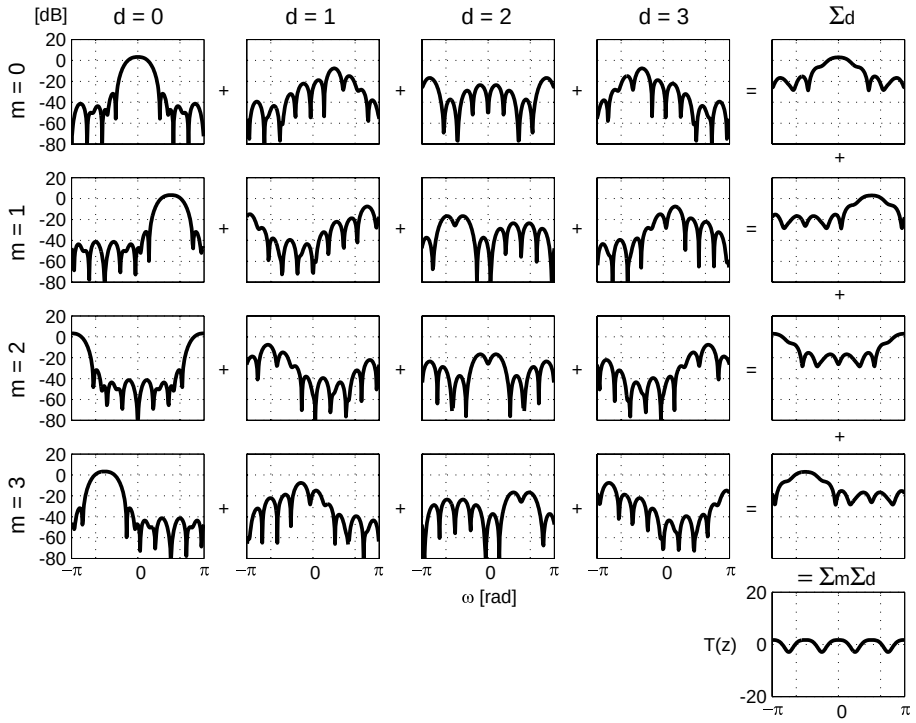

Figure 7: *The influence of $A_{m,d}(z)$ on the aliasing terms in $Y(z)$ for a filter bank with $M = D = 4$. The plots represent $|A_{m,d}(e^{j\omega})|^2$ in dB for different m and d. The sum over d of the terms is plotted in the right column and the sum of all terms, $T(z)$, is given below.*

The influence of $A_{m,d}(z)$, with $\xi_m(z) = 1$, on the aliasing terms is shown in Fig. 7 for a critically sampled filter bank and in Fig. 8 for an oversampled filter bank. The figures show how a number modulated versions of the input signal are filtered and added together to form the output. Clearly the non-modulated input signal contributes with the desired spectral content while the terms for $d > 0$ are undesired aliasing terms which arise from the modulated versions of the input signal.

The output signal in Eq. (14) can be rewritten in the more convenient form

$$Y(z) = \sum_{d=0}^{D-1} A_d(z) X(z W_D^d) \tag{16}$$

where

$$A_d(z) = \sum_{m=0}^{M-1} A_{m,d}(z), \qquad d = 0, \cdots, D-1. \tag{17}$$

The transfer functions, $A_d(z)$ for $d = 1, \ldots, D-1$, can be viewed as the transfer functions which give rise to the residual aliasing terms in the ouput signal. The function, $A_0(z)$, is the transfer function which gives rise to the desired output signal spectrum.

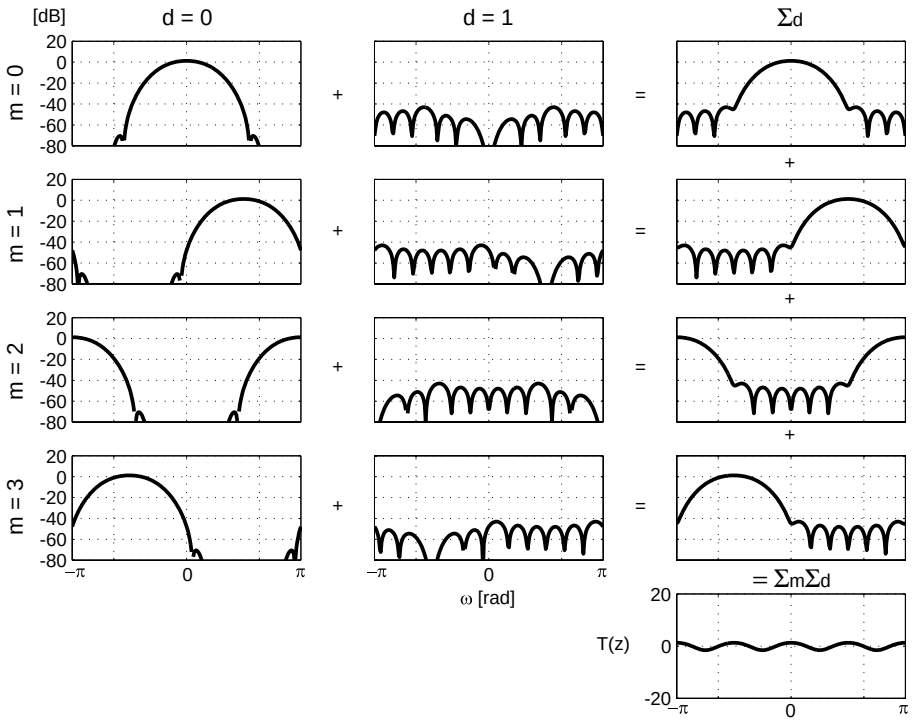

Figure 8: *The influence of $A_{m,d}(z)$ on the aliasing terms in $Y(z)$ for a filter bank with $M = 4$ and $D = \frac{M}{2} = 2$. The plots represent $|A_{m,d}(e^{j\omega})|^2$ in dB for different m and d. The sum over d of the terms is plotted in the right column and the sum of all terms, $T(z)$, is given below.*

The total response function, $T(z)$, for the filter-bank is given by the sum of the desired and undesired aliasing transfer functions, $A_d(z)$

$$T(z) = \sum_{d=0}^{D-1} A_d(z). \tag{18}$$

4 Analysis Filter Bank Design

4.1 Design Strategy

The analysis filter bank design problem reduces to the design of a single prototype analysis filter, $H(z)$, when the analysis filters are modulated versions of the prototype analysis filter according to Eq. (7). The purpose of the analysis filters is to split the original signal into a set of subband signals. Ideal analysis filters are bandpass filters with normalized center frequencies $\omega_m = 2\pi\frac{m}{M}$, $m = 0, \ldots, M-1$, and with bandwidth $\frac{2\pi}{M}$. The ideal filters have unit magnitude and zero phase in the passband while the stopband magnitude is zero. While zero phase filters require non-causality, the requirements need to be relaxed by using linear phase filters. FIR filters may have exact linear phase but they cannot possess the ideal magnitude requirements. Therefore approximations need to be made.

A straightforward way to design the prototype analysis filter is to design a lowpass filter, with a passband region centered around $\omega = 0$, and a minimum magnitude stopband region using filter design methods such as window techniques or the Parks-McClellan optimal equiripple FIR filter design method [18]. The bandwidth is controlled by the passband region and the attenuation of signal components in the stopband region leads to low-energy inband-aliasing terms. These methods cannot control the delay properties of the resulting filter. However, filter design techniques with complex approximation for filters with arbitrary phase exist [19].

The analysis filter bank may be designed so that the transfer functions of the analysis filters have power-complementary transfer functions, i.e. the sum of the squared filter magnitudes is unity [16],

$$\sum_{m=0}^{M-1} |H_m(e^{j\omega})|^2 = 1, \qquad \omega = [-\pi, \pi]. \tag{19}$$

If a prototype analysis filter is designed with this constraint in a modulated structure, the obtained bandwidth will be the maximum bandwidth in order to keep constant magnitude in the analysis filter bank. If the application does not demand the analysis filter bank to possess the power-complementary property, i.e. when a smaller bandwidth of the prototype analysis filters is allowed, a more general approach would be to define an arbitrary boundary frequency of the passband region.

An appropriate design criterion may be to minimize the following objective function

$$\epsilon_{\mathbf{h}} = \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega \tag{20}$$

where $H_d(z)$ is a desired frequency response of the prototype analysis filter in the passband region $\Omega_p = [-\omega_p, \omega_p]$. The desired frequency response is defined as

$$H_d(e^{j\omega}) = e^{-j\omega\tau_H}, \qquad \omega \in \Omega_p \tag{21}$$

where τ_H is the desired group delay of the prototype analysis filter and, consequently, the desired delay of the whole analysis filter bank.

By only specifying the power complementary constraint, or more general, a complex specification for the passband region, a suitable prototype filter cannot be obtained. The stopband regions of the prototype analysis filters must also be defined otherwise significant inband-aliasing distortion may occur in the subbands.

One approach to combat the undesired inband-aliasing is to minimize the energy in the stopband which in turn leads to minimal energy in the aliasing terms. A more appealing approach would be to address the inband-aliasing directly in the objective function. This can be done by complementing the design criterion with the minimization of *inband-aliasing distortion*. In Section 3, the $M \times D$ aliasing terms in all subband signals were identified. The design criterion is complemented by adding the energies in all $M \times (D-1)$ aliasing terms into the objective function in Eq. (20). For a critically sampled filter bank, the objective function becomes

$$\epsilon_{\mathbf{h}} = \alpha_{\mathbf{h}} + \beta_{\mathbf{h}}, \tag{22}$$

with the *Passband Response Error*

$$\alpha_{\mathbf{h}} = \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega \tag{23}$$

and the *Inband-Aliasing Distortion*

$$\beta_{\mathbf{h}} = \frac{1}{2\pi D^2} \sum_{m=0}^{M-1} \sum_{\substack{d=(M-m) \\ \text{mod } M}} \int_{-\pi}^{\pi} \left| H(e^{j\omega/D} W_M^m W_D^d) \right|^2 d\omega, \tag{24}$$

where all inband-aliasing terms are included. Taking a closer look at the summation of aliasing terms, for the critically sampled case, M equal terms can be identified, which are included M times in the summation. So, due to the modulated structure, it is sufficient to include only the terms in the first subband ($m = 0$), i.e. the terms for $d = 1, \ldots, D-1$. Therefore, $\beta_{\mathbf{h}}$ in Eq. (22) can be reduced to

$$\beta_{\mathbf{h}} = \frac{1}{2\pi D^2} \int_{-\pi}^{\pi} \sum_{d=1}^{D-1} \left| H(e^{j\omega/D} W_D^d) \right|^2. \tag{25}$$

Aliasing terms for a critically sampled filter bank case with $M = D = 4$ were shown in Fig. 5. When oversampling is applied, it is not trivial to distinguish the desired or undesired parts in the aliasing terms since the desired spectral content may appear in more than only one aliasing term. This is shown for the two-times oversampled case $M = 4$, $D = \frac{M}{2} = 2$ in Fig. 6. It can be seen from this example that the desired spectral content is spread over a maximum of two aliasing terms. Fig. 6 shows that inclusion of the aliasing terms only for the first subband is a

plausible approach, also in the two-times oversampled case, since the undesired spectral portions are suppressed. Generally, the approach of minimizing inband aliasing with Eq. (25) holds for oversampling by any decimation factor. Henceforth, only the critically sampled and the two times oversampled $D = \frac{M}{2}$) filter banks are considered. In the next sections, the objective function of the analysis filter-bank design is derived in terms of the prototype analysis filter, $H(z)$.

4.2 Passband Response Error

In this section, the quadratic form of the passband response error in Eq. (23) is derived, expressed in terms of the impulse response of the analysis prototype filter. The passband response error can be rewritten as

$$\begin{aligned}
\alpha_{\mathbf{h}} &= \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega \\
&= \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \left(H(e^{j\omega}) - H_d(e^{j\omega})\right)^* \left(H(e^{j\omega}) - H_d(e^{j\omega})\right) d\omega \\
&= \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \left[|H(e^{j\omega})|^2 - 2\mathrm{Re}\left\{H_d^*(e^{j\omega})H(e^{j\omega})\right\} + |H_d(e^{j\omega})|^2\right] d\omega. \quad (26)
\end{aligned}$$

where $H_d(z)$ is the desired frequency response, defined in Eq. (21). The prototype analysis filter response $H(z)$ is expressed in terms of its impulse response, $h(n)$, according to

$$H(z) = \sum_{n=0}^{L_{\mathbf{h}}-1} h(n) z^{-n} = \mathbf{h}^T \phi_{\mathbf{h}}(z) \quad (27)$$

where $\mathbf{h} = [h(0), \ldots, h(L_{\mathbf{h}} - 1)]^T$ and $\phi_{\mathbf{h}}(z) = [1, z^{-1}, \ldots, z^{-L_{\mathbf{h}}+1}]^T$. Substituting Eqs. (27) and (21) into Eq. (26) yields

$$\alpha_{\mathbf{h}} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \left[\mathbf{h}^T \phi_{\mathbf{h}}(e^{j\omega})\phi_{\mathbf{h}}^H(e^{j\omega})\mathbf{h} - 2\mathrm{Re}\left\{e^{j\omega\tau_H}\mathbf{h}^T\phi_{\mathbf{h}}(e^{j\omega})\right\} + 1\right] d\omega. \quad (28)$$

The passband response error, Eq. (28), can be rewritten in the quadratic form

$$\alpha_{\mathbf{h}} = \mathbf{h}^T\mathbf{A}\mathbf{h} - 2\mathbf{h}^T\mathbf{b} + 1 \quad (29)$$

where the $L_{\mathbf{h}} \times L_{\mathbf{h}}$ matrix $\mathbf{A}$ is

$$\mathbf{A} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \phi_{\mathbf{h}}(e^{j\omega})\phi_{\mathbf{h}}^H(e^{j\omega}) d\omega \quad (30)$$

and the $L_{\mathbf{h}} \times 1$ vector $\mathbf{b}$ is

$$\mathbf{b} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \mathrm{Re}\left\{e^{j\omega\tau_H}\phi_{\mathbf{h}}(e^{j\omega})\right\} d\omega. \quad (31)$$

Calculating the integrals for all entries in matrix $\mathbf{A}$ and vector $\mathbf{b}$, the following expression for the matrix entries $A_{m,n}$ is obtained

$$A_{m,n} = \frac{\sin(\omega_p(n-m))}{\omega_p(n-m)}, \tag{32}$$

and the vector entries b_m

$$b_m = \frac{\sin(\omega_p(\tau_H-m))}{\omega_p(\tau_H-m)}. \tag{33}$$

where $m = 0, \ldots, M-1$ and $n = 0, \ldots, L_{\mathbf{h}}-1$. The expressions for $A_{m,n}$ and b_m in Eqs. (32) and (33) are derived in appendix A.1 and A.2, respectively.

4.3 Inband-Aliasing Distortion

The inband-aliasing distortion term of the objective function, Eq. (25), is given by

$$\begin{aligned}\beta_{\mathbf{h}} &= \frac{1}{2\pi D^2}\int_{-\pi}^{\pi}\sum_{d=1}^{D-1}\left|H(e^{j\omega/D}W_D^d)\right|^2\, d\omega \\ &= \frac{1}{2\pi D^2}\int_{-\pi}^{\pi}\sum_{d=1}^{D-1}H(e^{j\omega/D}W_D^d)H^*(e^{j\omega/D}W_D^d)\, d\omega.\end{aligned} \tag{34}$$

Using Eq. (27), it can be rewritten as

$$\beta_{\mathbf{h}} = \frac{1}{2\pi D^2}\sum_{d=1}^{D-1}\mathbf{h}^T\left[\int_{-\pi}^{\pi}\phi_{\mathbf{h}}(e^{j\omega/D}W_D^d)\phi_{\mathbf{h}}^H(e^{j\omega/D}W_D^d)\, d\omega\right]\mathbf{h}, \tag{35}$$

which can be written in quadratic form

$$\beta_{\mathbf{h}} = \mathbf{h}^T\mathbf{C}\mathbf{h}, \tag{36}$$

where the $L_{\mathbf{h}} \times L_{\mathbf{h}}$ hermitian matrix $\mathbf{C}$ is defined as

$$\mathbf{C} = \frac{1}{2\pi D}\sum_{d=1}^{D-1}\int_{-\pi}^{\pi}\phi_{\mathbf{h}}(e^{j\omega/D}W_D^d)\phi_{\mathbf{h}}^H(e^{j\omega/D}W_D^d)\, d\omega. \tag{37}$$

Calculating the integral for the entries in matrix $\mathbf{C}$, the following expression for the matrix entry $C_{m,n}$ is obtained

$$C_{m,n} = \frac{\varphi(n-m)\sin(\pi(n-m)/D)}{\pi D(n-m)}, \tag{38}$$

where

$$\varphi(n) = D\sum_{k=-\infty}^{\infty}\delta(n-kD) - 1. \tag{39}$$

A derivation of Eq. (38) can be found appendix A.3.

4.4 The Optimal Prototype Analysis Filter

The objective function for the prototype analysis filter design, Eq. (22), expressed in terms of $\mathbf{h}$ is

$$\begin{aligned}\epsilon_{\mathbf{h}} &= \alpha_{\mathbf{h}} + \beta_{\mathbf{h}} \\ &= \mathbf{h}^T\mathbf{A}\mathbf{h} - 2\mathbf{h}^T\mathbf{b} + 1 + \mathbf{h}^T\mathbf{C}\mathbf{h} \\ &= \mathbf{h}^T(\mathbf{A}+\mathbf{C})\mathbf{h} - 2\mathbf{h}^T\mathbf{b} + 1. \end{aligned} \tag{40}$$

The solution to the design problem

$$\mathbf{h}_{\text{opt}} = \arg\min_{\mathbf{h}} \mathbf{h}^T\left(\mathbf{A}+\mathbf{C}\right)\mathbf{h} - 2\mathbf{h}^T\mathbf{b} + 1, \tag{41}$$

may be found by solving the set of linear equations

$$(\mathbf{A}+\mathbf{C})\mathbf{h} = \mathbf{b}. \tag{42}$$

5 Synthesis Filter Bank Design

5.1 Design Strategy

Similar to the design of analysis filter banks, presented in section 4, the design of synthesis filter banks reduces to the design of a single synthesis prototype filter. When designing the synthesis filter bank, the focus is on the performance of the analysis-synthesis filter bank as a whole. This implies that the synthesis filter bank is designed *given* an analysis filter bank, i.e. given the prototype analysis filter, designed using the method presented in the previous section. Different applications of filter banks require different strategies. The focus in the proposed method is on applications, which uses filtering operations in the subbands.

In the proposed design of the synthesis filter bank, the goal is to minimize amplitude and phase distortion of the analysis-synthesis filter bank and to minimize aliasing distortion in the output signal $Y(z)$. The objective function is the least square error

$$\epsilon_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|E(e^{j\omega})\right|^2 d\omega, \tag{43}$$

where

$$E(z) = T(z) - D(z) \tag{44}$$

denotes the total response error. Here, $T(z)$ is the complex-valued system response defined in Eq. (18). The desired complex-valued analysis-synthesis filter bank response $D(z)$ is defined for the special case with $\xi_m(z) = 1$,

$$D(z) = z^{-\tau_T}, \tag{45}$$

where τ_T is the desired total analysis-synthesis filter bank delay.

The complex error function in Eq. (44) can be written as

$$E(z) = E_0(z) + \sum_{m=0}^{M-1} \sum_{d=1}^{D-1} E_{m,d}(z). \tag{46}$$

The first term, $E_0(z)$, contains the desired spectral components and the summation in the second term contains all the undesired aliasing terms. The error function for the desired spectral content, $E_0(z)$, is defined as

$$E_0(z) = \sum_{m=0}^{M-1} E_{m,0}(z) = A_0(z) - D_0(z), \tag{47}$$

where the subband filters are unity, i.e. $\xi_m(z) = 1$. The total response excluding the undesired aliasing terms, $A_0(z)$, was defined in Eq. (17). The desired transfer function for $d = 0$ is

$$D_0(z) = \sum_{m=0}^{M-1} D_{m,0}(z) = z^{-\tau_T} \tag{48}$$

Note that $D_0(z) = D(z)$ according to (45). This implies that the desired transfer functions, $D_{m,d}(z)$, for $d > 0$ are zero, i.e. $D_{m,d} = 0$. The error functions for the undesired aliasing terms are defined as

$$E_{m,d}(z) = A_{m,d}(z) - D_{m,d}(z) = A_{m,d}(z). \tag{49}$$

Aliasing content may cancel out at the final reconstruction summation, so that the residual aliasing in the output signal $Y(z)$ is zero, even though the individual terms in the error function may have large energy, see Fig. 9 a). In applications without subband filtering, a direct minimization of Eq. (43) may be a plausible design strategy. In that case the filter bank will be a *nearly-perfect reconstruction filter bank*.

In applications where subband adaptive filtering is used, i.e. the subband filters, $\xi_m(z)$, are time-variant and unknown, the cancellation of aliasing terms at the

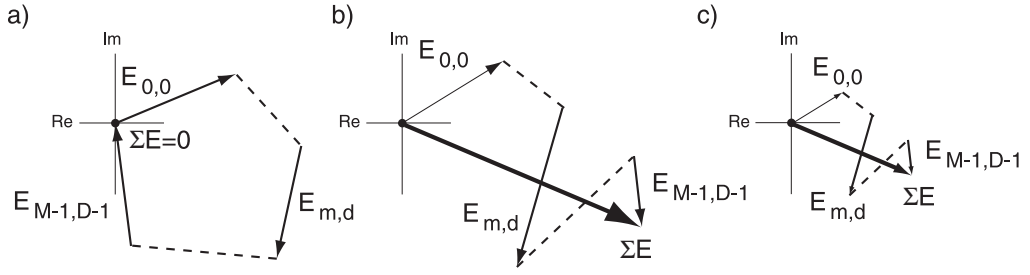

Figure 9: *Summation of complex errors, a) shows error cancellation with large errors, b) shows a large error case, and c) shows a case where individually minmized vectors, compared to b) yield a small total error.*

reconstruction summation cannot be controlled by a fixed prototype synthesis filter. To achieve this, the synthesis filter bank need to be time-variant and matched to the filters in the subbands.

However a fixed synthesis prototype filter may be designed so that the energy in the individual aliasing terms is minimized. The motive for this approach is illustrated in Fig. 9. The summation of errors in Eq. (46) may sum to zero in a perfect reconstruction case when the subband filters are known, as shown in Fig. 9 a).

When the magnitude and phase of error terms are altered by subband filters, which is depicted in Fig. 9 b), the error may be large since the individual vectors are large. A simplified example is shown in Fig. 10, for two terms. The perfect reconstruction case with $\xi_m(z) = 1$ is shown in Fig. 10 a). In Fig. 9 b), the maximum error is shown when the error vectors may be altered by the subband filters, $\xi_m(z)$. The grey zone denote the range in which the vectors may be altered. If the synthesis filter bank is designed so that the aliasing errors individually are small, the sum will also be small.

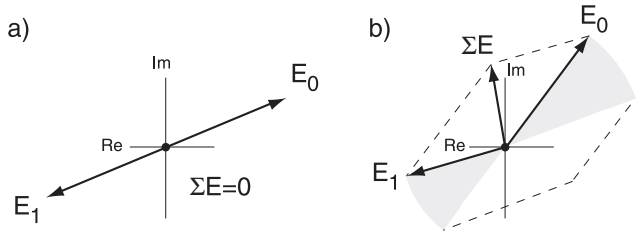

Figure 10: *Simple case with two errors, a) shows the zero error case, and b) shows the worst case where the grey zones denote the range in which the amplitude and phase of the vectors may be altered by subband filtering.*

The objective function is defined as

$$\epsilon_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|E_0(e^{j\omega})\right|^2 d\omega + \sum_{m=0}^{M-1}\sum_{d=1}^{D-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|E_{m,d}(e^{j\omega})\right|^2 d\omega, \tag{50}$$

where the energies of all undesired aliasing terms will be individually minimized independently of $\xi_m(z)$. In summary, the proposed objective function of the design method is

$$\epsilon_{\mathbf{g}}(\mathbf{h}) = \gamma_{\mathbf{g}}(\mathbf{h}) + \delta_{\mathbf{g}}(\mathbf{h}) \tag{51}$$

where

$$\gamma_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\sum_{m=0}^{M-1}A_{m,0}(e^{j\omega}) - D_0(e^{j\omega})\right|^2 d\omega \tag{52}$$

is the *Total Response Error* and

$$\delta_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{d=1}^{D-1}\sum_{m=0}^{M-1}\left|A_{m,d}(e^{j\omega})\right|^2 d\omega \tag{53}$$

is the *Residual Aliasing Distortion*.

In the next two sections, the two terms $\gamma_{\mathbf{g}}(\mathbf{h})$ and $\delta_{\mathbf{g}}(\mathbf{h})$ will be derived in terms of the impulse response of the synthesis prototype filter.

5.2 Total Response Error

The total response error is

$$\begin{aligned}\gamma_{\mathbf{g}}(\mathbf{h}) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|A_0(e^{j\omega}) - D_0(e^{j\omega})\right|^2 d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(A_0(e^{j\omega}) - D_0(e^{j\omega})\right)^*\left(A_0(e^{j\omega}) - D_0(e^{j\omega})\right) d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\left|A_0(e^{j\omega})\right|^2 - 2\mathrm{Re}\left\{D_0^*(e^{j\omega})A_0(e^{j\omega})\right\} + \left|D_0(e^{j\omega})\right|^2\right] d\omega, \quad (54)\end{aligned}$$

where $A_0(z)$ is defined in Eq. (17), and $\xi_m(z) = 1$ for $m = 0, \ldots, M-1$. As with the prototype analysis filter in Eq. (27), the prototype synthesis filter can be written in terms of its impulse response as follows

$$G(z) = \sum_{n=0}^{L_{\mathbf{g}}-1} g(n)z^{-n} = \mathbf{g}^T\phi_{\mathbf{g}}(z) \quad (55)$$

where $\mathbf{g} = [g(0), \ldots, g(L_{\mathbf{g}}-1)]^T$ and $\phi_{\mathbf{g}}(z) = [1, z^{-1}, \ldots, z^{-L_{\mathbf{g}}+1}]^T$.

The summation over m in Eq. (54) can be written in terms of the impulse responses $\mathbf{h}$ and $\mathbf{g}$. Inserting Eq. (27) and Eq. (55) into $A_0(z)$, with $\xi_m(z) = 1$, yields

$$\begin{aligned}A_0(z) &= \frac{1}{D}\sum_{m=0}^{M-1} H(zW_M^m)G(zW_M^m) \\ &= \frac{1}{D}\sum_{m=0}^{M-1} \mathbf{h}^T\phi_{\mathbf{h}}(zW_M^m)\phi_{\mathbf{g}}^T(zW_M^m)\mathbf{g} \\ &= \mathbf{h}^T\mathbf{\Psi}(z)\mathbf{g}, \quad (56)\end{aligned}$$

where

$$\mathbf{\Psi}(z) = \frac{1}{D}\sum_{m=0}^{M-1}\mathbf{\Phi}_{m,0}(z) \quad (57)$$

and

$$\mathbf{\Phi}_{m,d}(z) = \phi_{\mathbf{h}}(zW_M^mW_D^d)\phi_{\mathbf{g}}^T(zW_M^m). \quad (58)$$

Substituting Eq. (48), and Eq. (56) into Eq. (54) yields

$$\gamma_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\left|\mathbf{h}^T\mathbf{\Psi}(e^{j\omega})\mathbf{g}\right|^2 - 2\mathrm{Re}\left\{e^{j\omega\tau_T}\mathbf{h}^T\mathbf{\Psi}(e^{j\omega})\mathbf{g}\right\} + 1\right] d\omega, \quad (59)$$

which can be rewritten as

$$\gamma_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\mathbf{g}^T\mathbf{\Psi}^H(e^{j\omega})\mathbf{h}^*\mathbf{h}^T\mathbf{\Psi}(e^{j\omega})\mathbf{g} - 2\mathrm{Re}\left\{e^{j\omega\tau_T}\mathbf{h}^T\mathbf{\Psi}(e^{j\omega})\mathbf{g}\right\} + 1\right]d\omega. \quad (60)$$

Finally, the following quadratic form is obtained

$$\gamma_{\mathbf{g}}(\mathbf{h}) = \mathbf{g}^T\mathbf{E}\mathbf{g} - 2\mathbf{g}^T\mathbf{f} + 1, \quad (61)$$

where the $L_{\mathbf{g}} \times L_{\mathbf{g}}$ hermitian matrix $\mathbf{E}$ is defined as

$$\mathbf{E} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathbf{\Psi}^H(e^{j\omega})\mathbf{h}^*\mathbf{h}^T\mathbf{\Psi}(e^{j\omega})d\omega, \quad (62)$$

and the $L_{\mathbf{g}} \times 1$ vector $\mathbf{f}$ is defined as

$$\mathbf{f} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{Re}\left\{e^{j\omega\tau_T}\mathbf{\Psi}^T(e^{j\omega})\mathbf{h}\right\}d\omega. \quad (63)$$

Calculating the integral for all matrix entries, the following expression for the matrix entries $E_{p,q}$ is obtained

$$E_{p,q} = \frac{M^2}{D^2}\sum_{\kappa=-\infty}^{\infty} h^*(\kappa M - p)h(\kappa M - q) \quad (64)$$

Similarly, for the vector entries f_p

$$f_p = \frac{\lambda}{\pi D}h(\tau_T - p) \quad (65)$$

with

$$\lambda = \sum_{m=0}^{M-1}\cos(2\pi\tau_T m/M). \quad (66)$$

Here the total delay,τ_T, is assumed to be an integer. The derivations of these expressions are given in appendix A.4 and A.5

5.3 Residual Aliasing Distortion

The residual aliasing distortion term in Eq. (51) is given by

$$\delta_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi}\sum_{d=1}^{D-1}\sum_{m=0}^{M-1}\int_{-\pi}^{\pi}\left|A_{m,d}(e^{j\omega})\right|^2 d\omega. \quad (67)$$

Eq. (67) can be rewritten using the impulse response vectors $\mathbf{h}$ and $\mathbf{g}$, defined in Eq. (27) and Eq. (55), respectively, and the matrix $\mathbf{\Phi}_{m,d}(z)$, defined in Eq. (58) according to

$$\delta_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi D^2}\sum_{d=1}^{D-1}\sum_{m=0}^{M-1}\int_{-\pi}^{\pi}\left|\mathbf{h}^T\mathbf{\Phi}_{m,d}(e^{j\omega})\mathbf{g}\right|^2 d\omega. \quad (68)$$

This can be rewritten as

$$\delta_{\mathbf{g}}(\mathbf{h}) = \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \int_{-\pi}^{\pi} \mathbf{g}^T \mathbf{\Phi}_{m,d}^H(e^{j\omega}) \mathbf{h}^* \mathbf{h}^T \mathbf{\Phi}_{m,d}(e^{j\omega}) \mathbf{g} \, d\omega, \tag{69}$$

which finally leads to the quadratic form

$$\delta_{\mathbf{g}}(\mathbf{h}) = \mathbf{g}^T \mathbf{P} \mathbf{g} \tag{70}$$

where the $L_{\mathbf{g}} \times L_{\mathbf{g}}$ hermitian matrix $\mathbf{P}$ is defined as

$$\mathbf{P} = \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \int_{-\pi}^{\pi} \mathbf{\Phi}_{m,d}^H(e^{j\omega}) \mathbf{h}^* \mathbf{h}^T \mathbf{\Phi}_{m,d}(e^{j\omega}) \, d\omega. \tag{71}$$

Calculating the integral for all matrix entries, the following expression for $P_{p,q}$ is obtained

$$P_{p,q} = \frac{M}{D^2} \sum_{l=-\infty}^{\infty} h^*(l+q) h(l+p) \varphi(p-q), \tag{72}$$

where

$$\varphi(n) = D \sum_{k=-\infty}^{\infty} \delta(n - kD) - 1. \tag{73}$$

5.4 The optimal synthesis prototype filter

The optimal prototype analysis filter, in terms of minimal total response error and minimal energy in the aliasing components, is found by minimizing the objective function from Eq. (51)

$$\epsilon_{\mathbf{g}}(\mathbf{h}) = \gamma_{\mathbf{g}}(\mathbf{h}) + v \delta_{\mathbf{g}}(\mathbf{h}). \tag{74}$$

Here, a weighting factor v is introduced in order to enable emphasis on either the total response error ($0 < v < 1$) or the residual aliasing distortion ($v > 1$). Inserting Eq. (61) and Eq. (70) into (74) yields

$$\epsilon_{\mathbf{g}}(\mathbf{h}) = \mathbf{g}^T (\mathbf{E} + v\mathbf{P}) \mathbf{g} - 2\mathbf{g}^T \mathbf{f} + 1, \tag{75}$$

The solution

$$\mathbf{g}_{\text{opt}} = \arg\min_{\mathbf{g}} \epsilon_{\mathbf{g}}(\mathbf{h}), \tag{76}$$

can be found by solving the set of linear equations

$$(\mathbf{E} + v\mathbf{P}) \mathbf{g} = \mathbf{f}. \tag{77}$$

6 Design Parameters

6.1 Pre-Specified parameters

In order to use the design method, a number of parameters needs to be set. The most important parameter is the number of subbands, M. The filter lengths of the analysis and synthesis filters, $L_{\mathbf{h}}$ and $L_{\mathbf{g}}$, and the decimation factor, D, are then chosen. The filter lengths are commonly set to multiples of the decimation factor because it gives rise to the efficient polyphase implementation. The parameters mentioned here influence both performance and complexity, as will be seen in the following sections. Other design parameters are the passband boundary frequency, optional weighting in the synthesis filter bank optimization, the delay of the analysis filter bank, and the total delay.

6.2 Passband boundary frequency

The bandwidth of the prototype analysis filter is set by chosing the passband boundary frequency ω_p. The parameter determines the bandwidth of all analysis filters and it is therefore important, since it might influence the performance of the application. Prototype analysis filters, obtained from Eq. (41), with different passband boundary frequencies are shown in Fig. 11. The figure illustrates that a lower stop band level, and thus lower inband-aliasing distortion, can be obtained by chosing a smaller passband boundary frequency.

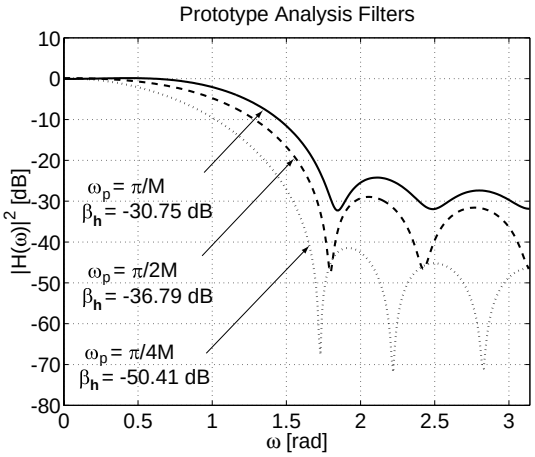

Figure 11: *Prototype analysis filter responses with* $L_{\mathbf{h}} = 8$, $M = 4$ *and* $D = 2$. *The frequency response of the filters is shown for passband boundaries* $\omega_p = \frac{\pi}{M}$, $\omega_p = \frac{\pi}{2M}$ *and* $\omega_p = \frac{\pi}{4M}$. *The corresponding inband-aliasing distortion,* $\beta_{\mathbf{h}}$, *is also shown.*

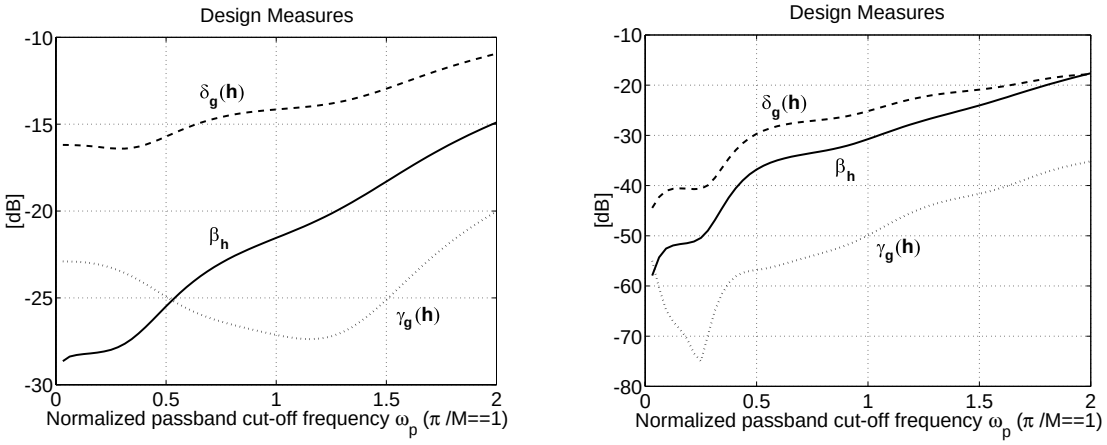

Figure 12: *Filter bank performance as a function of the passband boundary frequency. The solid line is the inband-aliasing distortion, the striped line is the residual aliasing distortion and the dotted line is the response error. The number of subbands is* $M = 4$ *and the decimation factor is* $D = 2$*. In the left figure, the filter lengths are set to* $L_{\mathbf{h}} = L_{\mathbf{g}} = 4$ *and in the right figure* $L_{\mathbf{h}} = L_{\mathbf{g}} = 8$*.*

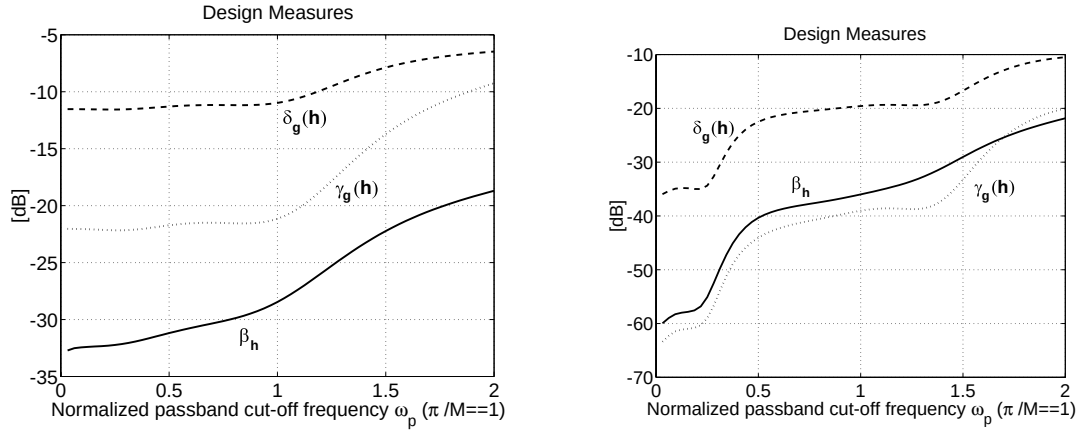

Figure 13: *Filter bank as a function of the passband boundary frequency. The solid line is the inband-aliasing distortion, the striped line is the residual aliasing distortion and the striped line is the total response error. The number of subbands is* $M = 8$ *and the decimation factor is* $D = 4$*. In the left figure, the filter lengths are set to* $L_{\mathbf{h}} = L_{\mathbf{g}} = 8$ *and in the right figure* $L_{\mathbf{h}} = L_{\mathbf{g}} = 16$*.*

The passband boundary frequency also affects the total response error and the residual aliasing distortion since the synthesis filter bank design has a dependency on the analysis filter bank design. Figs. 12 and 13 show how the inband-aliasing distortion, $\beta_{\mathbf{h}}$ in Eq. (25), total response error, $\gamma_{\mathbf{g}}(\mathbf{h})$ in Eq. (52), and the residual aliasing distortion, $\delta_{\mathbf{g}}(\mathbf{h})$ in Eq. (53), are affected by the passband boundary frequency parameters in the analysis filter bank design. All figures show that a passband boundary frequency lower than $\omega_p = \frac{\pi}{M}$ generally yields lower aliasing distortions. The amplitude error has local minima while the inband and residual aliasing distortion decrease monotonically when the passband boundary frequency is decreased.

6.3 Weighting in the Synthesis Filter Bank Design

The error weighting parameter, v, can be used to control the emphasis in the optimization on either the filter bank response error or the residual aliasing distortion. Figs. 14 and 15 show the total response error and the residual aliasing distortion as a function of the weighting factor. When the weighting factor is decreased ($v < 1$), the total response error decreases while the residual aliasing distortion converges to a certain level. By increasing the weighting factor the residual aliasing distortion can be decreased at the expense of the total response error, which converges to the maximal level of 0 dB.

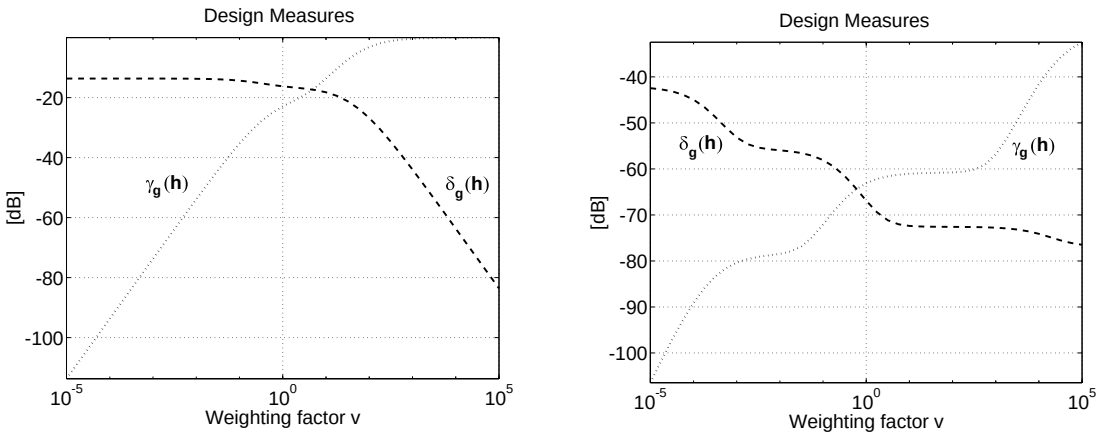

Figure 14: *Synthesis filter bank performance as a function of the weighting factor. The striped line is the residual aliasing distortion and the dotted line is the total response error. The number of subbands is $M = 4$ and the decimation factor is $D = 2$. The passband boundary frequency is set to $\omega_p = \frac{\pi}{8M}$. In the left figure, the filter lengths are set to $L_{\mathbf{h}} = L_{\mathbf{g}} = 4$ and in the right figure $L_{\mathbf{h}} = L_{\mathbf{g}} = 8$.*

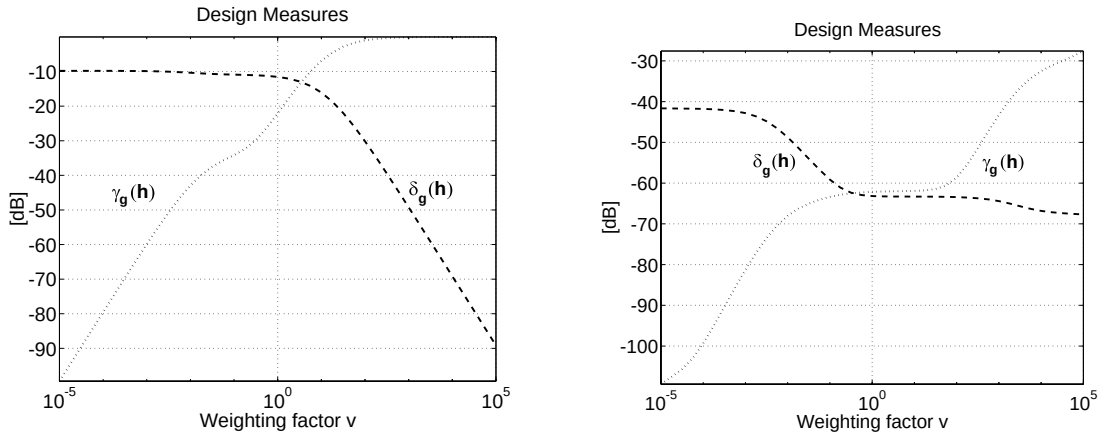

Figure 15: *Filter bank performance as a function of the weighting factor. The striped line is the residual aliasing distortion and the dotted line is the total response error. The number of subbands is* $M = 8$ *and the decimation factor is* $D = 4$*. The passband boundary frequency is set to* $\omega_p = \frac{\pi}{8M}$*. In the left figure, the filter lengths are set to* $L_{\mathbf{h}} = L_{\mathbf{g}} = 8$ *and in the right figure* $L_{\mathbf{h}} = L_{\mathbf{g}} = 16$*.*

6.4 Critical Sampling and Oversampling

Oversampling ($D < M$) will increase the degrees of freedom in the synthesis filter design. The total response error and the residual aliasing distortion can be decreased by choosing a smaller decimation factor than the critical decimation factor ($D = M$). The minimum values of the total response errors and the residual aliasing distortions for different decimation factors in some specific filter-bank design scenarios are shown in Fig. 16.

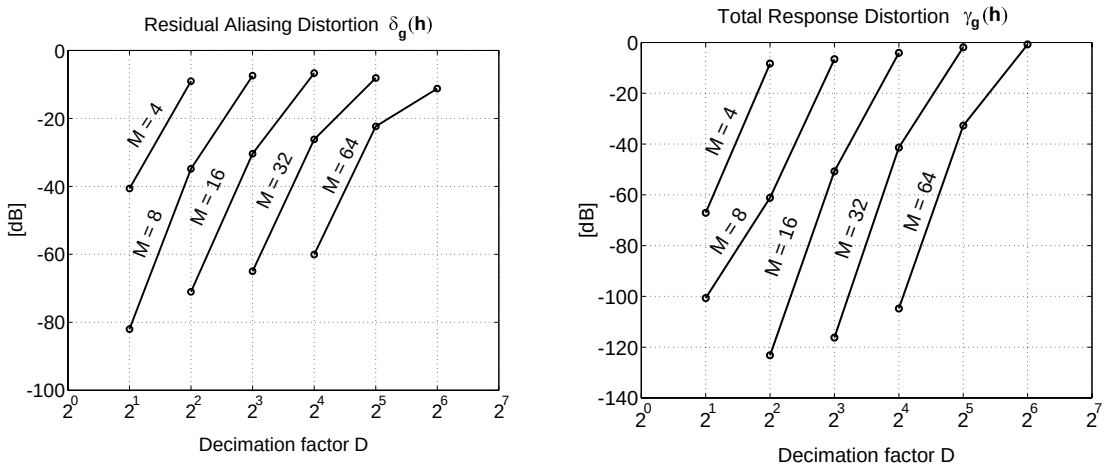

Figure 16: *Minimum total response error, $\gamma_{\mathbf{g}}(\mathbf{h})$, and minimum residual aliasing distortion, $\delta_{\mathbf{g}}(\mathbf{h})$, as a function of the decimation factor D. The filter lengths are set to $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M$. Each curve corresponds to a specific number of subbands, $M = 4, \ldots, 64$. In all cases, the passband boundary frequency is set to $\omega_p = \frac{\pi}{8M}$ and the prototype filters have linear phase. The black dots denote decimation factors chosen as powers of two.*

6.5 Analysis filter bank delay and total filter bank delay

The relation between reduced total delay and the filter bank performance is illustrated in Fig. 17. When setting the system delay at a fixed value, it can be of great importance to study how the delay of the analysis filter bank can be set in order to obtain low distortion. This is shown in Figs. 18 and 19. The inband-aliasing distortion has a unique minimum at $\tau_H = \frac{L_{\mathbf{h}}-1}{2}$ in all evaluation cases. Here the analysis prototype filters are linear phase filters with even-symmetric impulse responses, so that all degrees of freedom are used to obtain low inband-aliasing distortion.

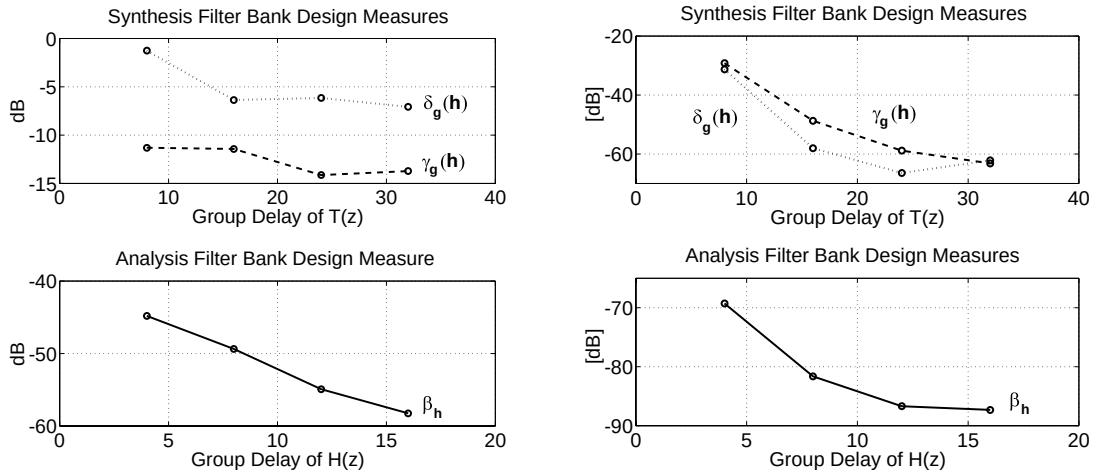

Figure 17: *Total response error and residual alisiasing distortion as a function of τ_H and τ_T. The number of subbands is $M = 4$. The decimation factor is set to $D = 2$. The filter lengths are set to $L_{\mathbf{h}} = L_{\mathbf{g}} = 8$ in the left figure and $L_{\mathbf{h}} = L_{\mathbf{g}} = 16$ in the right figure. The delay of the analysis filter-bank is set to $\tau_H = \frac{1}{2}L_{\mathbf{h}}$.*

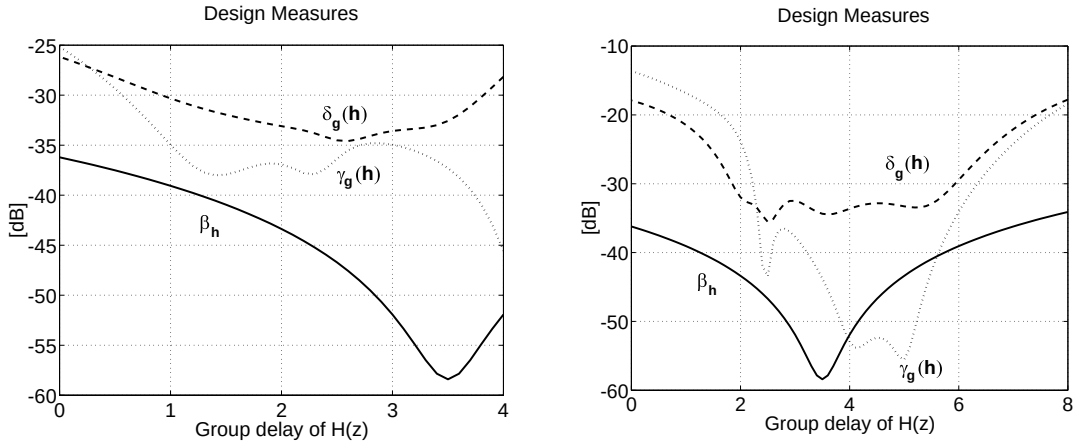

Figure 18: *Filter bank performance measures as a function of the analysis prototype filter delay, τ_H. The solid line is the inband-aliasing distortion, the striped line is the residual aliasing distortion, and the dotted line is the total response error. The number of subbands is $M = 4$, the decimation factor is $D = 2$ and the prototype filter lengths are set to $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M = 8$. The desired system delay is set to $\tau_T = 4$ for the left figure and $\tau_T = 8$ for the right figure. The group-delay of the prototype analysis filter τ_H is varied from 0 to τ_T, which means that τ_G varies from τ_T to 0 since $\tau_T = \tau_H + \tau_G$.*

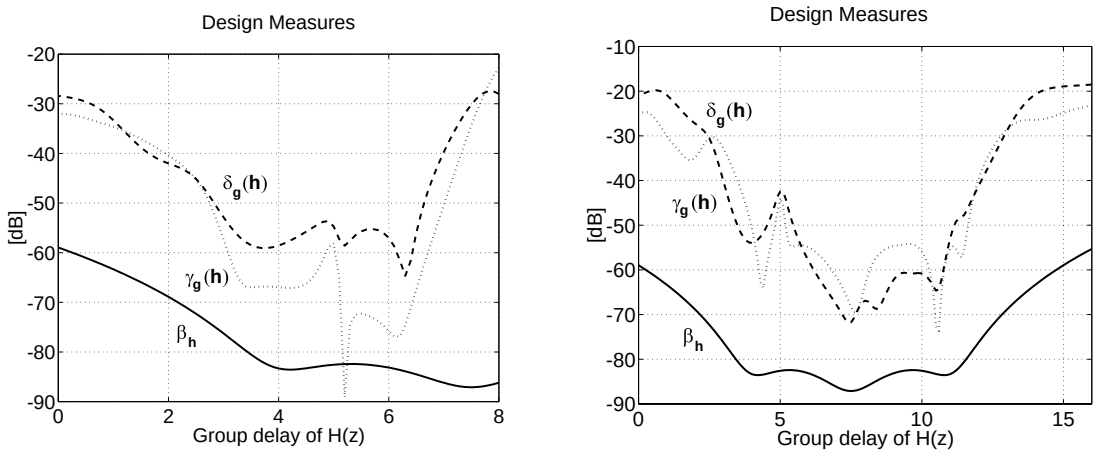

Figure 19: *Design errors as a function of the analysis prototype filter group-delay. The solid line is the inband aliasing distortion, the striped line is the residual aliasing distortion, and the dotted line is the total response error. The number of subbands is* $M = 4$*, the decimation factor is* $D = 2$ *and the prototype filter lengths are set to* $L_{\mathbf{h}} = L_{\mathbf{g}} = 4M = 16$*. The desired system delay is set to* $\tau_T = 8$ *for the left figure and* $\tau_T = 16$ *for the right figure. The delay of the prototype analysis filter* τ_H *is varied from 0 to* τ_T*, which means that* τ_G *varies from* τ_T *to* 0 *since* $\tau_T = \tau_H + \tau_G$*.*

7 Efficient Implementation and Computational Complexity of Modulated Filter Banks

In this section an efficient structure for two-times oversampled ($D = \frac{M}{2}$) analysis and synthesis polyphase modulated filter banks is derived.

7.1 Type I Polyphase Decomposition

First the *Type I Polyphase Decomposition* is introduced. The prototype analysis filter $H(z)$ can be decomposed into polyphase components according to

$$\begin{aligned} H(z) &= \sum_{n=-\infty}^{\infty} h(nK) z^{-nK} \\ &\quad + z^{-1} \sum_{n=-\infty}^{\infty} h(nK+1) z^{-nK} \\ &\quad \vdots \\ &\quad + z^{-K+1} \sum_{n=-\infty}^{\infty} h(nL+K-1) z^{-nK} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{K-1} z^{-k} \sum_{n=-\infty}^{\infty} h(nK+k) z^{-nK} \\
&= \sum_{k=0}^{K-1} z^{-k} E_k(z^K), \qquad (78)
\end{aligned}
$$

where K is the number of elements in the decomposition. The filters $e_k(n) = h(nK+k)$ are called the *type I polyphase components* of $H(z)$.

7.2 Analysis Filter Bank

Inserting the definition of Eq. 7 into Eq. 78, the polyphase decomposition of the m-th analysis filter in the analysis filter bank becomes

$$
\begin{aligned}
H_m(z) &= H(zW_M^m) \\
&= \sum_{k=0}^{K-1} (zW_M^m)^{-k} E_k([zW_M^m]^K) \\
&= \sum_{k=0}^{K-1} z^{-k} W_M^{-mk} E_k([zW_M^m]^K). \qquad (79)
\end{aligned}
$$

When the number of decompositions, K, is set equal to the decimation factor, i.e. $K = D = \frac{M}{2}$, the polyphase decomposition becomes

$$
H_m(z) = \sum_{k=0}^{D-1} z^{-k} W_M^{-mk} E_k([zW_M^m]^D) \qquad (80)
$$

Since

$$
W_M^{mD} = e^{-j\pi m} = \begin{cases} -1, & \text{when } m \text{ is odd} \\ 1, & \text{when } m \text{ is even} \end{cases} \qquad (81)
$$

it implies that odd and even subbands have to be treated separately. Sets of even and odd indices are defined by $\mathcal{M}_{\text{even}} = [0, 2, 4, \ldots, M-2]$ and $\mathcal{M}_{\text{odd}} = [1, 3, 5, \ldots, M-1]$. The decomposition of the analysis filters in the even subbands and the analysis filters in the odd subbands are

$$
H_m(z) = \sum_{k=0}^{D-1} z^{-k} W_M^{-mk} E_k(z^D), \qquad m \in \mathcal{M}_{\text{even}}, \qquad (82)
$$

and

$$
H_m(z) = \sum_{k=0}^{D-1} z^{-k} W_M^{-mk} E_k([-z]^D), \qquad m \in \mathcal{M}_{\text{odd}}. \qquad (83)
$$

The polyphase components for odd and even subbands are defined as $E_k(z)$ and $E'_k(z)$, respectively,

$$
\begin{aligned}
E_k(z) \quad &\longleftrightarrow \quad e_k(n) = h(nD+k), \\
E'_k(z) = E_k(-z) \quad &\longleftrightarrow \quad e'_k(n) = h(nD+k)(-1)^n. \qquad (84)
\end{aligned}
$$

Figure 20: *Noble Identity 1.*

The analysis filter bank structure with filters given in Eq. (82) and Eq. (83) consists of a delay line, the polyphase components, $E_k(z)$ and $E'_k(z)$, and two summations over k with coefficients W_M^{-mk}.

The implementation becomes efficient when the filter outputs are decimated by $D = \frac{M}{2}$ and the noble identity is applied, meaning that the decimation operators have been placed before the polyphase components, and thereby giving low rate filtering, see Fig. 20.

This involves that $E_k(z^D)$ are replaced by $E_k(z)$ and, similarly, for the odd subbands $E'_k(z^D)$ are replaced by $E'_k(z)$. The summation over k for each subband m with coeffcients W_M^{-mk} can be implemented efficiently using the IFFT algorithm. Since

$$W_M^{-mk} = e^{j2\pi mk/M} = e^{j2\pi\frac{k}{2}m/D} \tag{85}$$

for $k = 0, \ldots, D-1$ and $m \in \mathcal{M}_{\text{even}}$, an D-length FFT can be used for the even subbands. Similarly for the odd subbands,

$$W_M^{-mk} = e^{j2\pi mk/M} = e^{j2\pi\frac{(m-1)}{2}k/\frac{M}{2}} e^{j2\pi k/M} \tag{86}$$

for $k = 0, \ldots, D-1$ and $m \in \mathcal{M}_{\text{odd}}$, which enables the use of an D-length FFT when each channel k is multiplied with the correction factor $e^{j2\pi k/M} = W_M^{-k}$. Fig. 21 shows how the analysis filter bank is implemented.

7.3 Type II Polyphase Decomposition

In the synthesis filter bank, the *Type II Polyphase Decomposition* is used. The decomposition is basically the same as the Type I Decomposition with a slight difference

$$\begin{aligned} G(z) &= \sum_{k=0}^{K-1} z^{-k} F_{K-k-1}(z^K) \\ &= \sum_{k=0}^{K-1} z^{-[K-k-1]} F_k(z^K), \end{aligned} \tag{87}$$

where $F_k(z), k = 0, \ldots, K-1$ denote the type II polyphase components of $G(z)$. The first part of the equation is the original Type I Decomposition and the second part is the Type II equation. The difference is the order of summation.

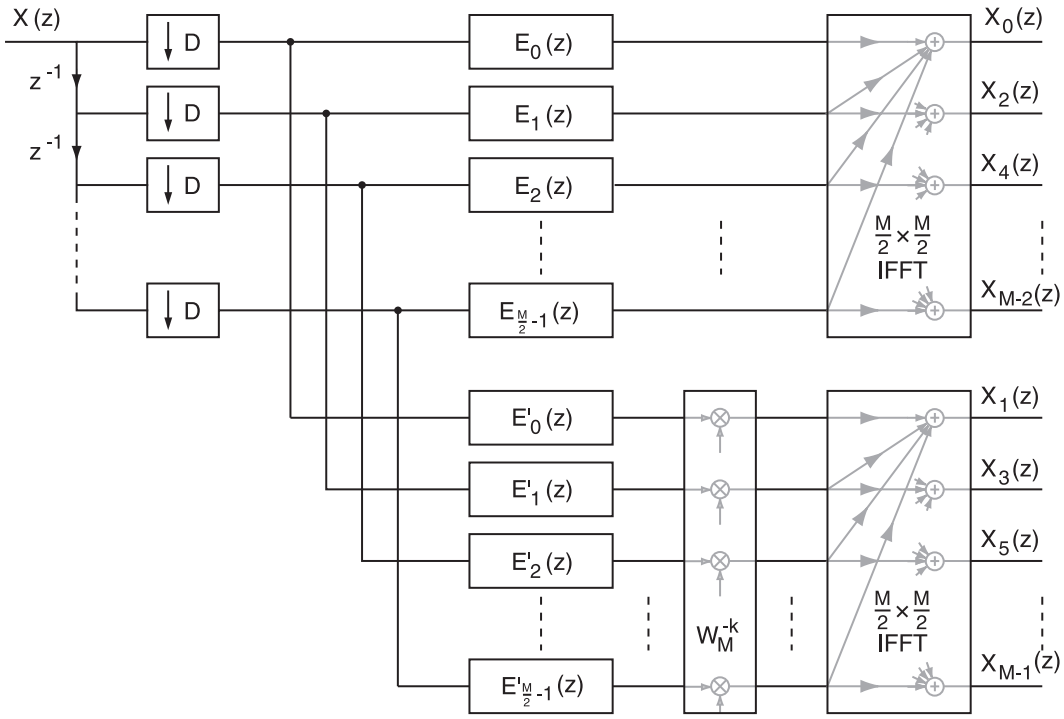

Figure 21: *Polyphase implementation of a modulated analysis filter bank with two-times oversampling.*

7.4 Synthesis Filter Bank

The Type II Polyphase Decomposition of the synthesis filter in subband m is

$$\begin{aligned} G_m(z) &= \sum_{k=0}^{D-1} z^{-[D-k-1]} W_M^{mk} F_k(z^D) \qquad m \in \mathcal{M}_{\text{even}} \\ G_m(z) &= \sum_{k=0}^{D-1} z^{-[D-k-1]} W_M^{mk} F_k([-z]^D) \qquad m \in \mathcal{M}_{\text{odd}}, \end{aligned} \tag{88}$$

where

$$\begin{aligned} F_k(z) \qquad &\longleftrightarrow \qquad f_k(n) = g(nD + D - k - 1), \\ F'_k(z) = F_k(-z) \qquad &\longleftrightarrow \qquad f'_k(n) = g(nD + D - k - 1)(-1)^n. \end{aligned} \tag{89}$$

Note that the number of elements in the decomposition is set to $K = D$, similar to the analysis filter bank. The output $Y(z)$ of the synthesis filter bank may be written in terms of the subband signals $Y_m(z)$ and the filters $G_m(z)$ according to

$$\begin{aligned} Y(z) &= \sum_{m \in \mathcal{M}_{\text{even}}} \sum_{k=0}^{D-1} W_M^{mk} Y_m(z) F_k(z^M) z^{-[D-k-1]} \\ &\quad + \sum_{m \in \mathcal{M}_{\text{odd}}} \sum_{k=0}^{D-1} W_M^{mk} Y_m(z) F'_k(z^M) z^{-[D-k-1]}, \end{aligned} \tag{90}$$

which may be rearranged as

$$Y(z) = \sum_{k=0}^{D-1} \left[\sum_{m \in \mathcal{M}_{\text{even}}} W_M^{mk} Y_m(z) F_k(z^M) + \sum_{m \in \mathcal{M}_{\text{odd}}} W_M^{mk} Y_m(z) F'_k(z^M) \right] z^{-[D-k-1]} \tag{91}$$

When the signals $Y_m(z)$ are interpolated and when Noble Identity 2 is applied, the polyphase components and the interpolation operations trade places, see Fig. 22.

Figure 22: *Noble Identity 2.*

As with the analysis filter bank, the summation over m for each k with coeffcients W_M^{mk} can be implemented efficiently using the FFT algorithm. For the odd subbands a correction by $e^{-j2\pi k/M} = W_M^k$ is needed after the transform operation. Fig. 23 shows the polyphase implemented synthesis filter bank.

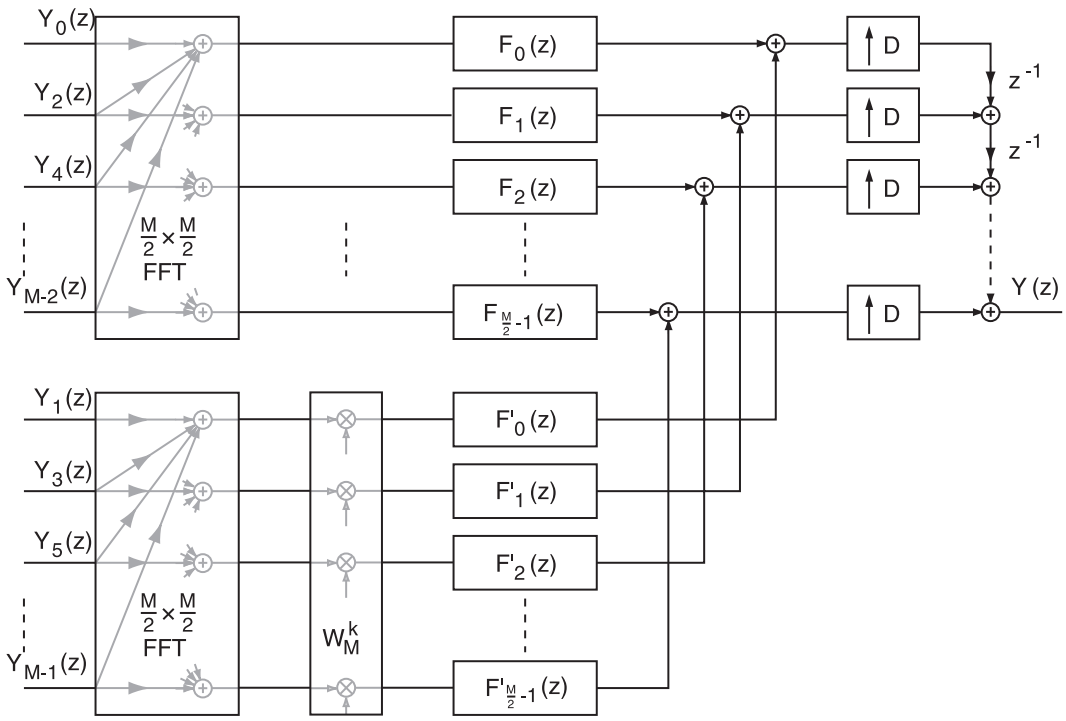

Figure 23: *Polyphase implementation of a modulated synthesis filter bank with two-times oversampling.*

7.5 Computational Complexity

This section discusses the computational complexity of critically sampled modulated filter banks and two-times oversampled modulated filterbanks. The filter banks are implemented using the polyphase implementation, described in the previous sections.

The complexity of signal processing structures is usually measured in terms of the number of additions and multiplications. These additions and multiplications can either be real or complex valued. Complex additions can be decomposed into two real additions and complex multiplications can be decomposed into four real multiplications and two real additions.

The number of real and complex operations in critically sampled filter banks are shown in Table 1. The number of operations in the oversampled filter bank are given in Table 2.

Implementation Part	Real Additions	Real Multiplications
Polyphase Components	$L - M$	L
IFFT	$2M \log_2 M$	$2M \log_2 M$
FFT	$2M \log_2 M$	$2M \log_2 M$
Polyphase Components	$L - M$	L
Delay + Sum Line	$M - 1$	0
Total	$2L + M(4 \log_2 M - 1) - 1$	$2L + 4M \log_2 M$

Table 1: *Computational complexity for the critically sampled polyphase filter bank structure.*

Implementation Part	Real Additions	Real Multiplications
Polyphase Components	$2(L - M)$	2L
Complex Correction	M	2M
IFFTs	$4M \log_2 \frac{M}{2}$	$4M \log_2 \frac{M}{2}$
FFTs	$4M \log_2 \frac{M}{2}$	$4M \log_2 \frac{M}{2}$
Complex Correction	M	2M
Polyphase Components	$2(L - M)$	2L
Delay + Sum Line	$\frac{M}{2} - 1$	0
Total	$4L + M(8 \log_2 \frac{M}{2} - 1\frac{1}{2}) - 1$	$4L + 4M(1 + 2 \log_2 \frac{M}{2})$

Table 2: *Computational complexity for the oversampled polyphase filter bank structure.*

The computation complexity of the polyphase implementation of an oversampled filter bank in relation to the complexity of a critically sampled filter bank is shown in Fig. 24. The ratio of the number of operations is plotted for $M = 4, \ldots, 512$, where M is a power of 2. The filter length is chosen as $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M$. The plots show that the oversampled structure is not twice as complex as the critically sampled structure. This is because by the FFT operations being less than half as complex in the oversampled polyphase implementation.

The computational complexity of the subband filtering structure is evaluated in Fig. 25. The subband filtering structure consists of analysis and synthesis polyphase implemented filter banks and filters in the subbands. The number of subbands varies between $M = 4, \ldots, 512$ while the length of the subband filters is correspondingly reduced from $L_\xi = 128, \ldots, 1$. The ratio between the complexity of the oversampled filter bank and the critically sampled filter bank is shown as a function of the number of subbands, M.

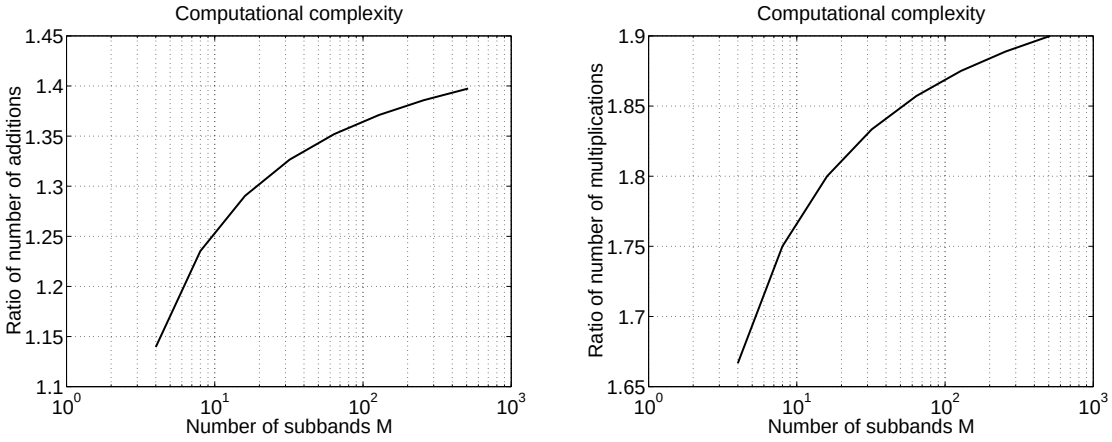

Figure 24: *Filter bank complexity. Comparison between critically sampled and oversampled filter bank implementations. The ratio describes how much more demanding the oversampled implementation is compared to the critical sampled implementation.*

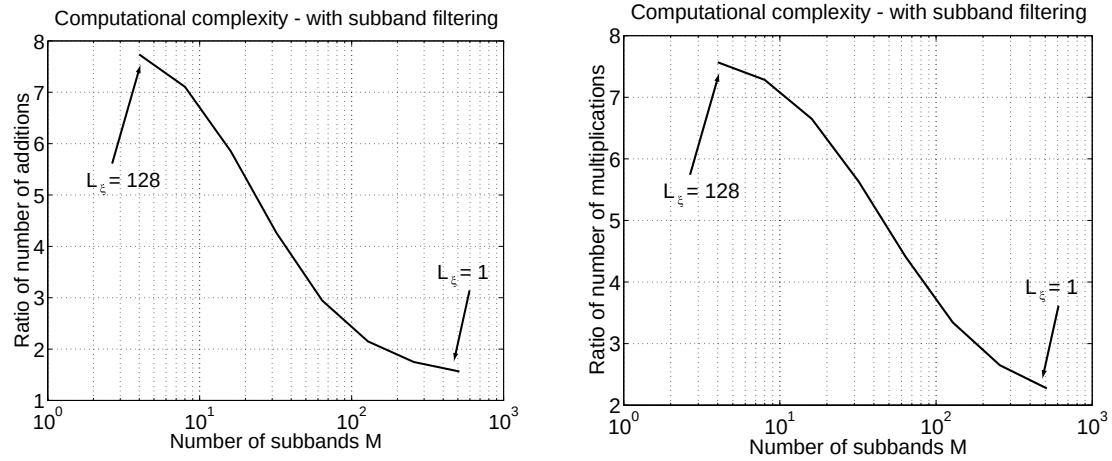

Figure 25: *Filter bank complexity. The figures show the complexity ratio including the subband filtering operations. In the critical sampled case, the least number of subbands is 4 with length 128 subband filters and the maximum number of subbands is 512, with 1-coefficient subband filters. In the oversampled case, the subband filters are twice as long.*

8 Filter Bank Examples

8.1 Filter Bank Cases

In this section four different filter banks are designed using the proposed method and they are compared against four conventional filter banks with known structure as reference. The number of subbands is set to $M = 8$ in all filter banks. The four reference filter banks are

1. A perfect reconstruction FFT filter bank with $L_{\mathbf{h}} = L_{\mathbf{g}} = M$ and $D = M$ (critically sampled). The filter bank is implemented using delay lines and FFT and IFFT operations. The prototype filters are rectangular windows.

2. A perfect reconstruction filter bank with $M = L_{\mathbf{h}} = L_{\mathbf{g}}$ and $D = M/2$ (oversampling). The prototype analysis filter is an M-point Hanning window, and the prototype synthesis filter is an M-point rectangular window.

3. A critically sampled, modulated filter bank with $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M$ and $D = M$. The analysis and synthesis filters are modulated from the same prototype low-pass linear phase filter designed using the window method with a Hamming window.

4. An oversampled modulated filter bank with $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M$ and $D = M/2$. The prototype analysis and synthesis filters are obtained as for the reference filter bank in case 3.

Four filter banks with the same number of subbands were designed, according to the method presented in part I. Note that these filter banks have the same structure as the filter banks in case 3 and 4 and that only the prototype filters differ.

5. A critically sampled modulated filter bank with $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M$ and $D = M$. The prototype analysis and synthesis filters are designed using the proposed method. The desired total delay is set to $\tau_T = 2M$.

6. An oversampled modulated filter bank with $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M$ and $D = M/2$. The prototype analysis and synthesis filters are designed using the proposed method. The desired total delay is set to $\tau_T = 2M$.

7. A critically sampled modulated filter bank with $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M$ and $D = M$. The prototype analysis and synthesis filters are designed using the proposed method. The desired total delay is reduced to $\tau_T = M$.

8. An oversampled modulated filter bank with $L_{\mathbf{h}} = L_{\mathbf{g}} = 2M$ and $D = M/2$. The prototype analysis and synthesis filters are designed using the proposed method. The desired total delay is reduced to $\tau_T = M$.

Note that the filter banks in case 5 and 6 have the same delay as in case 3 and 4. The delay reduced to half for the filter banks in cases 7 and 8.

8.2 Performance Evaluation

In appendix B, illustrations are presented for all eight filter bank cases. The frequency response and the impulse response of the prototype analysis and synthesis filters are plotted.

8.2.1 Inband-Aliasing

Fig. 26 shows the inband-aliasing distortion, which is one of the objectives of minimization in the analysis filter bank design. The measure is defined in Eq. (25). Clearly, the oversampled filter banks (even case numbers) have less inband-aliasing distortion compared to the critically sampled filter banks (odd case numbers). Comparing the oversampled filter bank cases 4, 6, and 8, which have the same structure, it can be seen that the optimal filter bank in case 6 has the least distortion. In case 8, the reduction of delay requires more degrees of freedom, which results in higher inband-aliasing distortion. In appendix B, the spectrum of the aliasing component in subband $m = 0$ is plotted.

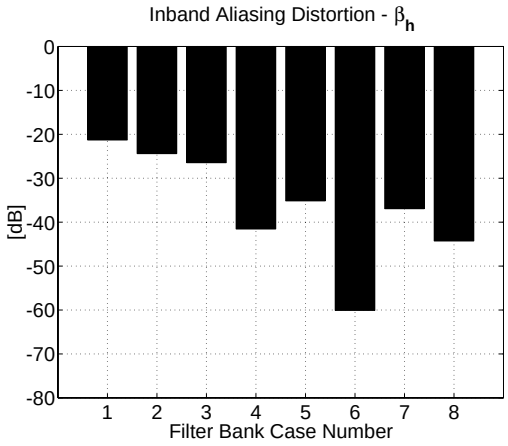

Figure 26: *Inband-Aliasing Distortion for the filter banks in the evaluation.*

8.2.2 Residual Aliasing

The residual aliasing energy, i.e. energy of the aliasing present in the output of the synthesis filter bank, is calculated according to

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\sum_{d=1}^{D-1}\sum_{m=0}^{M-1}A_{m,d}(e^{j\omega})\right|^2 d\omega. \tag{92}$$

Here, no subband filtering is applied, i.e. $\xi_m(z) = 1$. Fig. 27 a) shows the aliasing energies for the eight filter bank cases. Clearly, the perfect reconstruction filter

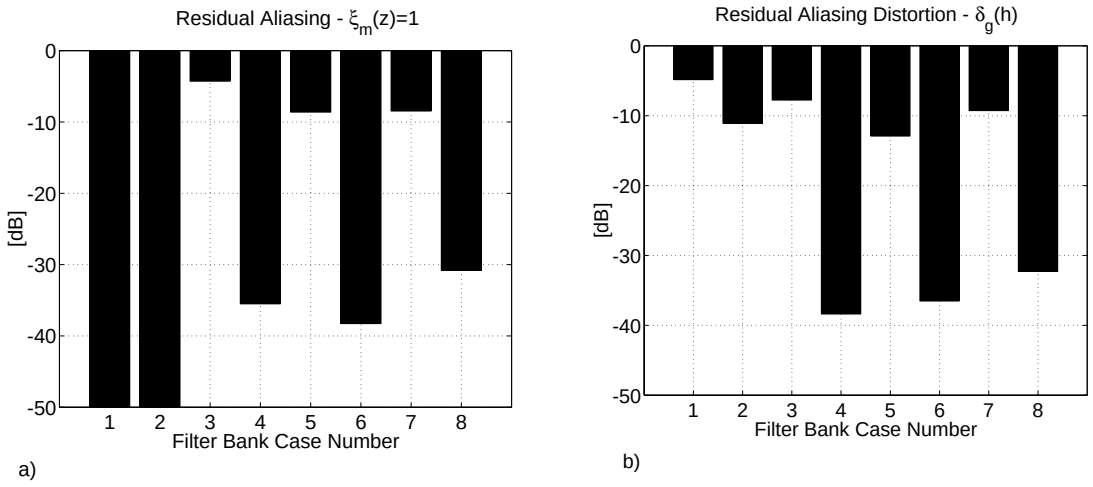

Figure 27: *a) Residual aliasing and b) Residual Aliasing Distortion measure for the filter banks in the evaluation.*

banks in case 1 and 2 have zero reconstruction aliasing. In the other cases, the oversampled filter banks have lower aliasing than the critically sampled filter banks. The residual aliasing distortion measure, defined in Eq. (53), which is one of the objectives of minimization in the synthesis filter bank design is shown in Fig. 27 b). In this case, the perfect reconstruction filter banks have the highest bound, while the optimal filter banks have the lowest bound, especially the oversampled filter banks.

8.2.3 Amplitude and Phase Distortion

The amplitude and phase distortion of the total filter bank are evaluated seperately. Both measures are included in the total response error in the synthesis filter bank design. The amplitude distortion is defined by

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Big||T(e^{j\omega})|-|D(e^{j\omega})|\Big|^2\,d\omega, \tag{93}$$

with $T(z)$ defined as in Eq. (18), and $D(z)$ defined as in Eq. (45). The measures are presented in Fig. 28. Clearly, the perfect reconstruction filter banks in cases 1 and 2 have zero amplitude distortion. The optimal filter banks have lowest amplitude distortion, especially the oversampled filter banks. The phase distortion, which is defined as

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Big|\angle T(e^{j\omega})-\angle D(e^{j\omega})\Big|\,d\omega, \tag{94}$$

is illustrated in Fig. 28. The optimal filter banks with oversampling have the lowest phase distortion, except for the perfect reconstruction filter banks in case 1 and 2, which have zero phase error.

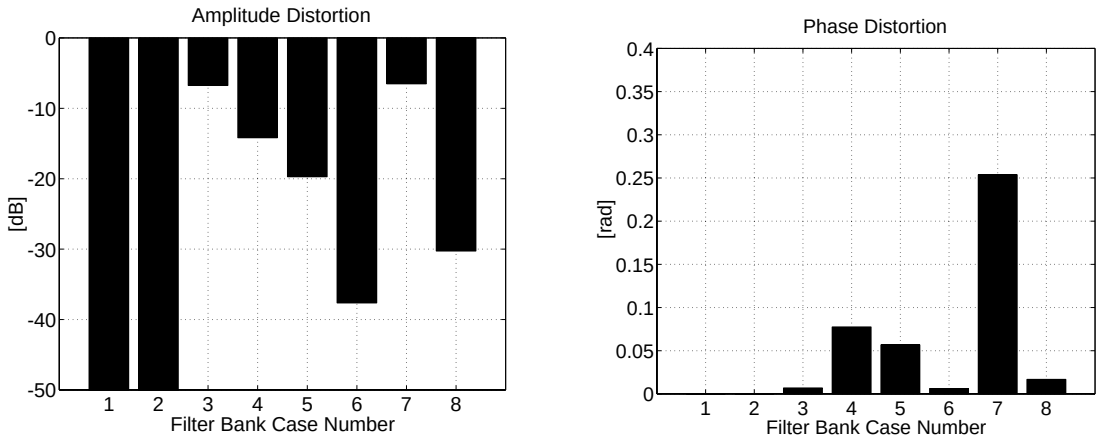

Figure 28: *Amplitude and Phase Distortion for the filter banks in the evaluation.*

9 Evaluation in Subband Acoustic Echo Cancellation

In this section, the filter banks presented in section 8 are evaluated in acoustic echo cancellation with real speech signals. Speech signals recorded in a conference room are used to evaluate the performance of the application when using different filter banks.

9.1 Optimal Subband Echo Cancelling

Subband acoustic echo cancellation is an application of subband system identification of the acoustic echo path in for example conference telephony, as illustrated in Fig. 29. Acoustic echo cancellation deals with the problem of far-end speech enter-

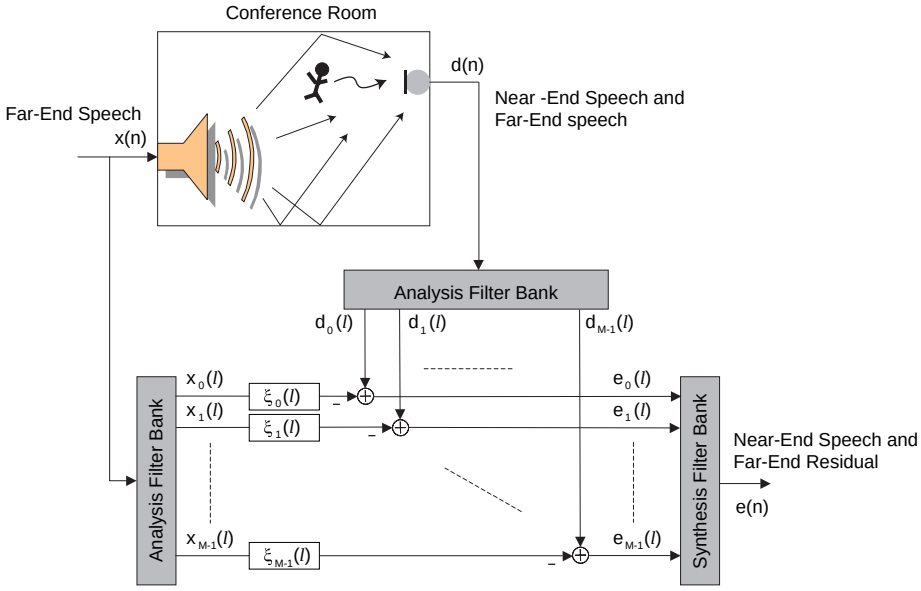

Figure 29: *Subband System Identification.*

ing the return path, mixed with the near-end speech. Only the near-end speech is desired. Hence, the returning far-end speech may cause annoying audible echoes at the far-end. Conference telephony is a situation where acoustic echo cancellation succesfully may be applied to remove far-end speech components from the return signal.

The far-end speech signal, emitted by the loudspeaker in the room, is denoted by $x(n)$. The near-end microphone signal is denoted by $d(n)$, and contains the near-end speech and the undesired far-end speech. The residual signal of the echo cancellation system is denoted by $e(n)$. In subband acoustic echo cancellation, the loudspeaker signal and the microphone signal are transformed to the subband domain by analysis filter banks. After processing of the signals in the subband

domain, the residual signal is transformed to the full band domain by a synthesis filter bank. In the subband domain, the signals are denoted by $x_m(l)$, $t_m(l)$ and $e_m(l)$. Fixed least squares optimal FIR filters, denoted by $\xi_m(z)$, are used in the subbands. The filters are individually optimized based on the input signals, $x(n)$ and $d(n)$. In the evaluation, white noise signals are used to determine the optimal filter coefficients. The coefficients are calculated according to

$$\hat{\xi}_m = \arg\min_{\xi} \sum_l |e_m(l)|^2 = \mathbf{R}^{-1}_{x_m x_m} \mathbf{r}_{d_m x_m} \tag{95}$$

Where $\xi_m = [\xi_m(0), \ldots, \xi_m(L_\xi - 1)]^T$ denotes the subband filters, where L_ξ is the filter length. The autocorrelation matrix, $\mathbf{R}_{x_m x_m}$, is defined as

$$\mathbf{R}_{x_m x_m} = \sum_l \mathbf{x}_m(l)\mathbf{x}^H_m(l), \tag{96}$$

and the autocorrelation vector, $\mathbf{r}_{d_m x_m}$, is defined as

$$\mathbf{r}_{d_m x_m} = \sum_l d_m(l)\mathbf{x}^H_m(l) \tag{97}$$

where $\mathbf{x}_m = [x_m(l), x_m(l-1), \ldots, x_m(l - L_\xi + 1)]^T$.

9.2 Evaluation with Speech Signals

The performance of subband acoustic echo cancellation with recorded speech signals in a conference room is evaluated for the filter bank types decribed in section 8. The identified system is a conference room, of which the impulse response is shown in Fig. 30. All results in this evaluation are based on two sequences of speech, a

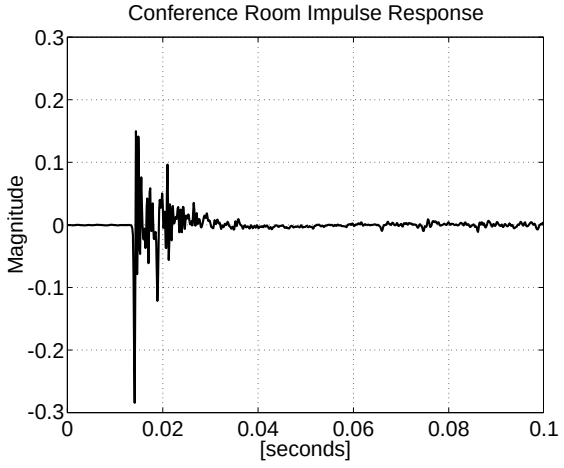

Figure 30: *Conference room impulse response.*

Case	Filter Bank	M	D	$L_{\mathbf{h}}$	$L_{\mathbf{g}}$	τ_H	τ_T
1	FFT	128	128	128	128	63.5	127
2	Hanning	128	64	128	128	63.5	127
3	Window/Crit	128	128	256	256	127.5	256
4	Window/Over	128	64	256	256	127.5	256
5	Opt/Crit	128	128	256	256	127.5	256
6	Opt/Over	128	64	256	256	127.5	256
7	Opt/Crit less delay	128	128	256	256	63.5	128
8	Opt/Over less delay	128	64	256	256	63.5	128

Table 3: *Pre-specified filter bank properties.*

near-end and a far-end utterance, sampled at 8 kHz and each with a duration of about 4 seconds. The properties of the filter banks are given in Table 3. All filter banks have 128 subbands. The number of filter weights, L_ξ, used in each subband filter is set to $L_\xi = 4$. Measures, which are used to evaluate the performance of the acoustic echo cancellation are the suppression of the far-end component in the microphone signal, $d(n)$, and the distortion of the near-end speech component in $e(n)$. The suppression of the acoustic echo is defined as

$$\mathcal{S} = \frac{\mathcal{C} \int_{-\pi}^{\pi} P_{t,FE}(\omega) d\omega}{\int_{-\pi}^{\pi} P_{e,FE}(\omega) d\omega} \tag{98}$$

where the normalization constant $\mathcal{C}$ is

$$\mathcal{C} = \frac{\int_{-\pi}^{\pi} P_{e,NE}(\omega) d\omega}{\int_{-\pi}^{\pi} P_{t,NE}(\omega) d\omega}. \tag{99}$$

Here, $P_t(\omega)$ and $P_e(\omega)$, are the Power Spectral Densities (PSD) of the microphone signal, $t(n)$, and the residual signal, $e(n)$. The far-end and near-end components in the signals are denoted by FE and NE, respectively. The distortion of the near-end speech is defined as

$$\mathcal{D} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|\frac{1}{\mathcal{C}} P_{e,NE}(\omega) - P_{t,NE}(\omega)|}{P_{t,NE}(\omega)} d\omega. \tag{100}$$

The suppression of the far-end signal and the distortion of the near-end speech for the different filter bank cases are presented in Fig. 31. The oversampled filter banks (even case numbers) give better suppression than the critically sampled cases. The oversampled filter bank in case 6 gives the best suppression. The distortion caused by oversampled filter banks is generally lower than the distortion caused by the critically sampled filter banks. The perfect reconstruction filter banks in cases 1 and 2 give the lowest near-end distortion. The suppression of the far-end component in the return signal, for filter bank case number 6, is shown in Fig. 32.

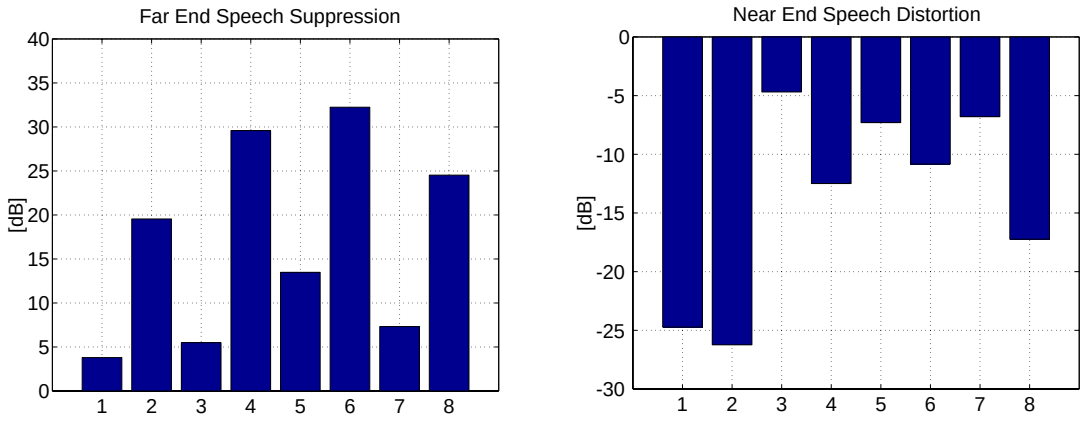

Figure 31: *Suppression of Acoustic Echo and Near-End Speech distortion in subband acoustic echo cancellation using the filter banks in the evaluation.*

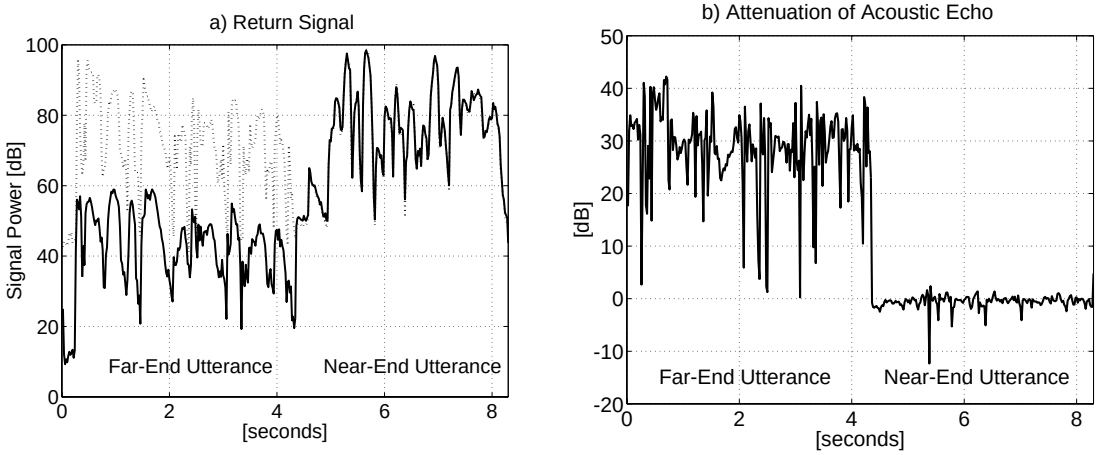

Figure 32: *a) Signal power of the far-end and the near-end utterances in the residual signal compared to the microphone signal. b) Suppression of the acoustic echo.*

In order to determine the level of aliasing present in the subband signals $x_m(l)$ and $d_m(l)$, $m = 0, \ldots, M-1$, and the output near-end speech, $e(n)$, the Signal-to-Aliasing Ratio (SAR) measure is introduced. The average subband SAR of $x_m(l)$ is defined as

$$\mathcal{SAR}_{x_m} = \frac{1}{M} \sum_{m=0}^{M-1} \frac{\int_{-\pi}^{\pi} P_{x_m}(\omega) d\omega}{\int_{-\pi}^{\pi} P_{x_m,A}(\omega) d\omega}, \tag{101}$$

where $P_{x_m,A}(\omega)$ denotes the PSD of the aliasing in $x_m(l)$. The subband Signal-to-Aliasing Ratio of the subband signals $X_m(z)$ and $T_m(z)$ is shown in Fig. 33 for the eight filter bank cases. The best performance is obtained with the optimal oversampled filter bank in case 6, and the other oversampled filter banks are better than the corresponding critically sampled filter banks.

The SAR of the residual signal, $e(n)$, is defined as

$$\mathcal{SAR}_e = \frac{\int_{-\pi}^{\pi} P_e(\omega) d\omega}{\int_{-\pi}^{\pi} P_{e,A}(\omega) d\omega}, \tag{102}$$

and is illustrated in Fig. 34 for the filter bank cases. The SAR of the residual signal is given for the near-end and far-end components separately since the residual aliasing depends on the filtering in the subbands, and the far-end subband signals, $x_m(l)$ are filtered while the near-end subband signals, $d_m(l)$, are not. The oversampled filterbanks give rise to a much better SAR in the far-end component, although the SAR is generally not high when comparing the SAR measures for the far-end component with the SAR measures for the near-end components. For the near-end component, the SAR is better for the perfect reconstruction filter banks and for the oversampled filter banks in general. Since the SAR is a relative measure, the suppression of the far-end speech has to be taken into consideration when comparing the far-end SAR values with the near-end SAR values.

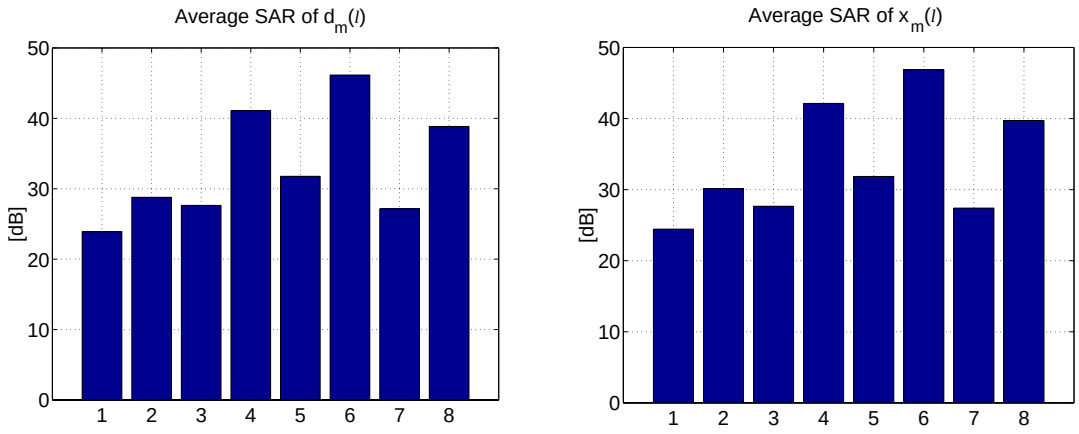

Figure 33: *Subband Signal-to-Aliasing Ratios for the subband microphone signal $d_m(l)$, and the subband loudspeaker signal, $x_m(l)$.*

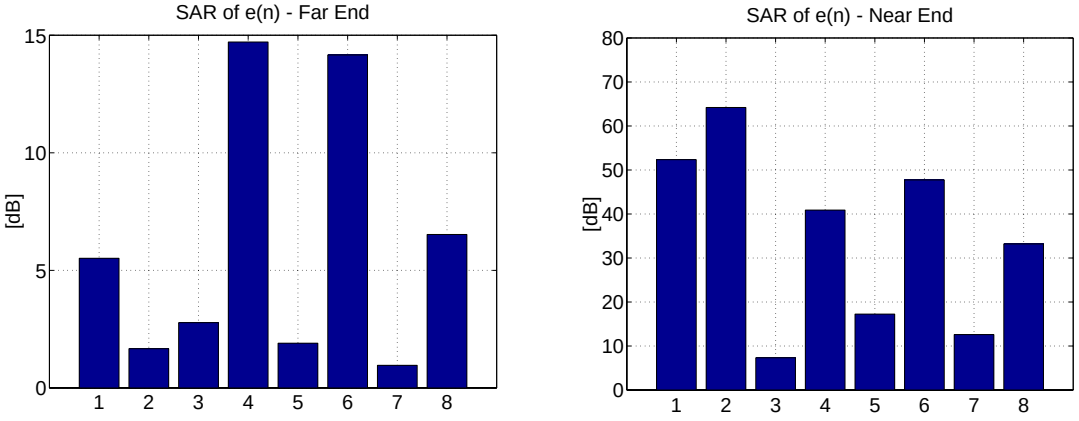

Figure 34: *Ouput Signal-to-Aliasing Ratios for the far-end and near-end components in $e(n)$.*

10 Evaluation in Optimal Subband Speech Enhancement

In this section, the filter bank types described in section 8 are evaluated in optimal subband speech enhancement.

10.1 Optimal Subband Speech Enhancement

Subband speech enhancement is used to remove disturbing components from speech signals. The input signal, $x(n)$ consists of desired components and undesired components, of which the latter usually consists of noise. The noise component is assumed to be uncorrelated with the desired component. An example of a signal is a microphone signal, where the desired component is speech and the undesired component consists of background noise. This noisy speech signal, $x(n)$, and the noise reference signal, $t(n)$ are transformed to the subband domain by analysis filter banks, see Fig. 35. Access to the background noise is needed to calculate the subband filter coefficients. The subband signals, $x_m(l)$, are modified by subband filters, $\xi_m(z)$, which depend on both the $x_m(l)$ and $d_m(l)$, so that the undersired noise is removed. In this evaluation, the subband filters consist of one filter coefficient. The residual, $e_m(l)$, contains the enhanced speech. The single parameter

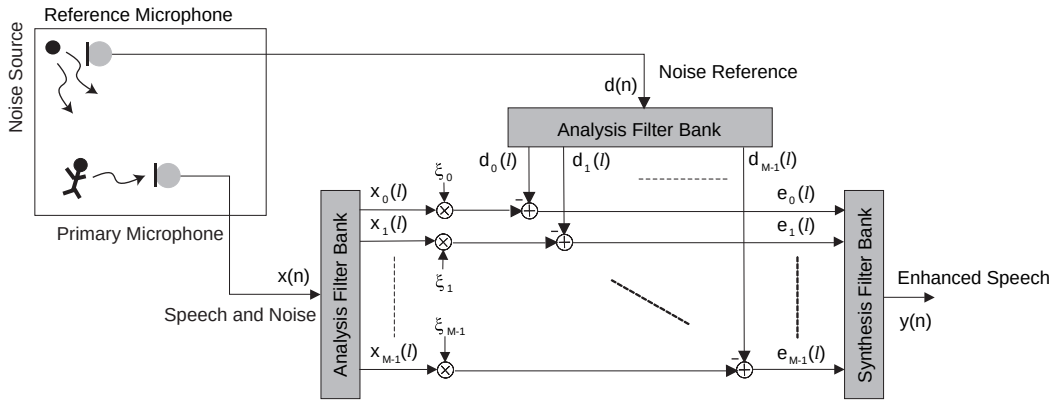

Figure 35: *Subband Speech Enhancement.*

optimal subband filter coefficients are calculated using the method of least squares, in each subband individually. Subband error signals are formed according to

$$e_m(l) = \xi_m x_m(l) - d_m(l). \tag{103}$$

The optimal solution of the subband filters is found by minimizing the least squares error

$$\hat{\xi}_m = \arg\min_{\xi} \sum_l |e_m(l)|^2 = \frac{r_{t_m x_m}(0)}{r_{x_m x_m}(0)}, \tag{104}$$

where ξ_m is the filter weight in subband m. The autocorrelation, $r_{x_m x_m}(0)$, is defined as

$$r_{x_m x_m}(0) = \sum_l x_m(l) x_m^*(ll), \tag{105}$$

and the crosscorrelation, $r_{t_m x_m}(0)$, is defined as

$$r_{t_m x_m}(0) = \sum_l d_m(l) x_m^*(l). \tag{106}$$

10.2 Performance Evaluation

The amount of noise suppression is measured by

$$\mathcal{S} = \frac{\mathcal{C} \int_{-\pi}^{\pi} P_{x,n}(\omega) d\omega}{\int_{-\pi}^{\pi} P_{e,n} d\omega}, \tag{107}$$

where the normalization constant, $\mathcal{C}$, is defined as

$$\mathcal{C} = \frac{\int_{-\pi}^{\pi} P_{e,s}(\omega) d\omega}{\int_{-\pi}^{\pi} P_{x,s}(\omega) d\omega}. \tag{108}$$

Here, P_x and P_e denote the Power Spectral Densities of the input noisy speech, $x(n)$ and the output enhanced speech $e(n)$. The indices n and s denote the noise and speech components in the signals. Speech distortion is evaluated using the following measure

$$\mathcal{S} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|\frac{1}{\mathcal{C}} P_{e,s}(\omega) - P_{t,s}(\omega)|}{P_{t,s}(\omega)} d\omega. \tag{109}$$

A noisy speech utterance of 4 seconds, sampled at 8 kHz is enhanced using the different filter banks. Background noise is are recorded in a open-air cefeteria, to which clean speech is added to form the noisy speech input signal. In Fig. 36, the noise suppression and the speech distortion measures are illustrated. The dependency of the application performance on the filter banks is very small. To determine the level of aliasing present in the subband signals $x_m(l)$, $m = 0, \ldots, M-1$, and the output enhanced speech, $e(n)$, Signal-to-Aliasing Ratio (SAR) measures are introduced. The average subband SAR of $x_m(l)$ is defined as

$$\mathcal{SAR}_{x_m} = \frac{1}{M} \sum_{m=0}^{M-1} \frac{\int_{-\pi}^{\pi} P_{x_m}(\omega) d\omega}{\int_{-\pi}^{\pi} P_{x_m,A}(\omega) d\omega}, \tag{110}$$

where $P_{x_m,A}(\omega)$ denotes the PSD of the aliasing component in $x_m(l)$. The SAR of $e(n)$ is defined as

$$\mathcal{SAR}_e = \frac{\int_{-\pi}^{\pi} P_e(\omega) d\omega}{\int_{-\pi}^{\pi} P_{e,A}(\omega) d\omega}. \tag{111}$$

where $P_{e,A}(\omega)$ denotes the PSD of the aliasing component in $e(n)$. The SAR measures are plotted in Fig. 37. The filter bank designs with the proposed method, clearly give higher SAR ratios than the other filter banks.

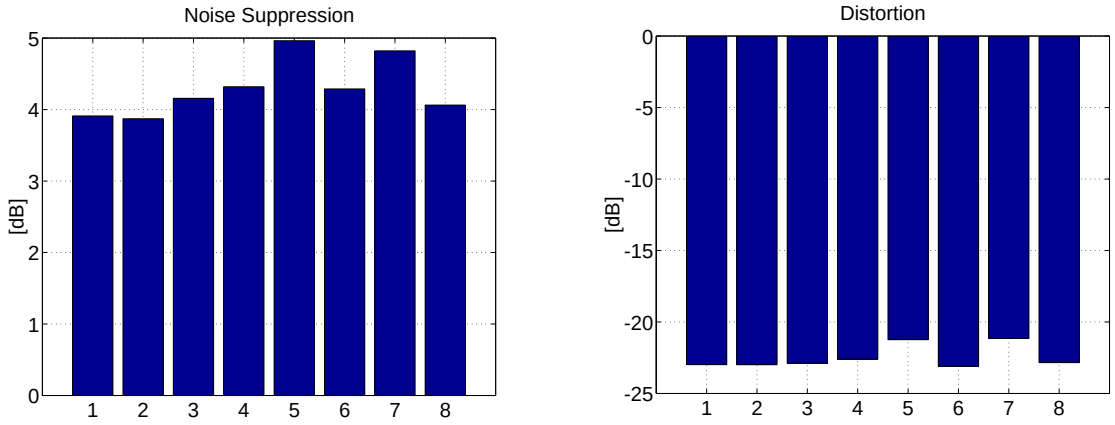

Figure 36: *Noise suppression and Speech Distortion in the speech enhancement scenarios, using different filter banks.*

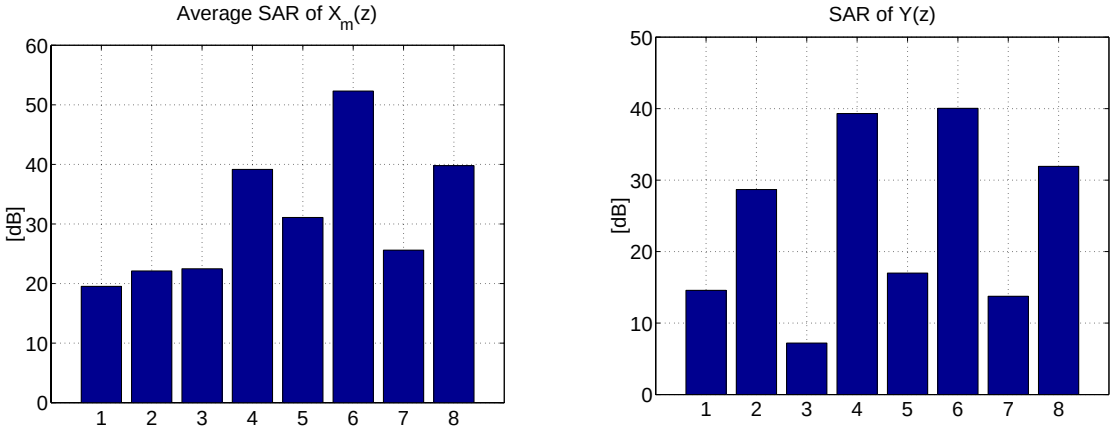

Figure 37: *Signal to Aliasing Ratio for $X_m(z)$ and $Y(z)$ in the scenarios.*

11 Evaluation in Spectral Estimation

11.1 Filter Bank Spectral Estimation

The problem of finding the *Power Spectral Density* (PSD) is to find the signal power distribution over frequency from a finite record of stationary data sequence [20]. Filter banks may be used to estimate the PSD. Some adaptive filtering applications use spectral estimates obtained from the analysis filter bank used for the signal decomposition. An example of such an application is spectral subtraction.

11.2 Spectral Subtraction

Spectral subtraction is a method to restore speech signals which are embedded in background noise. The method subtracts the estimated noise spectrum from the noisy speech spectrum, which results in an clean speech spectrum estimate. The signal model which is used in the spectral subtraction method is given by

$$x(n) = s(n) + v(n) \tag{112}$$

where $x(n)$ is the noisy signal, $s(n)$ is the clean speech signal and $v(n)$ is the additive noise. In the spectral domain this yields

$$P_x(\omega) = P_s(\omega) + P_v(\omega) \tag{113}$$

where $P(\omega)$ denotes *Power Spectral Density* and $s(n)$ and $v(n)$ are assumed uncorrelated. In this domain, the power spectrum estimate of the noise is subtracted from the power spectrum estimate of the noisy speech signal and it results in a power spectrum estimate of the clean speech signal

$$\hat{P}_s(\omega) = \hat{P}_x(\omega) - \hat{P}_v(\omega). \tag{114}$$

This can be expressed as a filtering operation

$$\hat{P}_s(\omega) = \left(1 - \frac{\hat{P}_v(\omega)}{\hat{P}_x(\omega)}\right) \hat{P}_x(\omega) = |Q(\omega)|^2 \hat{P}_x(\omega) \tag{115}$$

where

$$Q(\omega) = \sqrt{\left(1 - \frac{\hat{P}_v(\omega)}{\hat{P}_x(\omega)}\right)}. \tag{116}$$

The filtering operation in the frequency domain described by $Q(\omega)$ can be implemented using analysis and synthesis filter banks with complex-valued gain weights in the subbands. The analysis filter bank is used for the signal decomposition and for estimation of the input noisy speech spectrum and the input noise spectrum by dividing the input signal into noisy speech and noise-only periods.

11.3 Variance of Spectral Estimates

When an FFT filter bank is used (filter bank case 1 in the evaluation), the spectral estimate is referred to as the *Periodogram*. The periodogram may be expressed as follows [21]:

$$\hat{P}_x(m) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(k)e^{-j2\pi nm/N}\right|^2 = \frac{1}{N}\sum_{n=0}^{N-1}\sum_{l=0}^{N-1} x(n)x^*(l)e^{-j2\pi(n-l)m/N} \qquad (117)$$

here, the number of frequency bins, M, is equal to the number of observations, N. When using a modulated filter bank, see Fig. 38, where the number of frequency

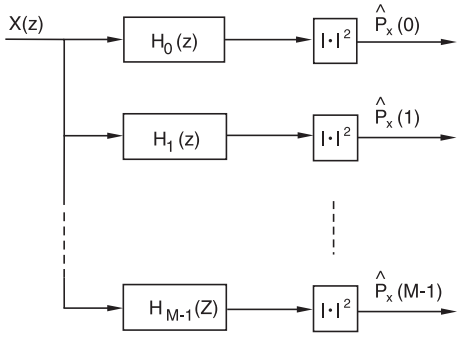

Figure 38: *Filter Bank Spectrum Estimation.*

bins M is equal or less than the number of observations N, the PSD estimate can be expressed as

$$\begin{aligned}\hat{P}_x(m) &= \frac{1}{L_{\mathbf{h}}U}\left|\sum_{n=0}^{L_{\mathbf{h}}-1} h(k)x(k)e^{-j2\pi nm/M}\right|^2 \\ &= \frac{1}{L_{\mathbf{h}}U}\sum_{n=0}^{L_{\mathbf{h}}-1}\sum_{l=0}^{L_{\mathbf{h}}-1} h(k)x(k)h^*(l)x^*(l)e^{-j2\pi(n-l)m/N}, \qquad (118)\end{aligned}$$

where

$$U = \frac{1}{L_{\mathbf{h}}}\sum_{n=0}^{L_{\mathbf{h}}-1}|h(n)|^2. \qquad (119)$$

Note that the number of observations equals the analysis filter length, $N = L_{\mathbf{h}}$.

The variance of the method may be evaluated for the special case of white Gaussian noise, as was shown for the Periodogram in [21]. The second order moment of the PSD estimate, according to Eq. (118), is

$$E\left\{\hat{P}_x(m_1)\hat{P}_x(m_2)\right\} = \frac{1}{L_{\mathbf{h}}^2 U^2}\sum_{p,q,r,s} E\left\{x(p)x^*(q)x(r)x^*(s)\right\} h(p)h^*(q)h(r)h^*(s)e^{-j2\pi[(p-q)\frac{m_1}{M}+(r-s)\frac{m_2}{M}]} \tag{120}$$

The moment factoring theorem for complex Gaussian random variables is

$$\begin{aligned} E\left\{x(p)x^*(q)x(r)x^*(s)\right\} &= E\left\{x(p)x^*(q)\right\}E\left\{x(r)x^*(s)\right\} + \ldots \\ &\quad \ldots + E\left\{x(p)x^*(s)\right\}E\left\{x(q)x^*(r)\right\}. \end{aligned} \tag{121}$$

and by substituting Eq. (121) into Eq. (120), the second-order moment becomes a sum of two terms. The first term contains the products of $E\left\{x(p)x^*(q)\right\}$ with $E\left\{x(r)x^*(s)\right\}$, which are equal to σ_x^4 when $p = q$ and $r = s$, and zero otherwise. The first term simplifies to σ_x^4 and the second term becomes

$$\frac{1}{L_{\mathbf{h}}^2 U^2}\sum_{p=0}^{L_{\mathbf{h}}-1}\sum_{q=0}^{L_{\mathbf{h}}-1}\sigma_x^4 h(p)h^*(q)e^{-j2\pi(p-q)(m_2-m_1)/M} = \sigma_x^4\frac{1}{L_{\mathbf{h}}^2 U^2}\left|H\left(m_2-m_1\right)\right|^2, \tag{122}$$

where

$$H(m) = \sum_{n=0}^{L_{\mathbf{h}}-1} h(n)e^{-j2\pi mn/M}. \tag{123}$$

Combining the terms, it follows that the second-order moment is

$$E\left\{\hat{P}_x(m_1)\hat{P}_x(m_2)\right\} = \sigma_x^4\left\{1 + \frac{1}{L_{\mathbf{h}}^2 U^2}\left|H\left(m_2-m_1\right)\right|^2\right\}. \tag{124}$$

Since

$$\mathrm{Cov}\left\{\hat{P}_x(m_1)\hat{P}_x(m_2)\right\} = E\left\{\hat{P}_x(m_1)\hat{P}_x(m_2)\right\} - E\left\{\hat{P}_x(m_1)\right\}E\left\{\hat{P}_x(m_2)\right\} \tag{125}$$

and $E\left\{\hat{P}_x(m)\right\} = \sigma_x^2$, the covariance of the modified periodogram is

$$\mathrm{Cov}\left\{\hat{P}_x(m_1)\hat{P}_x(m_2)\right\} = \frac{\sigma_x^4}{L_{\mathbf{h}}^2 U^2}\left|H\left(m_2-m_1\right)\right|^2. \tag{126}$$

Setting $m_1 = m_2$ the variance becomes

$$\mathrm{Var}\left\{\hat{P}_x(m)\right\} = \frac{\sigma_x^4}{L_{\mathbf{h}}^2 U^2}\left|\sum_{n=0}^{L_{\mathbf{h}}-1} h(n)\right|^2. \tag{127}$$

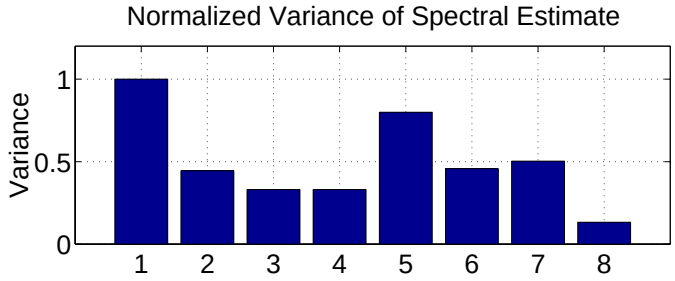

Figure 39: *Variance of Power Spectral Density estimates of WGN. The decimation factor is set to* $D = M = 32$ *in cases 5 and 7, and* $D = \frac{M}{2} = 16$ *in cases 6 and 8. The delay is set to* $\tau_H = \frac{L_{\mathbf{h}}-1}{2}$ *in cases 5 and 6, and* $\tau_H = \frac{M}{2}$ *in cases 7 and 8. The passband boundary frequency is set to* $\omega_p = \frac{\pi}{8M}$.

Fig. 39 shows the variance of PSD estimates using the filter bank types described in section 8. The number of subbands in this comparison is set to $M = 32$. The filter banks in case 5-8 are obtained from prototype analysis filters, designed according to Eq. (41), with filter length $L_{\mathbf{h}} = 2M = 64$. Obviously, the filter banks where the estimates are based on twice as many observations, with equal bandwidth, compared to the periodogram (case 1) have about half the variance. The spectral estimator in case 2 is based on the same amount of observations as in case 1, but has increased bandwith. This gives rise to the reduction of variance as can be seen in Fig. 39.

Time-adaptive algorithms, such as spectral subtraction, use averaging over time of spectral estimates. Since the sampling rate in the subbands of oversampled filter banks is higher than in the subbands of critically sampled filter banks, the variance is further reduced, depending on the decimation factor.

12 Conclusions

A two-step method for the design of modulated filter banks in subband filtering applications has been proposed. The method determines two low-pass prototype filters for the analysis filter bank and the synthesis filter bank, using unconstrained quadratic optimization. The delay of the analysis filter bank and the delay of the total filter bank may be controlled while the effects of inband-aliasing in the subband signals, and the residual aliasing in the output signal are minimized.

Filter banks with critical sampling and two times oversampling are designed using the suggested method, and they are compared with conventional filter bank types. Also filter banks with reduced delay have been included in the evaluation.

The filter banks have been evaluated in subband adaptive filtering applications, such as Acoustic Echo Cancellation and Speech Enhancement. It has been shown that the optimal filter banks, design with the proposed method can improve performance in these applications. Also the negative side effects, which are caused by the filter bank, such as aliasing and transmission delay, can be reduced without significant reduction of the performance.

Appendix

A Derivations

A.1 Derivation of matrix A

Matrix $\mathbf{A}$ is defined by

$$\mathbf{A} = \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} \phi_{\mathbf{h}}(e^{j\omega}) \phi_{\mathbf{h}}^H(e^{j\omega}) d\omega, \tag{128}$$

where $\phi_{\mathbf{h}}(z) = [1, z^{-1}, \ldots, z^{-L_{\mathbf{h}}+1}]^T$. A single matrix entry, $A_{m,n}$, is described by

$$\begin{aligned} A_{m,n} &= \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} e^{-j\omega m} e^{j\omega n} d\omega \\ &= \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} e^{j\omega(n-m)} d\omega \\ &= \frac{1}{2j\omega_p(n-m)} \left[e^{j\omega(n-m)} \right]_{-\omega_p}^{\omega_p} \\ &= \frac{1}{2j\omega_p(n-m)} \left[e^{j\omega_p(n-m)} - e^{-j\omega_p(n-m)} \right] \\ &= \frac{\sin(\omega_p(n-m))}{\omega_p(n-m)} \end{aligned} \tag{129}$$

A.2 Derivation of vector b

Vector $\mathbf{b}$ is defined by

$$\mathbf{b} = \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} \mathrm{Re}\left\{ e^{j\omega\tau_H} \phi_{\mathbf{h}}(e^{j\omega}) \right\} d\omega, \tag{130}$$

where $\phi_{\mathbf{h}}(z) = [1, z^{-1}, \ldots, z^{-L_{\mathbf{h}}+1}]^T$. A single vector entry, b_m, is described by

$$\begin{aligned} b_m &= \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} \mathrm{Re}\left\{ e^{j\omega\tau_H} e^{-j\omega m} \right\} d\omega \\ &= \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} \mathrm{Re}\left\{ e^{j\omega(\tau_H - m)} \right\} d\omega \\ &= \frac{1}{2\omega_p} \int_{-\omega_p}^{\omega_p} \cos(\omega(\tau_H - m)) d\omega \\ &= \frac{1}{2\omega_p(\tau_H - m)} \left[\sin(\omega(\tau_H - m)) \right]_{-\omega_p}^{\omega_p} \\ &= \frac{1}{2\omega_p(\tau_H - m)} \left[\sin(\omega_p(\tau_H - m)) - \sin(-\omega_p(\tau_H - m)) \right] \\ &= \frac{\sin(\omega_p(\tau_H - m))}{\omega_p(\tau_H - m)} \end{aligned} \tag{131}$$

A.3 Derivation of matrix C

Matrix $\mathbf{C}$ is defined by

$$\mathbf{C} = \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \int_{-\pi}^{\pi} \phi_{\mathbf{h}}(e^{j\omega/D} W_D^d) \phi_{\mathbf{h}}^H(e^{j\omega/D} W_D^d) d\omega, \tag{132}$$

where $\phi_{\mathbf{h}}(z) = [1, z^{-1}, \ldots, z^{-L_{\mathbf{h}}+1}]^T$. A single matrix entry, $C_{m,n}$, is described by

$$\begin{aligned} C_{m,n} &= \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \int_{-\pi}^{\pi} e^{-j\omega m/D} e^{j2\pi lm/D} e^{j\omega n/D} e^{-j2\pi ln/D} d\omega \\ &= \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} e^{j2\pi d(m-n)/D} \int_{-\pi}^{\pi} e^{j\omega(n-m)/D} d\omega \\ &= \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} e^{j2\pi d(m-n)/D} \frac{2D\sin(\pi(n-m)/D)}{(n-m)} \\ &= \frac{1}{\pi D} \sum_{d=1}^{D-1} e^{j2\pi d(m-n)/D} \frac{\sin(\pi(n-m)/D)}{(n-m)} \\ &= \frac{\sin(\pi(n-m)/D)}{\pi D(n-m)} \sum_{d=1}^{D-1} e^{j2\pi d(m-n)/D} \\ &= \frac{\sin(\pi(n-m)/D)}{\pi D(n-m)} \varphi(n-m), \end{aligned} \tag{133}$$

where

$$\varphi(n) = D \sum_{k=-\infty}^{\infty} \delta(n - kD) - 1 \tag{134}$$

A.4 Derivation of matrix E

Matrix $\mathbf{E}$ is defined by

$$\mathbf{E} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{\Psi}^H(e^{j\omega}) \mathbf{h}^* \mathbf{h}^T \mathbf{\Psi}(e^{j\omega}) d\omega, \tag{135}$$

where

$$\mathbf{\Psi}(z) = \frac{1}{D} \sum_{m=0}^{M-1} \phi_{\mathbf{h}}(zW_M^m) \phi_{\mathbf{g}}^T(zW_M^m). \tag{136}$$

Inserting the definition of $\mathbf{\Psi}(z)$ yields

$$\begin{aligned} \mathbf{E} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{1}{D} \sum_{m_1=0}^{M-1} \phi_{\mathbf{g}}^*(e^{j\omega} W_M^{m_1}) \phi_{\mathbf{h}}^H(e^{j\omega} W_M^{m_1}) \right] \mathbf{h}^* \mathbf{h}^T \left[\frac{1}{D} \sum_{m_2=0}^{M-1} \phi_{\mathbf{h}}(e^{j\omega} W_M^{m_2}) \phi_{\mathbf{g}}^T(e^{j\omega} W_M^{m_2}) \right] d\omega \\ &= \frac{1}{2\pi D^2} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \int_{-\pi}^{\pi} \phi_{\mathbf{g}}^*(e^{j\omega} W_M^{m_1}) \phi_{\mathbf{h}}^H(e^{j\omega} W_M^{m_1}) \mathbf{h}^* \mathbf{h}^T \phi_{\mathbf{h}}(e^{j\omega} W_M^{m_2}) \phi_{\mathbf{g}}^T(e^{j\omega} W_M^{m_2}) d\omega \end{aligned} \tag{137}$$

A single matrix entry, $E_{m,n}$, is described by

$$\begin{aligned} E_{p,q} &= \frac{1}{2\pi D^2} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} \sum_{i=0}^{L_{\mathbf{h}}-1} \int_{-\pi}^{\pi} \\ &\quad e^{j\omega p} e^{-j2\pi pm_1/M} e^{j\omega k} e^{-j2\pi km_1/M} h^*(k) h(i) e^{-j\omega i} e^{j2\pi im_2/M} e^{-j\omega q} e^{j2\pi qm_2/M} d\omega \\ &= \frac{1}{2\pi D^2} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} \sum_{i=0}^{L_{\mathbf{h}}-1} \\ &\quad e^{j2\pi(\frac{im_2}{M} + \frac{qm_2}{M} - \frac{pm_1}{M} - \frac{km_1}{M})} h^*(k) h(i) \int_{-\pi}^{\pi} e^{j\omega(p+k-i-q)} d\omega \end{aligned}$$

$$
\begin{aligned}
= & \frac{1}{2\pi D^2} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} \sum_{i=0}^{L_{\mathbf{h}}-1} \\
& e^{j2\pi(\frac{im_2}{M}+\frac{qm_2}{M}-\frac{pm_1}{M}-\frac{km_1}{M})} h^*(k)h(i) \frac{2\sin(\pi(p+k-i-q))}{(p+k-i-q)} \\
= & \frac{1}{\pi D^2} \sum_{k=0}^{L_{\mathbf{h}}-1} \sum_{i=0}^{L_{\mathbf{h}}-1} \frac{h^*(k)h(i)\sin(\pi(p+k-i-q))}{(p+k-i-q)} \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} e^{j2\pi(\frac{im_2}{M}+\frac{qm_2}{M}-\frac{pm_1}{M}-\frac{km_1}{M})}
\end{aligned}
\tag{138}
$$

The sum of exponentials is

$$
\sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} e^{j2\pi(\frac{im_2}{M}+\frac{qm_2}{M}-\frac{pm_1}{M}-\frac{km_1}{M})} = \begin{cases} M^2 & i+q=\kappa M,\ p+k=\kappa M,\ \kappa \in \mathbf{Z} \\ 0 & \text{otherwise} \end{cases} \tag{139}
$$

With

$$
\frac{\sin(\pi(p+k-i-q))}{(p+k-i-q)} = \begin{cases} \pi & p+k-i-q=0 \\ 0 & \text{otherwise} \end{cases} \tag{140}
$$

the final expression for a single matrix entry $E_{m,n}$ is

$$
E_{p,q} = \frac{M^2}{D^2} \sum_{\kappa=-\infty}^{\infty} h^*(\kappa M - p)h(\kappa M - q) \tag{141}
$$

A.5 Derivation of vector f

Vector $\mathbf{f}$ is defined by

$$
\mathbf{f} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ e^{j\omega\tau_T} \mathbf{\Psi}^T(e^{j\omega})\mathbf{h} \right\} d\omega \tag{142}
$$

Inserting the definition of $\mathbf{\Psi}(z)$ yields

$$
\mathbf{f} = \frac{1}{2\pi D} \sum_{m=0}^{M-1} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ e^{j\omega\tau_T} \phi_{\mathbf{g}}(e^{j\omega}W_M^m)\phi_{\mathbf{h}}^T(e^{j\omega}W_M^m) \right\} \mathbf{h} d\omega \tag{143}
$$

A single vector entry, f_p, is described by

$$
\begin{aligned}
f_p = & \frac{1}{2\pi D} \sum_{m=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} h(k) \int_{-\pi}^{\pi} \mathrm{Re}\left\{ e^{j\omega\tau_T} e^{-j\omega p} e^{j2\pi\frac{pm}{M}} e^{-j\omega k} e^{j2\pi\frac{km}{M}} \right\} d\omega \\
= & \frac{1}{2\pi D} \sum_{m=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} h(k) \int_{-\pi}^{\pi} \mathrm{Re}\left\{ e^{j\omega(\tau_T-p-k)} e^{-j2\pi(\frac{pm}{M}+\frac{km}{M})} \right\} d\omega \\
= & \frac{1}{2\pi D} \sum_{m=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} h(k)\cos(2\pi\left[\frac{pm}{M}+\frac{km}{M}\right]) \int_{-\pi}^{\pi} \cos(\omega\left[\tau_T-p-k\right]) d\omega \\
= & \frac{1}{2\pi D} \sum_{m=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} h(k)\cos(2\pi\left[\frac{pm}{M}+\frac{km}{M}\right]) \frac{2\sin(\pi(\tau_T-p-k))}{(\tau_T-p-k)} \\
= & \frac{1}{\pi D} \sum_{k=0}^{L_{\mathbf{h}}-1} \frac{h(k)\sin(\pi(\tau_T-p-k))}{(\tau_T-p-k)} \sum_{m=0}^{M-1} \cos(2\pi\left[\frac{pm}{M}+\frac{km}{M}\right]) \\
= & \frac{\lambda}{\pi D} h(\tau_T - p),
\end{aligned}
\tag{144}
$$

where

$$
\lambda = \sum_{m=0}^{M-1} \cos(2\pi\tau_T m/M) \tag{145}
$$

A.6 Derivation of matrix P

Matrix $\mathbf{P}$ is defined by

$$\mathbf{P} = \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \int_{-\pi}^{\pi} \mathbf{\Phi}_{m,d}^H(e^{j\omega}) \mathbf{h}^* \mathbf{h}^T \mathbf{\Phi}_{m,d}(e^{j\omega}) \, d\omega \tag{146}$$

Inserting the definition $\mathbf{\Phi}_{m,d}(z) = \phi_{\mathbf{h}}(zW_M^m W_D^d)\phi_{\mathbf{g}}^T(zW_M^m)$ yields

$$\begin{aligned}
\mathbf{P} &= \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \int_{-\pi}^{\pi} \left[\phi_{\mathbf{h}}(e^{j\omega} W_M^m W_D^d)\phi_{\mathbf{g}}^T(e^{j\omega} W_M^m)\right]^H \mathbf{h}^* \mathbf{h}^T \left[\phi_{\mathbf{h}}(e^{j\omega} W_M^m W_D^d)\phi_{\mathbf{g}}^T(e^{j\omega} W_M^m)\right] d\omega \\
&= \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \int_{-\pi}^{\pi} \phi_{\mathbf{g}}^*(e^{j\omega} W_M^m)\phi_{\mathbf{h}}^H(e^{j\omega} W_M^m W_D^d)\mathbf{h}^* \mathbf{h}^T \phi_{\mathbf{h}}(e^{j\omega} W_M^m W_D^d)\phi_{\mathbf{g}}^T(e^{j\omega} W_M^m) d\omega
\end{aligned} \tag{147}$$

A single matrix entry, $P_{p,q}$, is described by

$$\begin{aligned}
P_{p,q} &= \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} \sum_{i=0}^{L_{\mathbf{h}}-1} \int_{-\pi}^{\pi} \\
&\qquad e^{j\omega p} e^{-j2\pi pm/M} e^{j\omega k} e^{-j2\pi km/M} e^{-j2\pi kd/D} h^*(k) h(i) e^{-j\omega i} e^{j2\pi im/M} e^{j2\pi id/D} e^{-j\omega q} e^{j2\pi qm/M} d\omega \\
&= \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} \sum_{i=0}^{L_{\mathbf{h}}-1} e^{j2\pi(\frac{im}{M} + \frac{id}{D} + \frac{qm}{M} - \frac{pm}{M} - \frac{km}{M} - \frac{kd}{D})} h^*(k) h(i) \int_{-\pi}^{\pi} e^{j\omega(p+k-i-q)} d\omega \\
&= \frac{1}{2\pi D^2} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \sum_{k=0}^{L_{\mathbf{h}}-1} \sum_{i=0}^{L_{\mathbf{h}}-1} e^{j2\pi(\frac{im}{M} + \frac{id}{D} + \frac{qm}{M} - \frac{pm}{M} - \frac{km}{M} - \frac{kd}{D})} h^*(k) h(i) \frac{2\sin(\pi(p+k-i-q))}{(p+k-i-q)} \\
&= \frac{1}{\pi D^2} \sum_{k=0}^{L_{\mathbf{h}}-1} \sum_{i=0}^{L_{\mathbf{h}}-1} \frac{h^*(k) h(i) \sin(\pi(p+k-i-q))}{(p+k-i-q)} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} e^{j2\pi(\frac{im}{M} + \frac{id}{D} + \frac{qm}{M} - \frac{pm}{M} - \frac{km}{M} - \frac{kd}{D})}
\end{aligned} \tag{148}$$

The sum of complex exponentials can be rewritten according to

$$\sum_{d=1}^{D-1} \sum_{m=0}^{M-1} e^{j2\pi(\frac{im}{M} + \frac{id}{D} + \frac{qm}{M} - \frac{pm}{M} - \frac{km}{M} - \frac{kd}{D})} = \sum_{d=1}^{D-1} e^{j2\pi d(i-k)/D} \sum_{m=0}^{M-1} e^{j2\pi m(i+q-p-k)/M}, \tag{149}$$

of which the first part is

$$\sum_{d=1}^{D-1} e^{j2\pi d(i-k)/D} = \begin{cases} D-1, & i-k = \kappa D, \ \kappa \in \mathbf{Z} \\ -1, & \text{otherwise} \end{cases} \tag{150}$$

The second sum in the right part of Eq. (149) is

$$\sum_{m=0}^{M-1} e^{j2\pi m(i+q-p-k)/M} = \begin{cases} M, & i+q-p-k = \kappa M, \ \kappa \in \mathbf{Z} \\ 0, & \text{otherwise} \end{cases} \tag{151}$$

With

$$\frac{\sin(\pi(p+k-i-q))}{(p+k-i-q)} = \begin{cases} \pi & p+k-i-q = 0 \\ 0 & \text{otherwise} \end{cases} \tag{152}$$

the final expression for the single matrix entry $P_{p,q}$ is

$$P_{p,q} = \frac{M}{D^2} \sum_{l=-\infty}^{\infty} h^*(l+q) h(l+p) \varphi(p-q), \tag{153}$$

where

$$\varphi(n) = D \sum_{k=-\infty}^{\infty} \delta(n - kD) - 1 \tag{154}$$

B Examples

On the following pages, the impulse response and the frequency response of the prototype analysis filters and the prototype synthesis filters corresponding to the following filter bank types (see section 8) are shown:

All filter banks in this appendix have $M = 8$ subbands. The following six spectral properties of the filter banks are plotted.

- Amplitude Error - The amplitude error function, see Eq. (44)
$$\left||T(e^{j\omega})| - |D(e^{j\omega})|\right|^2$$
- Group Delay - The group delay of the total filter bank structure
$$\tau(\omega) = -\frac{d}{d\omega}\angle T(e^{j\omega})$$
- Phase Error - The phase error function
$$\left|\angle T(e^{j\omega}) - \angle D(e^{j\omega})\right|^2$$
- Inband-Aliasing - The inband-aliasing spectrum, see Eq. (10)
$$\frac{1}{D}\sum_{d=0}^{D-1} H(e^{j\omega/D}W_M^m W_D^d)$$
- Residual Aliasing - The residual aliasing spectrum, see Eq. (15)
$$\left|\sum_{d=0}^{D-1}\sum_{m=0}^{M-1} A_{m,d}(e^{j\omega})\right|^2$$
- Residual Aliasing Distortion - The residual aliasing distortion, see Eq. (53)
$$\sum_{d=0}^{D-1}\sum_{m=0}^{M-1}\left|A_{m,d}(e^{j\omega})\right|^2$$

B.1 FFT Filter Bank

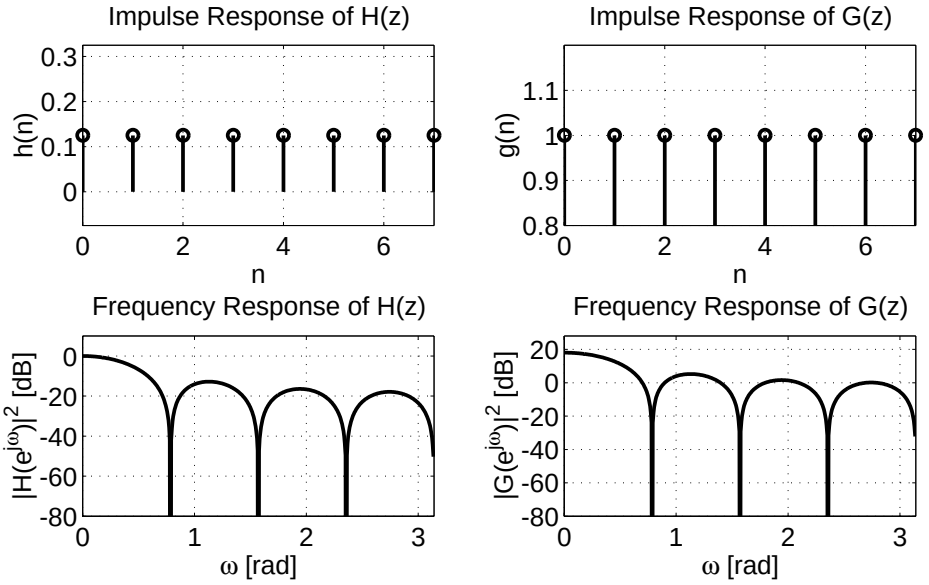

Figure 40: *Impulse response and frequency response of the prototype analysis filter and prototype synthesis filter.*

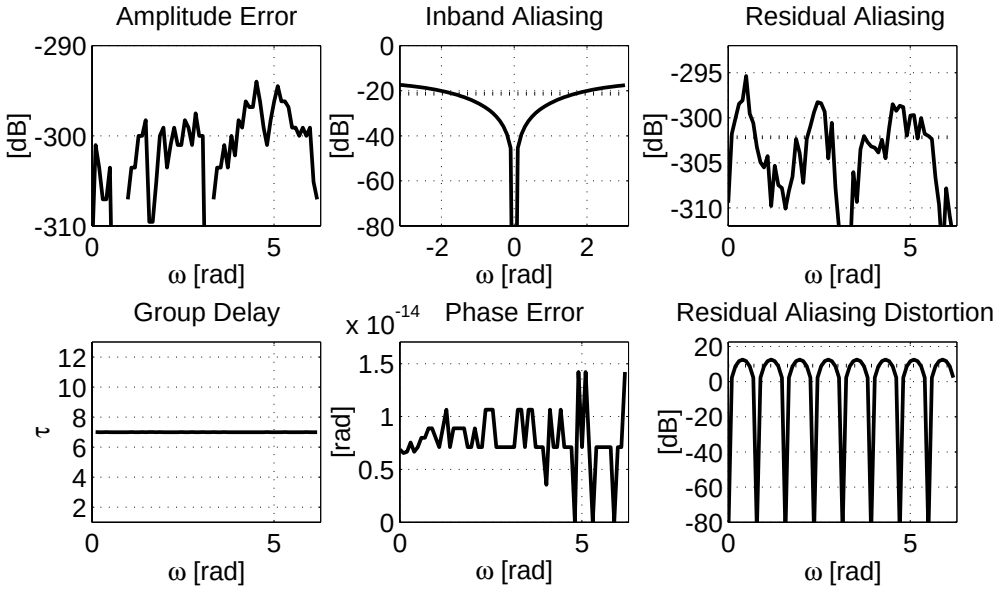

Figure 41: *Spectral properties of the total filter bank. The properties are described on page 75.*

B.2 Hanning Filter Bank

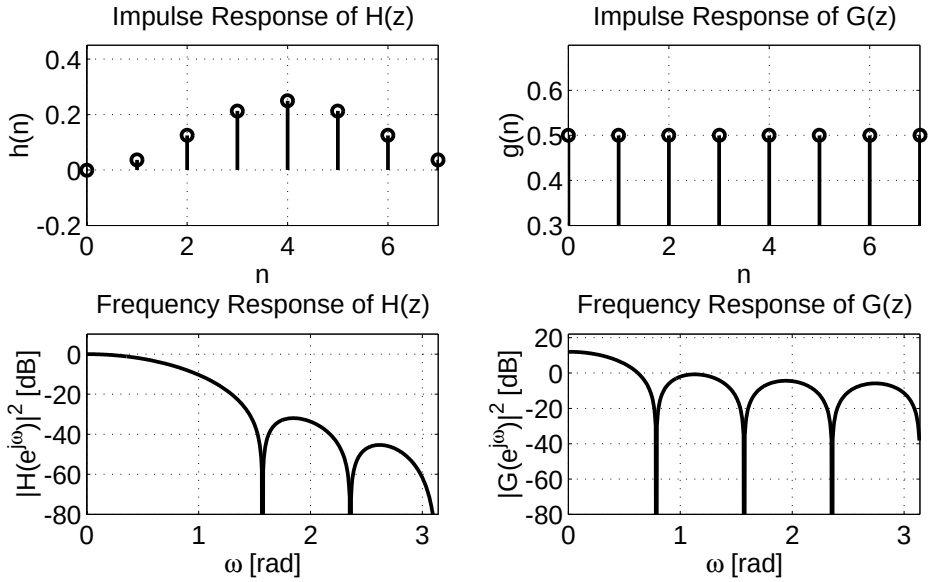

Figure 42: *Impulse response and frequency response of the prototype analysis filter and prototype synthesis filter.*

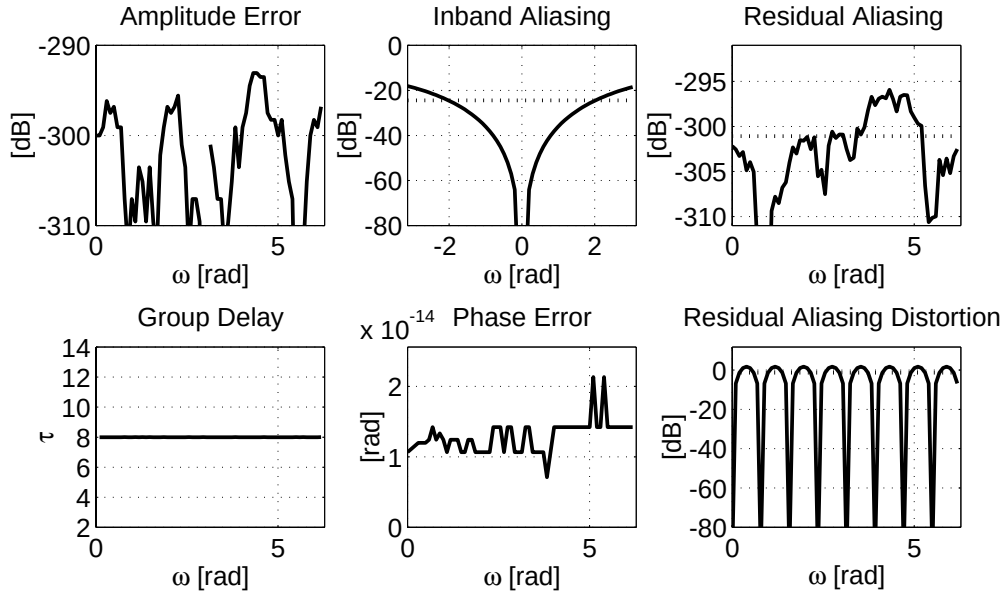

Figure 43: *Spectral properties of the total filter bank. The properties are described on page 75.*

B.3 Critically Sampled Filter Bank - Window Design Method

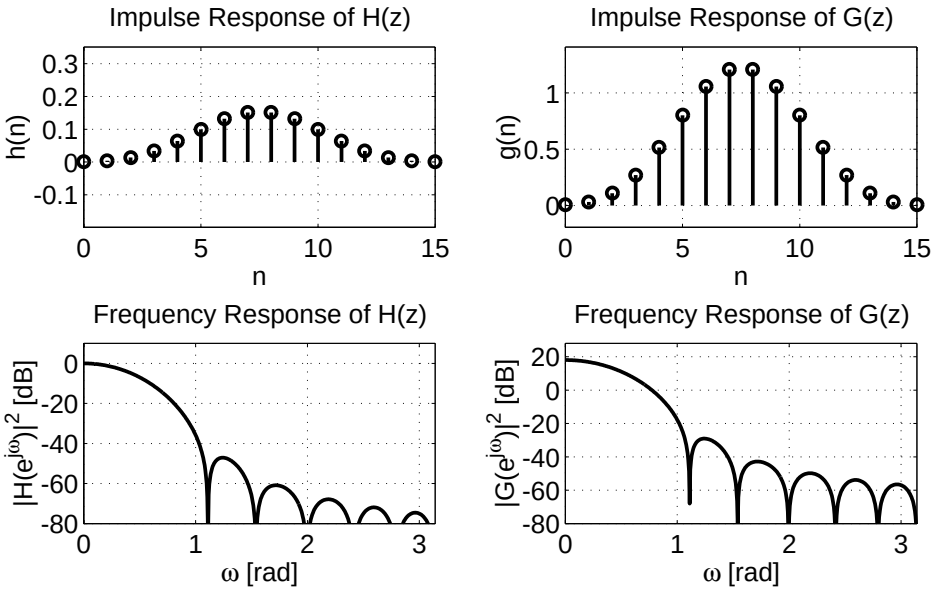

Figure 44: *Impulse response and frequency response of the prototype analysis filter and prototype synthesis filter.*

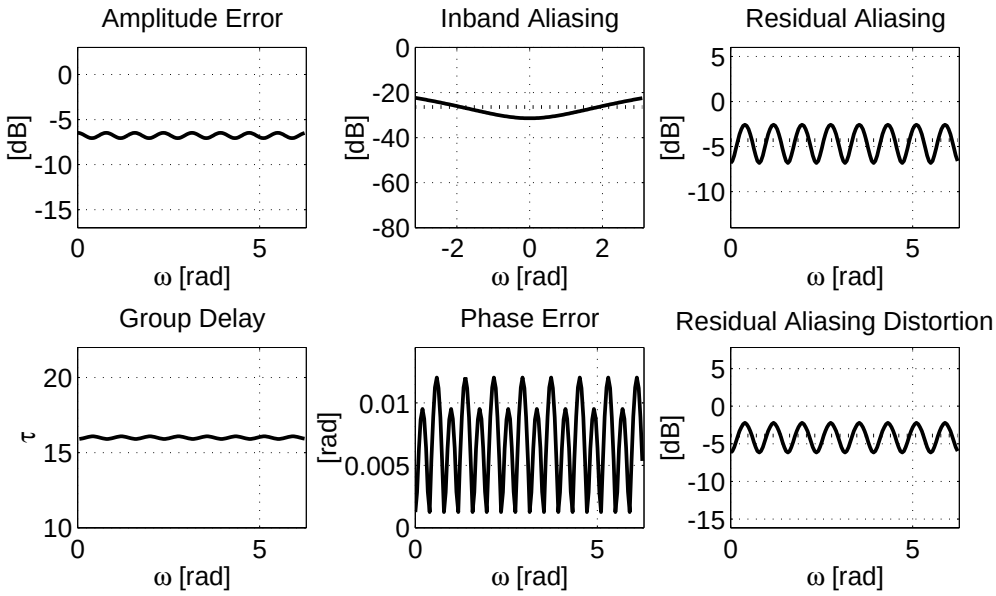

Figure 45: *Spectral properties of the total filter bank. The properties are described on page 75.*

B.4 Oversampled Filter Bank - Window Design Method

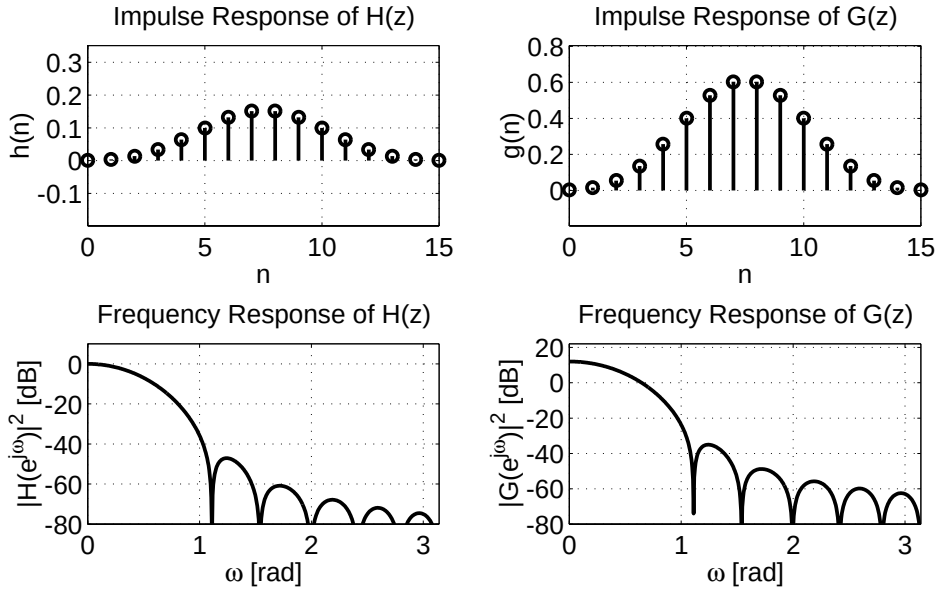

Figure 46: *Impulse response and frequency response of the prototype analysis filter and prototype synthesis filter.*

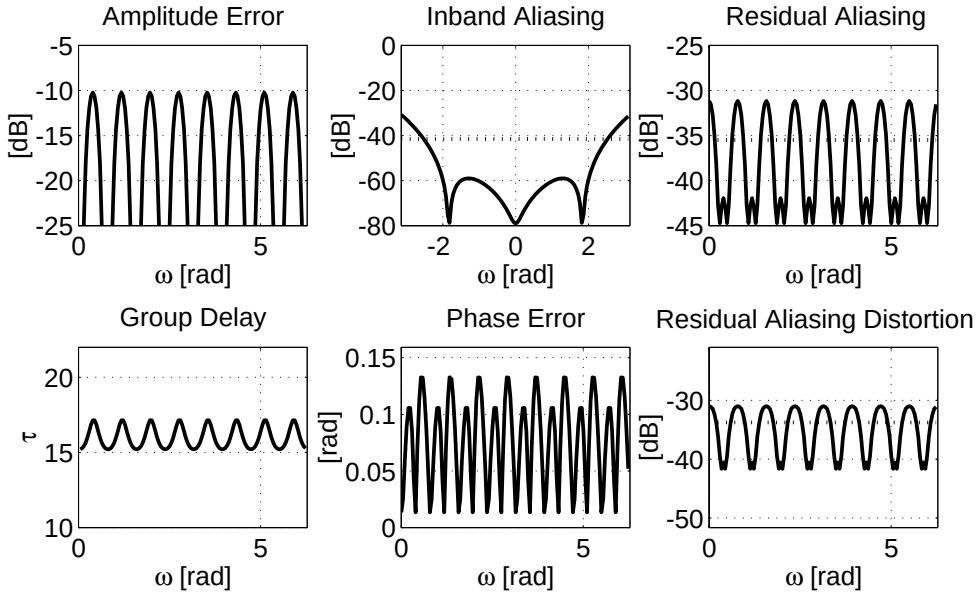

Figure 47: *Spectral properties of the total filter bank. The properties are described on page 75.*

B.5 Critically Sampled Filter Bank - Proposed Method

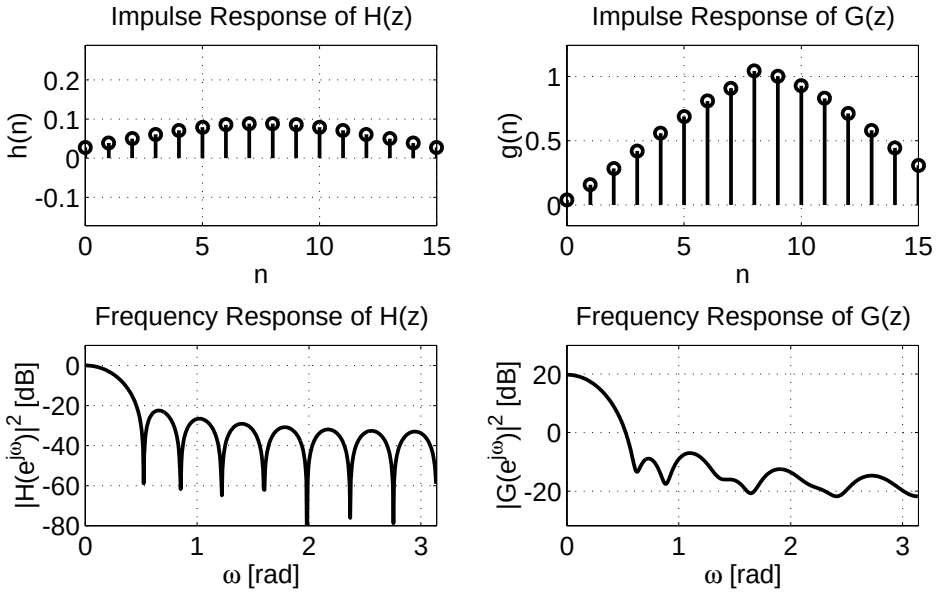

Figure 48: *Impulse response and frequency response of the prototype analysis filter and prototype synthesis filter.*

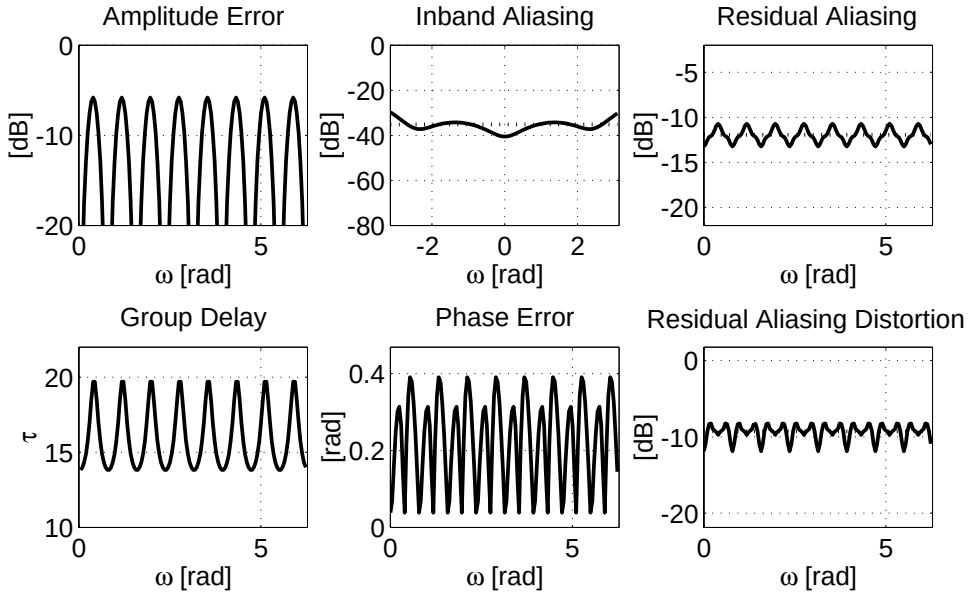

Figure 49: *Spectral properties of the total filter bank. The properties are described on page 75.*

B.6 Oversampled Filter Bank - Proposed Method

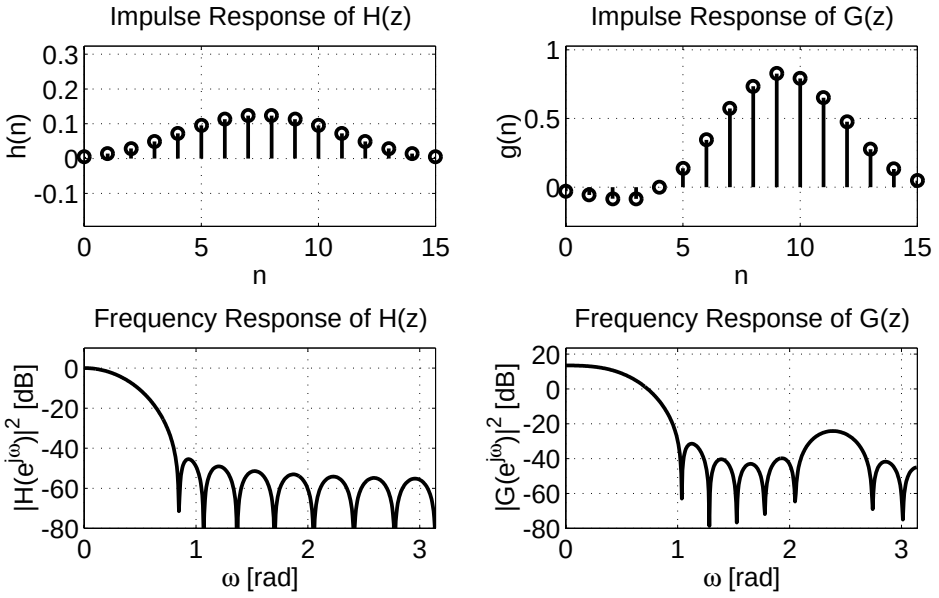

Figure 50: *Impulse response and frequency response of the prototype analysis filter and prototype synthesis filter.*

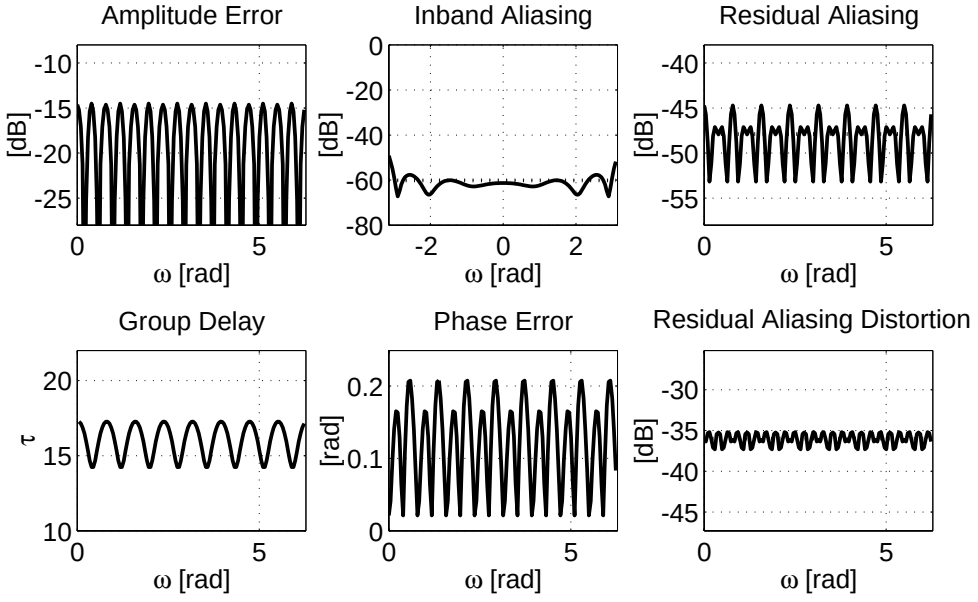

Figure 51: *Spectral properties of the total filter bank. The properties are described on page 75.*

B.7 Critically Sampled Filter Bank with less delay - Proposed Method

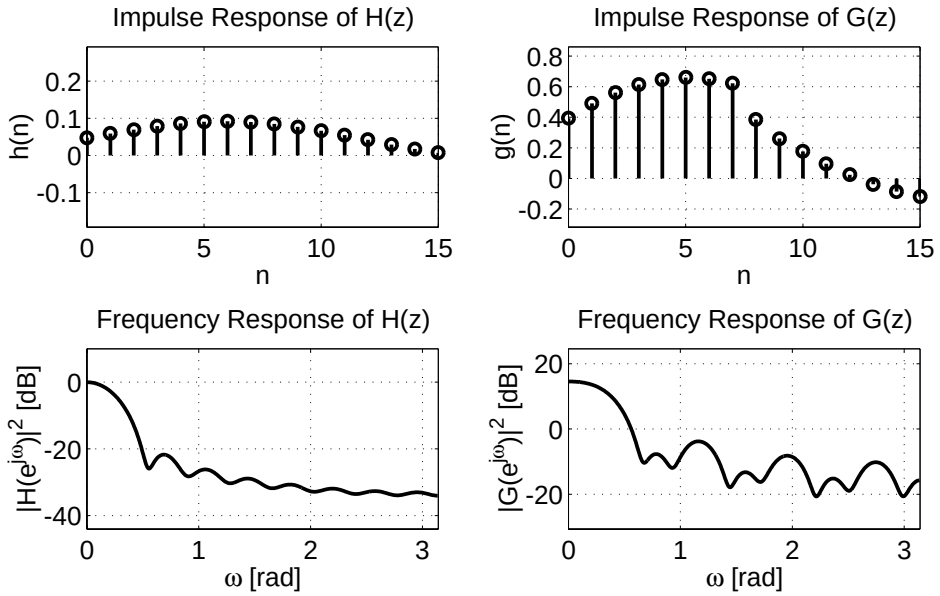

Figure 52: *Impulse response and frequency response of the prototype analysis filter and prototype synthesis filter.*

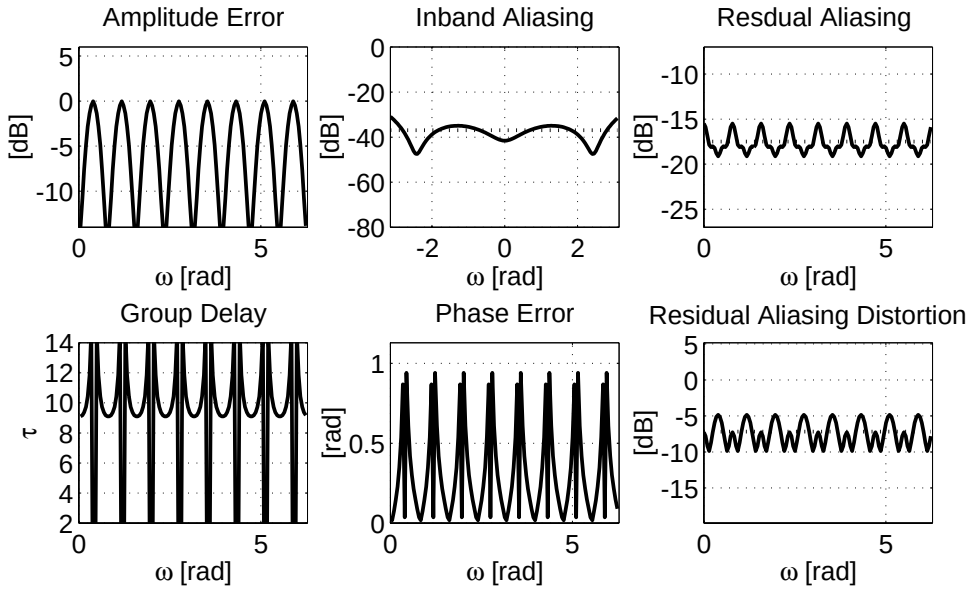

Figure 53: *Spectral properties of the total filter bank. The properties are described on page 75.*

B.8 Oversampled Filter Bank with less delay - Proposed Method

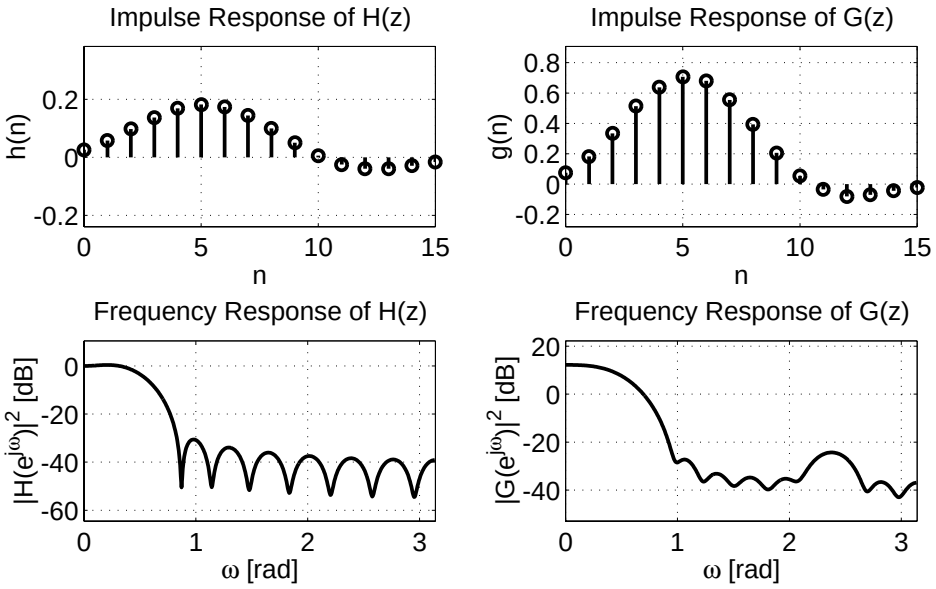

Figure 54: *Impulse response and frequency response of the prototype analysis filter and prototype synthesis filter.*

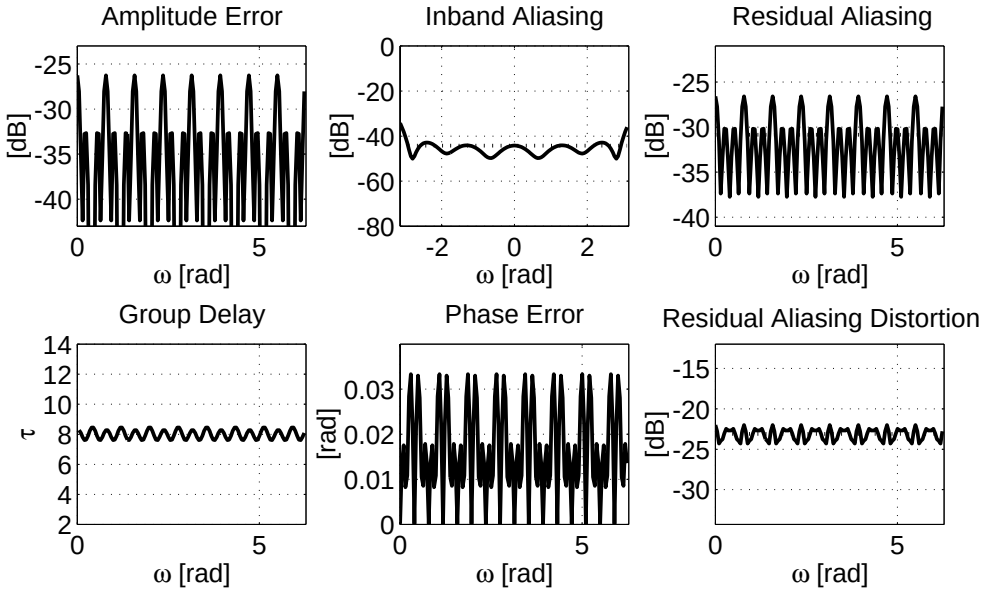

Figure 55: *Spectral properties of the total filter bank. The properties are described on page 75.*

References

[1] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands," in Proc. ICASSP'88, pp. 1572-1575, 1988.

[2] W. Kellermann, "Analysis and design of multirate systems for cancellation of acoustic echoes," in Proc. ICASSP'88, pp. 2570-2573, 1988.

[3] A. Gilloire, M. Vetterli, "Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation," IEEE Transactions on Signal Processing, pp. 1862-1875, vol. 40, issue 8 , Aug. 1992.

[4] P. A. Naylor, O. Tanrikulu A. G. Constantinidis, "Subband Adaptive Filtering for Acoustic Echo Control using Allpass Polyphase IIR filter banks," IEEE Trans. on Speech and Audio Processing, vol. 6, no. 2, Mar. 1998.

[5] S. L. Gay, J. Benesty (Editors), "Acoustic Signal Processing for Telecommunication," Kluwer Academic Publishers, 2000.

[6] T. Gülzow, A. Engelsberg, U. Heute, "Comparison of a discrete wavelet transformation and a nonuniform polyphase filter bank applied to spectral subtraction speech enhancement," Signal Processing 64, pp. 5-19, 1998.

[7] N. Grbić, "Speech Signal Extraction - A Multichannel Approach," University of Karlskrona/Ronneby, pp. 72-75, Nov. 1999, ISBN 91-630-8841-X.

[8] J. M. de Haan, N. Grbić, I. Claesson, S. Nordholm, "Design of Oversampled Uniform DFT Filter Banks with Delay Specification using Quadratic Optimization," in Proc. ICASSP, May 2001.

[9] Q.-G. Liu, B. Champagne, D. K. C. Ho, "Simple design of oversampled uniform DFT filter banks with applications to subband acoustic echo cancellation," Signal Processing 80, pp. 831-847, 2000.

[10] M. Harteneck, S. Weiss, R. W. Stewart, "Design of Near Perfect Reconstruction Oversampled Filter Banks for Subband Adaptive Filters," IEEE Trans. on Circuits and Systems, vol. 46, no. 8, pp. 1081-1085, Aug 1999.

[11] G. D. T. Schuller, "Modulated Filter Banks with Arbitrary System Delay: Efficient Implementations and the Time-Varying Case," IEEE Trans. on Signal Processing, vol. 48, no. 3, pp. 737-748, Mar. 2000.

[12] J. Kliewer, "Simplified Design of Linear-Phase Prototype Filters for Modulated Filter Banks," European Signal Conference, pp. 1191-1194, Sept. 1996.

[13] J. Kliewer, A. Mertins, "Design of Paraunitary Oversampled Cosine-Modulated Filter Banks," IEEE Proc. Acoustics, Speech and Signal Processing, pp. 2073-2076, Apr. 1997.

[14] P. N. Heller, T. Karp, T. Q. Nguyen, "A General Formulation of Modulated Filter Banks," IEEE Trans. on Signal Processing, vol. 47, no. 4, Apr. 1999.

[15] R. E. Chrochiere, L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall, 1983

[16] P. P. Vaidyanathan, *Multirate Systems and filter banks*, Prentice-Hall, 1993.

[17] N. J. Fliege, *Multirate Digital Signal Processing*, Wiley, 1994.

[18] T. W. Parks, C. S. Burrus, *Digital Filter Design* , John Wiley & Sons, Inc. 1987.

[19] D. R. Xiangkun Chen, T. W. Parks, "Design of FIR filters in the Complex Domain," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-35, no. 2, Feb. 1979.

[20] P. Stoica, R. Moses, *Introduction to Spectral Analysis*, Prentice Hall, 1997, ISBN 0-13-258419-0.

[21] M. H. Hayes, *Statistical Digital Signal processing and Modeling*, John Wiley & Sons, 1996.

[22] S. Haykin, *Adaptive Filter Theory*, Prentice Hall Int. Inc., 1996, ISBN 0-13-397985-7.

Part II

Design of Oversampled Uniform DFT Filter Banks with Delay Specification using Quadratic Optimization

Part II is published as:

J. M. de Haan, N. Grbić, I. Claesson and S. Nordholm, "Design of Oversampled Uniform DFT Filter Banks with Delay Specification using Quadratic Optimization," accepted for publication at the International Conference on Acoustics, Speech, and Signal Processing, ICASSP, Salt Lake City, USA, May 2001.

Design of Oversampled Uniform DFT Filter Banks with Delay Specification using Quadratic Optimization

J. M. de Haan, N. Grbić, I. Claesson, S. Nordholm*

Department of Telecommunications and Signal Processing
Blekinge Institute of Technology
Ronneby, Sweden

*Australian Telecommunications Research Institute
Curtin University
Perth, Australia

Abstract

Subband adaptive filters have been proposed to avoid the drawbacks of slow convergence and high computational complexity associated with time domain adaptive filters. Subband processing introduces transmission delays caused by the filter bank and signal degradations due to aliasing effects. One efficient way to reduce the aliasing effects is to allow a higher sample rate than critically needed in the subbands and thus reduce subband signal degradation. We suggest a design method, for uniform DFT filter banks with any oversampling factor, where the total filter bank group delay may be specified, and where the aliasing and magnitude/phase distortions are minimized.

1 Introduction

Subband adaptive filtering has arised as an alternative for conventional time domain adaptive filtering, [1]. The main reason is the reduction in computational complexity and the increase in convergence speed for the adaptive algorithm, achieved by dividing the algorithm into subbands, [2]. The computational savings comes from the fact that time domain convolution becomes decoupled in the subbands, at a lower sample rate, [3].

Subband analysis and synthesis is often performed using multirate filter banks, [4]. Non-ideal filters in the filter bank cause aliasing of the subband signals. This aliasing can be cancelled in the synthesis bank when certain conditions are met by the synthesis filters and in the subband processing. However, even if aliasing distortion in the filter bank output is cancelled in this way, the inband aliasing is still present in the subband adaptive filter input signals and, consequently, the adaptive filters are perturbed and the overall performance of the system is reduced, [5].

Several solutions to the subband filtering problem have been suggested in literature. Non-critical decimation has been suggested in [1], where filter bank delay aspects, and amplitude distortions, have not especially been taken into consideration. The use of cross filters, [5], has been suggested to explicitly filter out the aliasing components. A delayless structure has been proposed in [6], where the actual filtering is performed in the time domain, with consequences of higher computational complexity. The computational complexity also increases significantly with cross band filters.

We use a uniform DFT-modulated FIR filter bank for the subband transformations. Modulated filter banks provide a computationally efficient implementation, due to the polyphase implementation [4], and great design simplicity. The main contribution in this paper is the suggested design method, where the filter bank response error and the inband and output aliasing errors are minimized simultaneously, while the total filter bank group-delay is pre-specified. The influence of the filter bank performance is evaluated on an RLS subband beamformer [7].

2 The Uniform DFT Modulated filter bank

In this section we will derive an input-output expression for analysis-synthesis DFT filter banks with arbitrary decimation factor. Two sets of M filters, $H_m(z)$ and $G_m(z)$, form a uniform DFT analysis filter bank and synthesis filter bank, respectively, when they are related to prototype filters, $H(z)$ and $G(z)$, as

$$\begin{aligned} H_m(z) &= H(zW_M^m) = \mathbf{h}^T\phi(zW_M^m) \\ G_m(z) &= G(zW_M^m) = \mathbf{g}^T\phi(zW_M^m) \\ &\quad \text{for} \;\; m = 0, \ldots, M-1 \end{aligned} \tag{1}$$

where $W_M = e^{-j2\pi/M}$, $\mathbf{h} = [h(0), \ldots, h(L_h - 1)]^T$, $\mathbf{g} = [g(0), \ldots, g(L_g - 1)]^T$ and $\phi(z) = [1, z^{-1}, \ldots, z^{-(L-1)}]^T$. Each subband signal is decimated by a factor D. An efficient implementation of such a filter bank is given in [8]. For simplicity of derivation we study the direct form realization of the filter banks given in Fig. 1. The input signal $X(z)$ is filtered by the analysis filters $H_m(z)$ and decimated by

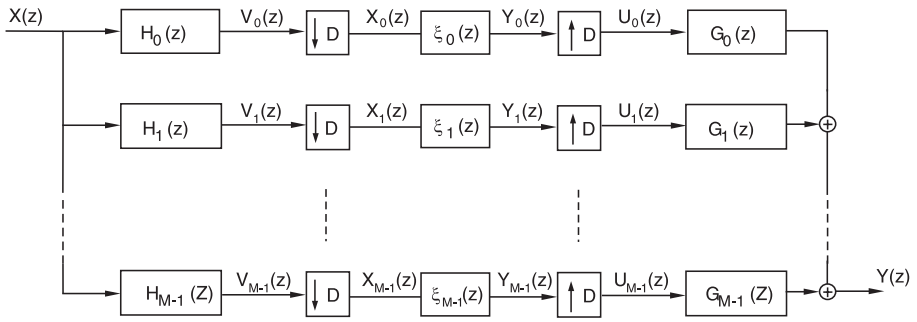

Figure 1: *Direct form realization for analysis and synthesis filter banks.*

the factor D according to

$$\begin{aligned} X_m(z) &= \frac{1}{D}\sum_{d=0}^{D-1} H(z^{\frac{1}{D}}W_M^m W_D^d)X(z^{\frac{1}{D}}W_D^d) \\ & \quad m = 0, \cdots, M-1 \end{aligned} \tag{2}$$

where $W_D = e^{-j2\pi/D}$. In the synthesis filter bank, the subband signals Y_m are interpolated by the interpolation factor D, filtered by the synthesis filters $G_m(z)$ and then added together to form the output signal

$$Y(z) = \frac{1}{D}\sum_{d=0}^{D-1} X(zW_D^d)\sum_{m=0}^{M-1}\xi_m(z^D)H(zW_M^m W_D^d)G(zW_M^m). \tag{3}$$

Here $\xi_m(z)$ is the application dependent filtering operation in subband no. m.

3 Analysis Filter Bank Design

Since the analysis filters are related to a single prototype analysis filter according to Eq. (1), the analysis filter bank design problem reduces to the design of a single filter. The ideal prototype analysis filter is a low pass filter with cut-off frequency $\omega_p = \pi/M$. Since FIR filters are not ideally frequency selective, approximations need to be made. The analysis prototype filter will be designed by defining a pass-band region in which the filter response should be flat, while minimizing the inband aliasing error. By minimizing the inband aliasing, the prototype filter will obtain low-pass characteristics since this is similar to maximizing stop band attenuation.

The passband response error in the passband region $\Omega_p = [-\omega_p, \omega_p]$ is defined as

$$\varepsilon_P(\mathbf{h}) = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega \tag{4}$$

where $H_d(z)$ is the desired frequency response. The desired frequency response is

$$H_d(e^{j\omega}) = e^{-j\omega\tau_H} \qquad \omega \in \Omega_p \tag{5}$$

where τ_H is the desired group delay of the analysis prototype filter and with that the desired delay of the analysis filter bank. With the analysis prototype filter written in terms of its impulse response, $H(z) = \mathbf{h}^T\phi(z)$, we substitute Eq. (5) into Eq. (4)

$$\varepsilon_P(\mathbf{h}) = \mathbf{h}^T\mathbf{A}\mathbf{h} - 2\mathbf{h}^T\mathbf{b} + 1 \tag{6}$$

where

$$\mathbf{A} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \phi(e^{j\omega})\phi^H(e^{j\omega})d\omega \tag{7}$$

and

$$\mathbf{b} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \mathrm{Re}\left\{e^{j\omega\tau_H}\phi(e^{j\omega})\right\}d\omega. \tag{8}$$

In order to minimize the inband aliasing we will define the *Inband Aliasing Error* and express it in terms of the analysis prototype filter. From Eq. (2), the sum of the inband aliasing terms $\mathcal{X}_m(z)$ in subband signal $X_m(z)$ is described by

$$\mathcal{X}_m(z) = \frac{1}{D}\sum_{d=1}^{D-1} H(z^{\frac{1}{D}}W_M^m W_D^d)X(z^{\frac{1}{D}}W_D^d) \tag{9}$$

for $m = 0, \cdots, M-1$. In the ideal case with the ideal prototype filter, and thus with zero aliasing in the subband signals $X_m(z)$, the frequency response part in each term is zero and thus $\mathcal{X}_m(z) = 0$. In the non-ideal case with FIR filters we would like to minimize the energy in each term. Since the analysis filters are related by Eq. (1) it is sufficient to minimize the energy in the aliasing terms of the first subband ($m = 0$). The sum of the power magnitudes of the aliasing terms in the first subband is

$$D_0(e^{j\omega}) = \frac{1}{D}\sum_{d=1}^{D-1} |H(e^{j\omega/D}W_D^d)|^2. \tag{10}$$

Subsequently, the inband aliasing error expressed in terms of the impulse response of the prototype analysis filter is

$$\varepsilon_{D_0}(\mathbf{h}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_0(e^{j\omega})d\omega = \mathbf{h}^T\mathbf{C}\mathbf{h} \tag{11}$$

where the hermitian matrix $\mathbf{C}$ is

$$\mathbf{C} = \frac{1}{2\pi D}\sum_{d=1}^{D-1}\int_{-\pi}^{\pi} \phi(e^{j\omega/D}W_D^d)\phi^H(e^{j\omega/D}W_D^d)d\omega. \tag{12}$$

The optimal analysis prototype filter with respect to minimal passband response error and minimal energy in the aliasing components is found by minimizing the

sum of the passband response error in Eq. (6) and the inband aliasing error in Eq. (11)

$$\begin{aligned}\varepsilon_{tot}(\mathbf{h}) &= \varepsilon_P(\mathbf{h}) + \varepsilon_{D_0}(\mathbf{h}) \\ &= \mathbf{h}^T(\mathbf{A}+\mathbf{C})\mathbf{h} - 2\mathbf{h}^T\mathbf{b} + 1 \end{aligned} \tag{13}$$

that is, by solving the set of linear equations

$$\mathbf{h} = \arg\min_{\mathbf{h}} \varepsilon_{tot}(\mathbf{h}) = (\mathbf{A}+\mathbf{C})^{-1}\mathbf{b}. \tag{14}$$

4 Synthesis Filter Bank Design

Given the analysis filter bank with an analysis prototype filter $H(z)$ designed as described in Section 3, we will design an optimal synthesis filter bank which minimizes the amplitude and phase distortion of the total filter bank system and also minimizes the output aliasing distortion.

The first part in the design of the synthesis filter bank is the minimization of amplitude and phase distortion. We will derive the *System Response Error* expressed in terms of the impulse reponses of the prototype filters $H(z) = \mathbf{h}^T\phi(z)$ and $G(z) = \mathbf{g}^T\phi(z)$. The system response error is defined by

$$\varepsilon_T = \frac{1}{2\pi}\int_{-\pi}^{\pi} |T(e^{j\omega}) - T_d(e^{j\omega})|^2 d\omega. \tag{15}$$

From Eq. (3), we can express the total filter bank system response in terms of $\mathbf{h}$ and $\mathbf{g}$, with $\xi_m(z) = 1, m = 0, \ldots, M-1$

$$T(z) = \frac{1}{D}\sum_{d=0}^{D-1}\sum_{m=0}^{M-1} H(zW_M^m W_D^d)G(zW_M^m) = \mathbf{h}^T\mathbf{\Psi}(z)\mathbf{g} \tag{16}$$

where

$$\mathbf{\Psi}(z) = \frac{1}{D}\sum_{d=0}^{D-1}\sum_{m=0}^{M-1} \phi(zW_M^m W_D^d)\phi^T(zW_M^m). \tag{17}$$

The desired filter bank response is

$$T_d(e^{j\omega}) = e^{-j\omega\tau_T} \tag{18}$$

where τ_T is the desired total filter bank delay. Substituting Eq. (18) and Eq. (16) into Eq. (15) yields

$$\varepsilon_T(\mathbf{g}) = \mathbf{g}^T\mathbf{E}\mathbf{g} - 2\mathbf{g}^T\mathbf{f} + 1 \tag{19}$$

where

$$\mathbf{E} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbf{\Psi}^H(e^{j\omega})\mathbf{h}^*\mathbf{h}^T\mathbf{\Psi}(e^{j\omega})d\omega \tag{20}$$

and

$$\mathbf{f} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Re}\left\{e^{j\omega\tau_T}\mathbf{\Psi}^T(e^{j\omega})\mathbf{h}\right\} d\omega. \tag{21}$$

From Eq. (3) we know that the aliasing terms $\mathcal{Y}(z)$ in the filter bank output signal $Y(z)$ are described by the sum of all repeated spectra

$$\mathcal{Y}(z) = \frac{1}{D}\sum_{d=1}^{D-1} X(zW_D^d)\sum_{m=0}^{M-1}\xi_m(z^D)H(zW_M^mW_D^d)G(zW_M^m). \tag{22}$$

In the ideal case, the aliasing terms in the output signal are zero, i.e. $\mathcal{Y}(z) = 0$. In this case the prototype filters are such that the products of H and G in Eq. (22) are zero for all terms. In the non-ideal case we wish to minimize the energy in all aliasing terms. We define the sum of power magnitudes

$$D(e^{j\omega}) = \frac{1}{D}\sum_{d=1}^{D-1}\sum_{m=0}^{M-1}\left|H(e^{j\omega}W_M^mW_D^d)G(e^{j\omega}W_M^m)\right|^2. \tag{23}$$

We can rewrite Eq. (23) using the impulse responses $\mathbf{h}$ and $\mathbf{g}$

$$D(e^{j\omega}) = \frac{1}{D}\sum_{d=1}^{D-1}\sum_{m=0}^{M-1}\left|\mathbf{h}^T\mathbf{\Phi}_{m,d}(e^{j\omega})\mathbf{g}\right|^2 \tag{24}$$

where

$$\mathbf{\Phi}_{m,d}(z) = \phi(zW_M^mW_D^d)\phi^T(zW_M^m). \tag{25}$$

The output aliasing error is defined as

$$\varepsilon_D = \frac{1}{2\pi}\int_{-\pi}^{\pi} D(e^{j\omega})d\omega = \mathbf{g}^T\mathbf{P}\mathbf{g} \tag{26}$$

where

$$\mathbf{P} = \frac{1}{2\pi D}\sum_{d=1}^{D-1}\sum_{m=0}^{M-1}\int_{-\pi}^{\pi}\mathbf{\Phi}_{m,d}^H(e^{j\omega})\mathbf{h}^*\mathbf{h}^T\mathbf{\Phi}_{m,d}(e^{j\omega})\ d\omega. \tag{27}$$

The optimal synthesis prototype filter in terms of minimal system response error and minimal energy in the output aliasing terms is found by minimizing the total error function

$$\begin{aligned}\varepsilon_{tot}(\mathbf{g}) &= \varepsilon_T(\mathbf{g}) + \varepsilon_D(\mathbf{g}) \\ &= \mathbf{g}^T(\mathbf{E}+\mathbf{P})\mathbf{g} - 2\mathbf{g}^T\mathbf{f} + 1\end{aligned} \tag{28}$$

that is, by solving the set of linear equations system

$$\mathbf{g} = \arg\min_{\mathbf{g}}\varepsilon_{tot}(\mathbf{g}) = (\mathbf{E}+\mathbf{P})^{-1}\mathbf{f}. \tag{29}$$

5 Evaluation

We have designed two critically ($D = M$) and two oversampled ($D = \frac{1}{2}M$) decimated filter banks with 64 subbands and prototype analysis and synthesis filter lengths $L_h = L_g = 128$. The decimation factor is set to $D = 64$ or $D = 32$ and the group delay is specified as $\tau_T = 128$ or $\tau_T = 64$, which gives four scenarios in total. The group delay of the prototype analysis filter is set to $\tau_H = \frac{1}{2}\tau_T$. Table 1 shows the filter bank performance measures after optimization, for the four scenarios. In this table, the *System Phase Error* is defined as

$$\epsilon = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\angle T(e^{j\omega}) - \angle T(e^{j0}) + \tau_T\omega| d\omega. \tag{30}$$

Case	$L = 128, M = 64$	ε_{D_0}	ε_D	ε_T	ϵ
1	$D = 64, \tau_T = 128$	-51.3220	-9.5093	-6.6266	0.0393
2	$D = 64, \tau_T = 64$	-50.2648	-8.9925	-3.1576	0.0718
3	$D = 32, \tau_T = 128$	-71.8347	-28.9326	-23.8421	0.0022
4	$D = 32, \tau_T = 64$	-58.0498	-23.3649	-19.9155	0.0239
		dB	dB	dB	rad

Table 1: *Filter bank performance for the four filter bank cases in the evaluation. The table shows the inband aliasing error ε_{D_0}, the output aliasing error ε_D, the system response error ε_T and the system phase error ϵ.*

We evaluate the performance of the filter banks in the case of a subband RLS beamformer with real data recorded in a hands-free car situation, [7, 8]. In this situation we have a target signal, an interference signal causing echo at the far end of the communication link, and background noise, see Fig. 2.

A linear array of six microphones is used and it was mounted in a car on the visor at the passenger side. The distance between the speaker position and the microphone array is 350 mm and the position is perpendicular to the array axis at the center point. The spacing between adjacent elements in the array is 50 mm. A known white noise sequence is emitted from a human shaped doll in order to determine the LS beamformer weights. The noisy background is recorded separately from the target and interference signals, in a car running at 110 km/h on a normal asphalt road. Recordings of the background signal, the real speech target signal and the interference signal serve as evaluation signals for the beamformer performance.

In order to measure the performance of the beamformer, we introduce the normalized distortion quantity

$$\mathcal{D} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |C_d\hat{P}_{y_S}(\omega) - \hat{P}_{x_S}(\omega)| d\omega \tag{31}$$

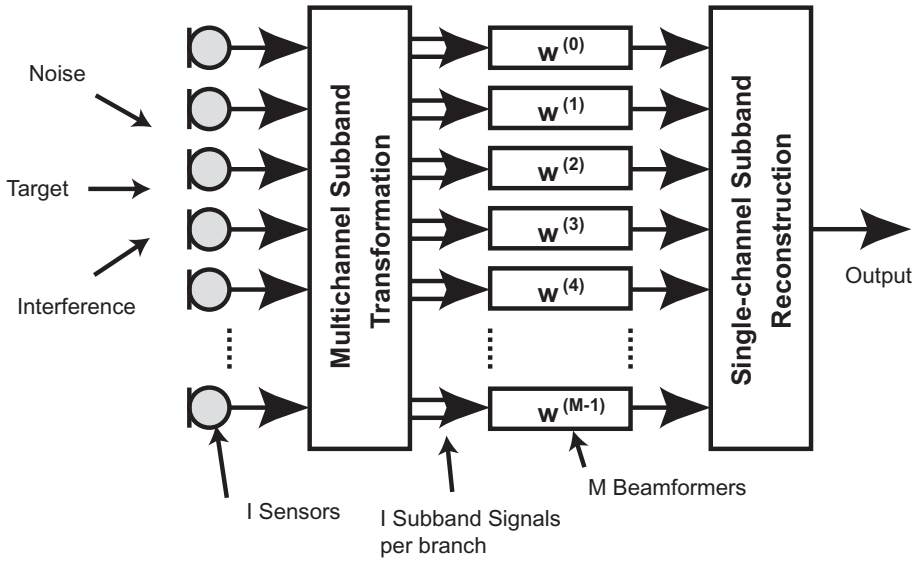

Figure 2: *Subband FIR Beamforming Structure.*

with

$$C_d = \frac{\int_{-\pi}^{\pi} \hat{P}_{x_S}(\omega) d\omega}{\int_{-\pi}^{\pi} \hat{P}_{y_S}(\omega) d\omega}, \tag{32}$$

the normalized interference and noise suppression

$$\mathcal{S}_I = \frac{\int_{-\pi}^{\pi} \hat{P}_{y_I}(\omega) d\omega}{C_d \int_{-\pi}^{\pi} \hat{P}_{x_I}(\omega) d\omega}, \quad \mathcal{S}_N = \frac{\int_{-\pi}^{\pi} \hat{P}_{y_N}(\omega) d\omega}{C_d \int_{-\pi}^{\pi} \hat{P}_{x_N}(\omega) d\omega}. \tag{33}$$

Here, $\hat{P}_x$ is a PSD estimate of a single sensor observation and $\hat{P}_y$ is a PSD estimate of the beamformer output. The indices S, I and N denote the target speech component, the interference component and the noise component, respectively. The beamformer performance measures for the four scenarios are presented in Table 2.

Case	$L = 128, M = 64$	$\mathcal{S}_N$	$\mathcal{S}_I$	$\mathcal{D}$
1	$D = 64, \tau_T = 128$	9.4410	9.6890	-22.9364
2	$D = 64, \tau_T = 64$	5.8847	8.5513	-22.3555
3	$D = 32, \tau_T = 128$	13.0571	12.9027	-28.3018
4	$D = 32, \tau_T = 64$	12.0605	12.6365	-26.6816
		dB	dB	dB

Table 2: *Performance measures for beamformer output in relation to the reference microphone input, when each signal component, speech, noise and echo components are active individually.*

The results show that the beamformer performs better when oversampling is applied instead of critical sampling. They also show that the delays caused by the filter banks can be reduced at the expence of a minor deterioration of the beamformer performance. Fig. 3 shows short-time power estimates of the reference microphone signal and the beamformer output signal for the subband beamformer in scenario 3. We can observe that the background noise is suppressed by about 13 dB and that the interference signal (the male speaker) is suppressed by about 12 dB.

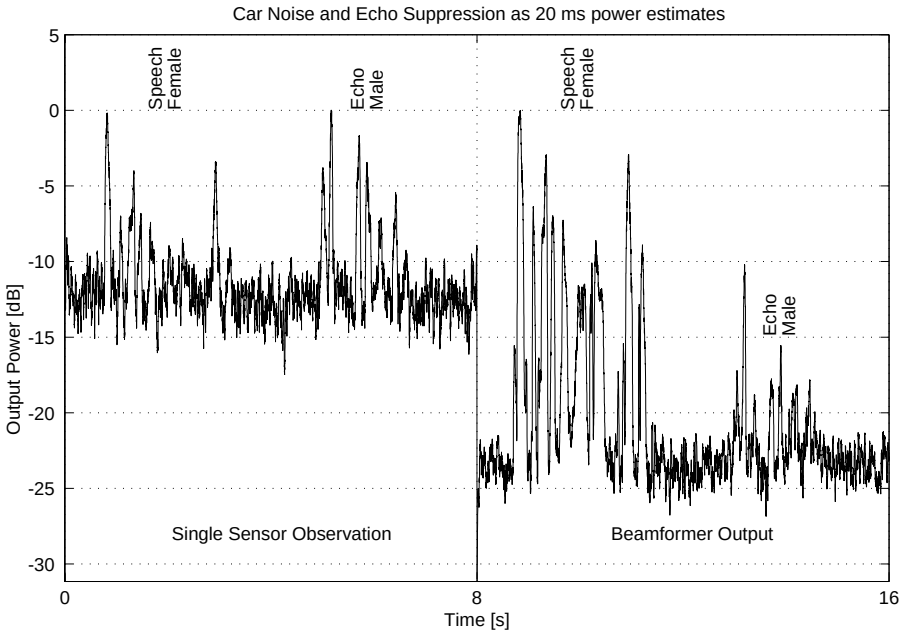

Figure 3: *Short time (20 ms) power estimates of the reference microphone input signal (left) and the beamformer output signal (right). The figure corresponds to scenario 3 in the evaluation.*

6 Conclusions

We have proposed an efficient design method for a uniform DFT filter bank with the possibility of a pre-specified filter bank group delay. The method minimizes the inband and output aliasing error as well as the overall filter-bank transfer function's phase and amplitude deviation. The evaluation on a subband beamformer shows that the accuracy is dependent on both the group delay and the aliasing effects. Subband oversampling allows for a decrease in aliasing and amplitude errors, which in turn increases the performance significantly.

References

[1] W. Kellermann, "Analysis and design of multirate systems for cancellation of acoustic echoes," in Proc. ICASSP'88, pp. 2570-2573, 1988.

[2] S. Haykin, *Adaptive Filter Theory* , Prentice Hall Int. Inc., 1996, ISBN 0-13-397985-7.

[3] P. P. Vaidyanathan, "Orthonormal and biorthonormal filter banks as convolvers, and convolutional coding gain," IEEE Trans. on Signal Processing, vol. 41, no. 6, pp. 2110-2130, Jun. 1993.

[4] P. P. Vaidyanathan, *Multirate Systems and filter banks*, Prentice Hall, 1993.

[5] A. Gilloire, M. Vetterli, "Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation," IEEE Transactions on Signal Processing, pp. 1862-1875, vol. 40, issue 8 , Aug. 1992.

[6] D. R Morgan, J. C. Thi, "A delayless subband adaptive filter architecture," IEEE Trans. Signal Processing, vol. 43, pp. 1819-1830, 1995.

[7] M. Dahl, I. Claesson "Acoustic Noise and Echo Cancelling with Microphone Array," IEEE Trans. on Vehicular Technology, vol. 48, No. 5, pp. 1518-1526, Sep. 1999.

[8] N. Grbić, *Speech Signal Extraction - A Multichannel Approach*, Licenciate Thesis, University of Karlskrona/Ronneby, pp. 72-75, Nov. 1999, ISBN 91-630-8841-X.

Part III

Design of Oversampled Uniform DFT Filter Banks with Reduced Inband Aliasing and Delay Constraints

Part III is submitted for publication as:

N. Grbić, J. M. de Haan, S. Nordholm and I. Claesson and "Design of Oversampled Uniform DFT Filter Banks with Reduced Inband Aliasing and Delay Constraints," submitted for presentation at the Sixth International Symposium on Signal Processing and its Applications, ISSPA, Kuala-Lumpur, Malaysia, Aug. 2001.

Design of Oversampled Uniform DFT Filter Banks with Reduced Inband Aliasing and Delay Constraints

N. Grbić, J. M. de Haan, S. Nordholm* and I. Claesson

Department of Telecommunications and Signal Processing
Blekinge Institute of Technology
Ronneby, Sweden

*Australian Telecommunications Research Institute
Curtin University
Perth, Australia

Abstract

Subband adaptive filters have been proposed to avoid the drawbacks of slow convergence and high computational complexity associated with time domain adaptive filters. Subband processing introduces transmission delays caused by the filter bank and signal degradations due to aliasing effects. One efficient way to reduce the aliasing effects is to allow a higher sample rate than critically needed in the subbands and thus reduce subband signal degradation. We suggest a design method, for a uniform DFT filter bank with any over sampling factor, where the total filter bank group delay may be specified, and where the aliasing and magnitude/phase distortions are minimized.

1 Introduction

Subband adaptive filtering has arised as an alternative for conventional time domain adaptive filtering, [1]. The main reason is the reduction in computational complexity and the increase in convergence speed for the adaptive algorithm, by dividing the algorithm into subbands, [2]. The computational savings comes from the fact that time domain convolution becomes decoupled in the subbands, at a lower sample rate, [3].

Subband analysis and synthesis is often performed using multirate filter banks, [4]. Non-ideal filters in the filter bank cause aliasing of the subband signals. This aliasing can be cancelled in the synthesis bank when certain conditions are met by the synthesis filters and in the subband processing. However, even if this aliasing is cancelled in this way, the inband aliasing is still present in the subband adaptive filter input signals and, consequently, the adaptive filters are perturbed and the overall performance of the system is reduced, [5].
Several solutions to the subband filtering problem have been suggested in the literature. Non-critical decimation has been suggested in [1], where filter bank delay aspects, and amplitude distortions, have not especially been taken into consideration. The use of cross filters, [5], has been suggested to explicitly filter out the aliasing components. A delayless structure has been proposed in [6], where the actual filtering is performed in the time domain, with consequences of higher computational complexity. The computational complexity also increases significantly with cross band filters. The use of allpass IIR filter banks gives very high sidelobe attenuation and has been shown to be computationally efficient, while keeping the aliasing effects low, [7]. Non-linear phase distortions and appearance of narrowband high energy aliasing terms may be noticed at the subband boundaries with this approach.
We use an uniform DFT modulated FIR filter bank for the subband transformations. Modulated filter banks provide a computationally efficient implementation, due to the polyphase implementation [4], and great design simplicity. The main contribution in this paper is the suggested design method, where the inband and the reconstruction aliasing are minimized simultaneously, while the total filter bank group delay is pre-specified. A numerical comparison of a real room identification shows that both the inband aliasing and the filter bank delay affects the identification accuracy.

2 The Uniform DFT Modulated filter bank

A set of M filters forms a uniform DFT analysis filter bank when they are related to a prototype filter, $h(n)$, as

$$H_m(z) = H(zW_M^m) = \sum_{n=-\infty}^{\infty} h(n)(zW_M^m)^{-n} \qquad m = 0, \cdots, M-1 \tag{1}$$

where $W_M = e^{-j2\pi/M}$. Each subband signal is decimated by a factor, D. An efficient implementation of such filter bank is given in [8]. In order to analyze the magnitude and aliasing effects caused by the filter bank, a direct form realization of the filter bank (see Fig. 1) is beneficial when it comes to uniformity of the derivations for different decimation factors. The subband filters in the synthesis

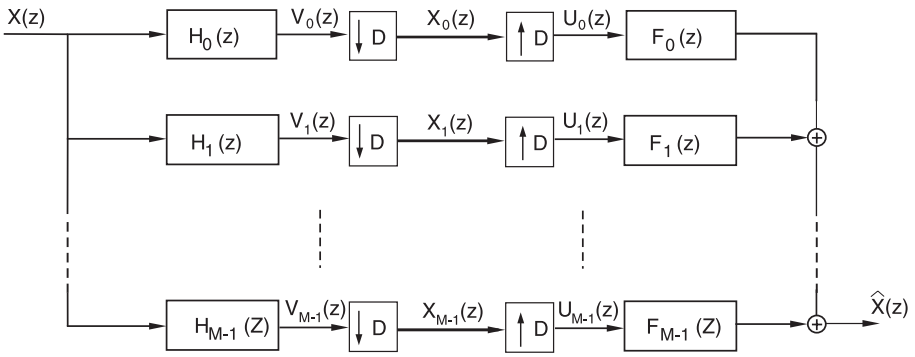

Figure 1: *Direct form filter bank realization.*

filter bank are the same as for the analysis filter bank, i.e. $F_m(z) = H_m(z)$. Each branch signal, $V_m(z)$, is simply a filtered version of the input signal

$$V_m(z) = H_m(z)X(z) = H(zW_M^m)X(z). \tag{2}$$

The decimators cause a summation of repeated and expanded spectra of the input signal

$$\begin{aligned} X_m(z) &= \frac{1}{D}\sum_{d=0}^{D-1} V_m(z^{\frac{1}{D}}W_D^d) \\ &= \frac{1}{D}\sum_{d=0}^{D-1} H(z^{\frac{1}{D}}W_M^m W_D^d)X(z^{\frac{1}{D}}W_D^d) \end{aligned} \tag{3}$$

where $W_D = e^{-j2\pi/D}$. The interpolators have a compressing effect

$$U_m(z) = X_m(z^D) = \frac{1}{D}\sum_{d=0}^{D-1} H(zW_M^m W_D^d)X(zW_D^d). \tag{4}$$

By yet another filtering operation by the reconstruction filters we state a relation between the input signal $X(z)$ and the output signal $\hat{X}(z)$

$$\hat{X}(z) = \sum_{d=0}^{D-1} A_l(z)X(zW_D^d) \tag{5}$$

where

$$A_l(z) = \frac{1}{D}\sum_{m=0}^{M-1} H(zW_M^m W_D^d)H(zW_M^m). \tag{6}$$

The transfer functions, $A_l(z), \quad d = 1, \cdots, D-1$, can be viewed as aliasing gains. The function, $A_0(z)$, is the magnitude gain for the original input signal spectrum. The magnitude and phase response of the overall transfer function, $T(z)$, and the

distortion function, $D(z)$, can be described in terms of the repeated spectrum gains, $A_l(z)$, as

$$T(z) = \sum_{d=0}^{D-1} A_l(z) \qquad D(z) = \sum_{d=1}^{D-1} A_l(z). \tag{7}$$

For each subband signal, we define a measure of the inband aliasing as the total gain of all repeated spectra in that subband. Due to the modulated structure of the filterbank, the inband aliasing gains will be the same for all subband signals. Thus, we only measure the inband aliasing for the first subband, $D_0(z)$, given by

$$D_0(z) = \frac{1}{D} \sum_{d=1}^{D-1} |H(z^{\frac{1}{D}} W_D^d)|. \tag{8}$$

The inband aliasing given in Eq. (8) is an upper bound on the actually incurred aliasing term, since cancellation of the terms in the sum may occur. However, filtering in the subbands will alter the cancellation effect and the objective is to minimize the upper bound in order to successfully tackle adaptive filtering in the subbands.

3 Prototype filter optimization

We want to find a prototype filter $\mathbf{h}$ which minimizes the overall error function $E(\mathbf{h})$

$$E(\mathbf{h}) = \alpha\epsilon_1(\mathbf{h}) + \beta\epsilon_2(\mathbf{h}) + \gamma\epsilon_3(\mathbf{h}) + \delta\epsilon_4(\mathbf{h}) \tag{9}$$

where the factors α, β, γ and δ are weight factors and the average amplitude error, ϵ_1, is defined as

$$\epsilon_1(\mathbf{h}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |1 - |T(e^{j\omega})|^2| d\omega \tag{10}$$

and the average phase error, ϵ_2, is defined as

$$\epsilon_2(\mathbf{h}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\angle T(e^{j\omega}) - \angle T(e^{j0}) + \tau_T \omega| d\omega \tag{11}$$

where τ_T is the desired group delay of the filter bank response $T(z)$. The average aliasing distortion, ϵ_3, is defined as

$$\epsilon_3(\mathbf{h}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D(e^{j\omega})|^2 d\omega \tag{12}$$

and the average aliasing distortion in the first subband, ϵ_4, is defined as

$$\epsilon_4(\mathbf{h}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_0(e^{j\omega})|^2 d\omega. \tag{13}$$

The minimization of Eq. (9) is a very involved non-linear optimization problem. We suggest an iterative two step procedure, where we optimize Eq. (9) based

on two design parameters, a passband boundary frequency, ω_p, and a stopband boundary frequency, ω_s. For given design parameters, a prototype filter $\mathbf{h}$, can be created using methods outlined in [9] and [10]. This is a complex domain filter design method which allows also for optimization on the group delay. The objective in this design is to minimize the function, $J(\omega_s, \omega_p, \tau_T)$, on a sampled grid of frequencies

$$J(\omega_s, \omega_p, \tau_T) = \sum_{i=1}^{N} |H_d(\omega_i) - H(\omega_i)|^2 \tag{14}$$

where

$$H_d(\omega) = \begin{cases} e^{-j\omega\tau_T/2} & \omega \in \Omega_p \\ 0 & \omega \in \Omega_s \end{cases} \tag{15}$$

is the desired complex filter specification with the passband region defined as $\Omega_p = [0, \omega_p]$ and the stopband region defined as $\Omega_s = [\omega_s, \pi]$. The predetermined total desired filter bank group delay is τ_T, and

$$H(\omega) = \sum_{n=0}^{L-1} h(n) e^{-j\omega n} \tag{16}$$

is the frequency response of the prototype filter.

The objective of the optimization is then,

$$[\omega_p, \omega_s] = \arg \min_{\omega_p > 0, \omega_s > \omega_p} E(\mathbf{h}(\omega_p, \omega_s, \tau_T)). \tag{17}$$

The algorithm follows:

1. **Initialization phase.** The passband and stopband frequencies are initialized with $\omega_{p,0} = \frac{\pi}{M}$ and $\omega_{s,0} = \frac{\pi}{D}$, respectively. Initial step sizes $\zeta_{p,0}$ and $\zeta_{s,0}$ are set. Iteration index, i, is set to 0.

2. **Design phase.** The prototype filter $\mathbf{h}(\omega_{p,i}, \omega_{s,i}, \tau_T)$, is designed such that Eq. (14) is minimized.

3. **Optimization phase.** The filter specification frequencies ω_p and ω_s are adapted according to

$$\omega_{p,i+1} = \omega_{p,i} - \zeta_{p,i} \mathrm{sgn}(\dot{\omega}_{p,i}) \tag{18}$$

$$\omega_{s,i+1} = \omega_{s,i} - \zeta_{s,i} \mathrm{sgn}(\dot{\omega}_{s,i}) \tag{19}$$

where

$$\dot{\omega}_p = \frac{E_i - E_{i-1}}{\omega_{p,i} - \omega_{p,i-1}} \qquad \dot{\omega}_s = \frac{E_i - E_{i-1}}{\omega_{s,i} - \omega_{s,i-1}} \tag{20}$$

where we have used discretized approximations of the gradient. The step sizes are exponentially decreased. Index i is increased by one and steps **2** and **3** are continued until a stop criterion is met.

4 Evaluation

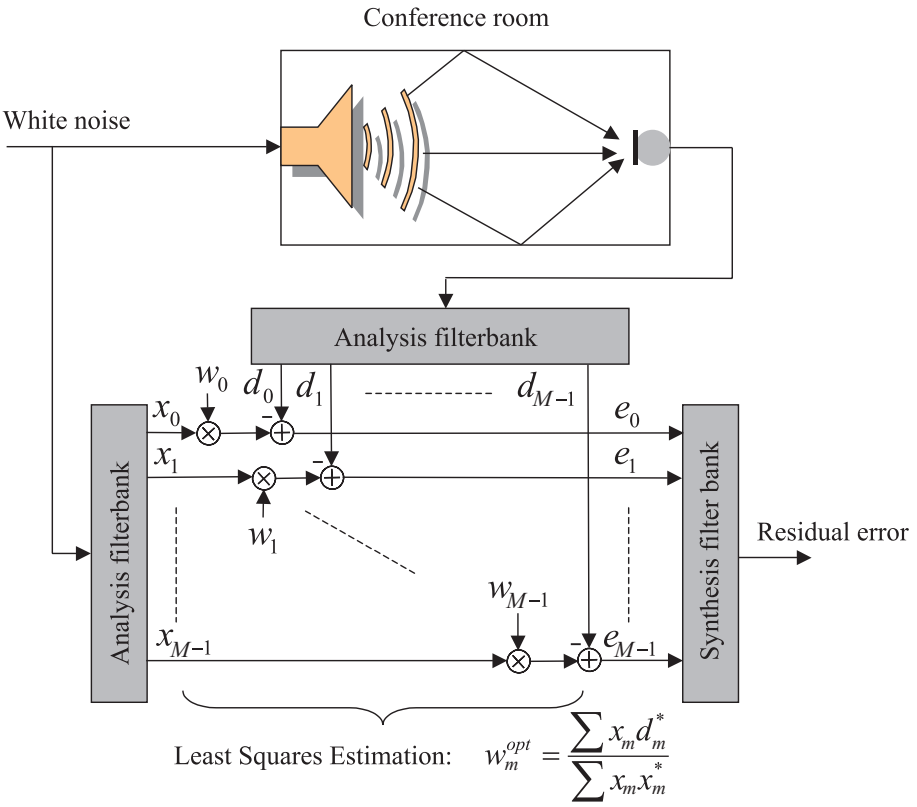

Figure 2: *Subband room impulse response identification.*

We have designed two critical and two non-critical decimated filter banks with 32 subbands and each with decimation factor 32 and 16, respectively. The length of the prototype filter is 128 and the group delay is specified as $\tau_T = 128$ and $\tau_T = 64$. Table 1 shows the final distortion measures after optimization, for the four scenarios. We evaluate the performance of the subband implementations in the case of a real room impulse response estimation. A white noise sequence is emitted through a loudspeaker in a conference room and by using a microphone observation as a desired signal, we identify the acoustic path, see Fig. 2. We use the least squares estimation method, [2], individually in each subband, and compare the fullband FIR filter identification with subband identifications achieved with the filter banks given in Table 1. The average spectral error of the estimations, are given in Table 2, together with an FFT filter bank implementation. The system responses of the critically decimated filter bank identifications are shown in Fig. 3, while Fig. 4 show the response of the non-critically decimated filter banks, together with the real room system response. It can be seen that the variations are much larger with the critically decimated subband implementation, especially at the subband boundaries.

$L = 128, M = 32$		$\tau_T = 128$		$\tau_T = 64$	
$D = 32$	ϵ_1	-12.56	dB	-3.42	dB
	ϵ_2	0.03	rad	0.76	rad
	ϵ_3	-44.57	dB	-43.52	dB
	ϵ_4	-12.28	dB	-9.48	dB
$D = 16$	ϵ_1	-21.55	dB	-9.85	dB
	ϵ_2	0.02	rad	0.05	rad
	ϵ_3	-79.22	dB	-68.02	dB
	ϵ_4	-48.31	dB	-36.58	dB

Table 1: *Average distortion measures, $\epsilon_1, \ldots, \epsilon_4$ for critical and non-critical decimation with two specified delay cases.*

$L = 128, M = 32$	$\tau_T = 128$	$\tau_T = 64$
$D = 32$	-8.17 dB	-7.74 dB
$D = 16$	-8.81 dB	-9.16 dB
FFT Filter Bank	-5.20 dB	
Fullband	-9.84 dB	

Table 2: *Average spectral error of least squares solutions to a system identification of an acoustic path in a conference room. Four subband cases with 32 weights and an FFT filter bank are compared to the fullband solution.*

The estimation accuracy for the non-critical decimated filter banks are close to the fullband solution and significantly better than the critically decimated cases.

5 Conclusions

We have proposed an efficient design method for a uniform DFT filter bank with the possibility of a prespecified filter bank group delay. The optimization minimizes the inband aliasing components as well as the overall filter bank transfer function's phase and amplitude deviation. A real room transfer function estimation shows that the accuracy is dependent on both the group delay and the aliasing effects. Subband oversampling decreases the inband aliasing, which in turn increases the estimation accuracy. The gain with over sampling is more significant when reduced delay filter banks are used.

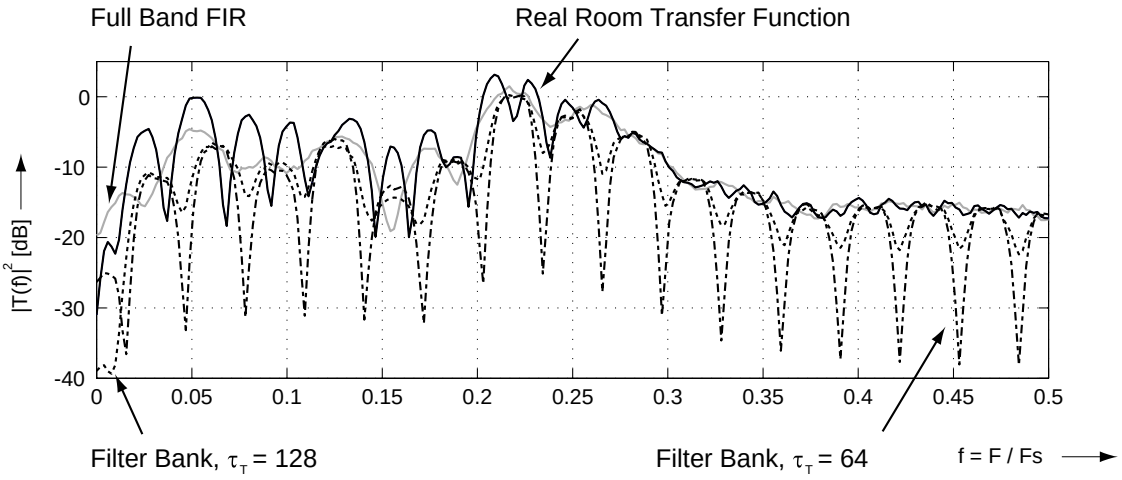

Figure 3: *System responses of the real room transfer function estimates for the fullband implementation and the critically decimated subband implementations,* $L = 128$, $M = 32$, $D = 32$. *It can be seen that the identification error is relatively large.*

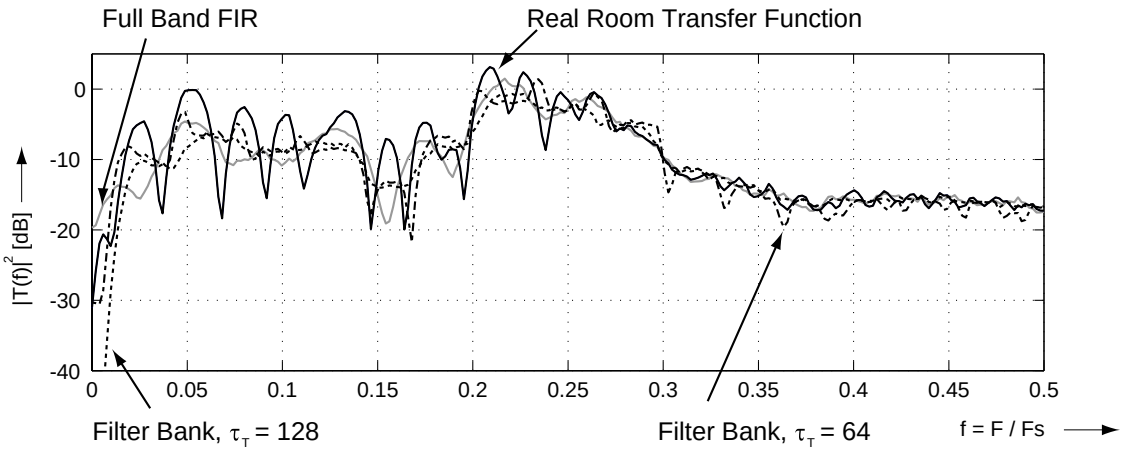

Figure 4: *System responses of the real room transfer function estimates for the fullband and the non-critically decimated subband implementations,* $L = 128$, $M = 32$, $D = 16$. *It can be seen that the identification error is reduced compared to Fig. 3.*

References

[1] W. Kellermann, "Analysis and design of multirate systems for cancellation of acoustic echoes," in Proc. ICASSP'88, pp. 2570-2573.

[2] S. Haykin, *Adaptive Filter Theory* , Prentice Hall Int. Inc., 1996, ISBN 0-13-397985-7.

[3] P. P. Vaidyanathan, "Orthonormal and biorthonormal filter banks as convolvers, and convolutional coding gain," IEEE Trans. on Signal Processing, vol. 41, no. 6, pp. 2110-2130, Jun. 1993.

[4] P. P. Vaidyanathan, *Multirate Systems and filter banks*, Prentice Hall, 1993.

[5] A. Gilloire, M. Vetterli, "Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation," IEEE Transactions on Signal Processing, pp. 1862-1875, vol. 40, issue 8 , Aug. 1992.

[6] D. R Morgan, J. C. Thi, "A delayless subband adaptive filter architecture," IEEE Trans. Signal Processing, vol. 43, pp. 1819-1830, 1995.

[7] P. A. Naylor, O. Tanrikulu A. G. Constantinidis, "Subband Adaptive Filtering for Acoustic Echo Control using Allpass Polyphase IIR filter banks," IEEE Trans. on Speech and Audio Processing, vol. 6, no. 2, Mar. 1998.

[8] N. Grbić, *Speech Signal Extraction - A Multichannel Approach*, Licenciate Thesis, University of Karlskrona/Ronneby, pp. 72-75, Nov. 1999, ISBN 91-630-8841-X.

[9] T. W. Parks, C. S. Burrus, *Digital Filter Design*, John Wiley and Sons, Inc., 1987, ISBN 0-471-82896-3.

[10] D. R. Xiangkun Chen, T. W. Parks, "Design of FIR filters in the Complex Domain," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-35, no. 2, Feb. 1979.