A New Factorization and Structure for Cosine Modulated Filter Banks with Variable System Delay

Gerald Schuller, University of Hannover
Institut fuer Theoretische Nachrichtentechnik und Informationsverarbeitung
30167 Hannover, Germany
email schuller@tnt.uni-hannover.de, phone: +49-511-7625027

Abstract

A new design method for biorthogonal modulated filter banks is presented. It is based on a cascade of simple matrices, and it has some properties that have not been reported before. It represents filter banks with arbitrary overall system delay and filter length, it is shown that almost all cosine modulated filter banks can be described by this structure, and that it leads to a more efficient implementation than previous structures. Imposing certain symmetries on the matrices can be used to design low delay filter banks with identical (except for the sign) baseband impulse responses for the analysis and synthesis filter bank.

1. Introduction

The system considered here is an N band (N even) cosine modulated FIR filter bank with critical downsampling and perfect reconstruction. The analysis part consists of N analysis filters with impulse responses $h_k(n)$ with subsequent downsampling by N, where k is the frequency or band index. Its input signal is $x(n)$, its output are the N subbands $y_k(m)$, where m is the time index at the lower sampling rate. The synthesis part consists of upsamplers by N followed by N synthesis filters with impulse responses $g_k(n)$. The output of these filters is summed to form the reconstructed signal $\hat{x}(n)$. The filter bank has perfect reconstruction if $\hat{x}(n) = x(n - n_d)$, where n_d is the system delay.

Filter banks used so far usually have orthogonal filters and a standard system delay which equals the length of its filters minus one sample, assumed its filters are of equal length. Their delays are determined by the lower sampling rate, i.e. they have system delays which are restricted to be integer multiples of N, minus one.

The filter bank presented here is not orthogonal but biorthogonal, and the system delay can be pre-specified in the design process independently of the filter length. This makes filter banks with non-symmetric filters and a lower than usual delay possible, which is important for applications like speech and audio coding. Its delay does not need to be an integer multiple of N, it can be specified in terms of the higher sampling rate. It also possesses an computationally efficient implementation and the perfect reconstruction property is maintained even if low precision arithmetic is used for its implementation.

The impulse responses have the form

$$h_k(n) = h(n) \cdot \cos\left(\frac{\pi}{N}(k+0.5)(n+0.5+n_0)\right)$$

$$g_k(n) = h'(n) \cdot \frac{2}{N} \cdot \cos\left(\frac{\pi}{N}(k+0.5)(n+0.5-N+n_0')\right)$$

$k = 0, \ldots, N-1, n = 0, \ldots, LN-1$
Their length is LN, which can include leading or trailing zeros. The factor $2/N$ is just a normalization which simplifies the following notation. $h(n)$ and $h'(n)$ are the analysis and synthesis baseband prototype filters respectively. For causality $h(n) = 0$ and $h'(n) = 0$ for $n < 0$. n_0 and n_0' can be limited to $-N \leq n_0, n_0' \leq N$ due to the periodicity of the cosine function. The modulation function of g_k has a shift of N because it better suits the form for perfect reconstruction, as will be seen.

1.1. Definitions

Boldface letters denote matrices or vectors. ":=" means "defined as". A polynomial matrix $\mathbf{f}(z)$ is causal if it contains no positive powers of z. $\mathbf{I}$ is the $N \times N$ identity matrix, the anti-diagonal matrix is defined as

$$\mathbf{J} := \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

$$\mathbf{diag}(x_0,\ldots,x_{N-1}) := \begin{bmatrix} x_0 & 0 & \cdots & 0 \\ 0 & x_1 & \cdots & 0 \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & x_{N-1} \end{bmatrix}$$

$$\mathbf{S}(z) := \begin{bmatrix} 0 & 0 & \cdots & 0 & z \\ 1 & 0 & \cdots & & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

$\mathbf{diag}$ is an $N \times N$ diagonal matrix and $\mathbf{S}(z)$ is a shift matrix. The rank of a (square) matrix $\mathbf{f}$ is $\mathrm{rank}(\mathbf{f})$ and defined as the number of linearly independent row or column vectors of $\mathbf{f}$. If the rank is equal to the number of rows or columns of $\mathbf{f}$ it has a full rank. $[\mathbf{f}]_{n,k}$ denotes the element at the n'th row and k'th column of the matrix $\mathbf{f}$.

2. The Polyphase Notation

For an N-band analysis/synthesis filter bank, the input is represented by an N-dimensional row vector $\mathbf{x}(n)$ composed of the downsampled input components

$$\mathbf{x}(m) := [x(mN), \ldots, x(mN+N-1)]$$

Taking the z-transform of each element we obtain the vector

$$\mathbf{X}(z) = [X_0(z), \ldots, X_{N-1}(z)]$$

similar for the reconstructed signal $\hat{\mathbf{X}}(z)$. For every block of N input samples, N output samples are produced. These outputs are $y_k(m)$ where $k = 0, 1, \ldots, N-1$ and are also expressed as the vector $\mathbf{y}(m)$ with corresponding z-transform row vector $\mathbf{Y}(z) = [Y_0(z), \ldots, Y_{N-1}(z)]$. The analysis filters $h_k(n)$ that convert $\mathbf{x}(m)$ into $\mathbf{y}(m)$ are represented as an analysis polyphase filter matrix $\mathbf{P_a}$, its elements are

$$[\mathbf{P_a}(z)]_{n,k} := \sum_{m=0}^{L-1} h_k(mN+N-1-n)z^{-m}$$

$n, k = 0, \ldots, N-1$, where k is the frequency index and n the time index. The analysis section can now be completely described by the equation

$$\mathbf{Y}(z) = \mathbf{X}(z) \cdot \mathbf{P_a}(z)$$

The synthesis polyphase matrix is

$$[\mathbf{P_s}(z)]_{k,n} := \sum_{m=0}^{L-1} g_k(mN+n)z^{-m}$$

So that the output of the synthesis filter bank can be written as

$$\hat{\mathbf{X}}(z) = \mathbf{Y}(z) \cdot \mathbf{P_s}(z)$$

Perfect reconstruction means $\mathbf{P_a}(z) \cdot \mathbf{P_s}(z) = z^{-d} \cdot \mathbf{S}^{n_t}(z)$ The system delay n_d is the delay introduced by the above matrices plus the blocking delay of length $N-1$, which results from forming input blocks $\mathbf{X}(z)$ of length N before processing them. It results to $n_d = d \cdot N + N - 1 - n_t$. n_t can be used for the "fine tuning" of the system delay.

3. The New Factorization

Modulated filter banks have certain symmetries in their impulse responses that can be used for the design and efficient implementation of the filter bank. The important point here is that their polyphase matrices can be written as a product of a sparse "filter matrix" with polynomial elements, $\mathbf{F_a}(z)$, $\mathbf{F_s}(z)$, a transform matrix $\mathbf{T}$, and the shift matrix $\mathbf{S}(z)$. They all can be implemented efficiently. $\mathbf{T}$ can basically be any transform matrix, but here it is assumed to be a Discrete Cosine Transform type IV matrix, defined as

$$[\mathbf{T}]_{n,k} := \cos(\frac{\pi}{N}(k+0.5)(n+0.5)) \ , \quad 0 \le n, k < N$$

This means the polyphase matrices can be written as the product

$$\mathbf{P_a}(z) = \mathbf{S}^{n_a}(z) \cdot \mathbf{F_a}(z) \cdot \mathbf{T}$$

$$\mathbf{P_s}(z) = \mathbf{T}^{-1} \cdot \mathbf{F_s}(z) \cdot \mathbf{S}^{n_s}(z)$$

n_a, n_s are in the range of $0 < n_a, n_s \le N$, with $n_a = n_0$ if $n_0 > 0$, else $n_a = n_0 + N$, and $n_s = n'_0$ if $n'_0 > 0$, else $n_s = n'_0 + N$. The filter matrices can be obtained with $\mathbf{F_a}(z) := \mathbf{S}^{-n_a}(z) \cdot \mathbf{P_a}(z) \cdot \mathbf{T}^{-1}$ and $\mathbf{F_s}(z) := \mathbf{T} \cdot \mathbf{P_s}(z) \cdot \mathbf{S}^{-n_s}(z)$. They have a sparse, "bi-diagonal" form

$$\mathbf{F_a}(z) = [\mathbf{diag}(P_{N-1}(z), \ldots, P_0(z)) \cdot \mathbf{J} +$$

$$+ z^{-1} \cdot (-1)^{i_a} \mathbf{diag}(-P_{2N-1}(z), \ldots, -P_N(z))] \cdot \mathbf{J}^{i_a}$$

$$\mathbf{F_s}(z) = \mathbf{J}^{i_s} \cdot [\mathbf{diag}(P'_{N-1}(z), \ldots, P'_0(z)) \cdot \mathbf{J} +$$

$$+ z^{-1} \cdot (-1)^{i_s} \mathbf{diag}(P'_N(z), \ldots, P'_{2N-1}(z))]$$

where $i_a = 0$ if $n_a = n_0$, else $i_a = 1$, and $i_s = 0$ if $n_s = n'_0$, else $i_s = 1$, and with

$$P_k(z) = \sum_{m=-\infty}^{\infty} h(m2N + k - n_a)(-1)^m z^{-2m} \qquad (1)$$

$k = 0, \ldots, 2N-1$,

$$P'_k(z) = \sum_{m=-\infty}^{\infty} h'(m2N + k - n_s)(-1)^m z^{-2m} \qquad (2)$$

It is assumed that $i_a = i_s$ because this results in the suitable form for perfect reconstruction (see also sec. 5 eqs. 7, 8). These filter matrices could now already be used for the design of filter banks. To perfectly reconstruct a signal from a given analysis filter bank the synthesis filter matrix needs to be the inverse of the analysis filter matrix, multiplied with a delay z^{-d} to make it causal. But this approach may lead to IIR synthesis filters, which may not be stable. There would also be no direct control over the system delay, which is determined by the additional delay z^{-d}. The goal is now to obtain FIR analysis and also FIR synthesis filters with the perfect reconstruction property, to have control over the overall system delay, and to obtain a structure for an efficient implementation. This is done by constructing the filter matrices as a product or cascade of two basic types of simpler matrices. The simple matrices have an inverse, which is FIR, have different system delays associated with them, and are sparse with only a few elements unequal to 1 or 0, which leads to an efficient implementation. The design process then consists of choosing the matrices for the desired properties (system delay, filter length) and then to optimize the coefficients of the resulting cascade for the desired frequency response. These simple matrices are described in the following.

Zero-Delay Matrices– They increase the filter length but not the system delay.

$$\mathbf{E}_i(z) := \mathbf{J} + z^{-1} \cdot \mathbf{diag}(0, \ldots, 0, e^i_{N/2}, \ldots, e^i_{N-1})$$

$$\mathbf{G}_i(z) := \mathbf{J} + z^{-1} \cdot \mathbf{diag}(g^i_0, \ldots, g^i_{N/2-1}, 0, \ldots, 0)$$

where e^i_j, g^i_j are matrix coefficients, and i denotes different sets of coefficients ($i > 0$). Observe that their inverse is causal, so that no multiplication with a delay is necessary.

$$\mathbf{E}^{-1}_i(z) = \mathbf{J} + z^{-1} \cdot \mathbf{diag}(-e^i_{N-1}, \ldots, -e^i_{N/2}, 0, \ldots, 0)$$

$$\mathbf{G}^{-1}_i(z) = \mathbf{J} + z^{-1} \cdot \mathbf{diag}(0, \ldots, 0, -g^i_{N/2-1}, \ldots, -g^i_0)$$

Maximum-Delay Matrices– They also increase the filter length, but especially the system delay.

$$\mathbf{A}_i(z) := z^{-1} \cdot \mathbf{J} + \mathbf{diag}(0, \ldots, 0, a^i_{N/2}, \ldots, a^i_{N-1})$$

$$\mathbf{B}_i(z) := z^{-1} \cdot \mathbf{J} + \mathbf{diag}(b^i_0, \ldots, b^i_{N/2-1}, 0, \ldots, 0)$$

The matrix $\mathbf{B}_0(z)$ uses also coefficients on the anti-diagonal,

$$\mathbf{B}_0(z) := [z^{-1} \cdot \mathbf{J} \cdot \mathbf{diag}(b^0_N, \ldots, b^0_{2N-1}) +$$

$$+\mathbf{diag}(b^0_0, \ldots, b^0_{N/2-1}, 0, \ldots, 0)] \cdot \mathbf{J}^{i_a}$$

Their inverse need a multiplication with z^{-2} to to obtain a causal matrix.

$$z^{-2} \cdot \mathbf{A}^{-1}_i(z) = z^{-1} \cdot \mathbf{J} +$$

$$+\mathbf{diag}(-a^i_{N-1}, \ldots, -a^i_{N/2}, 0, \ldots, 0),$$

$$z^{-2} \cdot \mathbf{B}^{-1}_i(z) = z^{-1} \cdot \mathbf{J} +$$

$$+\mathbf{diag}(0, \ldots, 0, -b^i_{N/2-1}, \ldots, -b^i_0),$$

$$z^{-2} \cdot \mathbf{B}^{-1}_0(z) = \mathbf{J}^{i_a}[z^{-1} \cdot \mathbf{diag}(\hat{b}^0_N, \ldots, \hat{b}^0_{2N-1}) \cdot \mathbf{J} +$$

$$+\mathbf{diag}(0, \ldots, 0, \hat{b}^0_{N/2}, \ldots, \hat{b}^0_{N-1})]$$

with $\hat{b}^0_{N/2+j} = -b^0_j/(b^0_{N+j} b^0_{2N-1-j})$, $j = 0...N/2-1$, and $\hat{b}^0_{N+j} = 1/b^0_{2N-1-j}$, $j = 0...N-1$

A product of these matrices has to have the shape of the filter matrix, i.e. it must have a bi-diagonal shape and the distribution of the even and odd powers of z must be as in the filter matrix. This ensures that the resulting polyphase matrix leads to a modulated filter bank. The following products or cascades have this property. The analysis filter matrix is

$$\mathbf{F_a}(z) = \mathbf{B}_0(z) \cdot \prod_{i=1}^{\mu-1} \mathbf{H}_i(z) \cdot \prod_{j=1}^{\nu} \mathbf{L}_j(z)$$

where ν and μ are the number of zero-delay matrices and maximum-delay matrices resp. The synthesis filter matrix for perfect reconstruction is

$$\mathbf{F_s}(z) = z^{-d} \mathbf{F_a}^{-1}(z) =$$

$$= \prod_{j=\nu}^{1} \mathbf{L}^{-1}_j(z) \cdot \prod_{i=\mu-1}^{1} \left(z^{-2}\mathbf{H}^{-1}_i(z)\right) \cdot \left(z^{-2}\mathbf{B}^{-1}_0(z)\right)$$

where $d = 2\mu$. $\mathbf{H}_i$ and $\mathbf{L}_i$ are defined as $\mathbf{H}_i(z) := \mathbf{B}_i(z)$ if $i_a = 0$, else $\mathbf{H}_i(z) := \mathbf{A}_i(z)$. $\mathbf{L}_i(z) := \mathbf{E}_i(z)$ if $\mathbf{H}_i = \mathbf{B}_i$ and μ is even or $\mathbf{H}_i = \mathbf{A}_i$ and μ is odd, else $\mathbf{L}_i(z) := \mathbf{G}_i(z)$. The coefficients of $\mathbf{B}_0(z)$ which lead to coefficients of $\mathbf{S}^{n_a}(z) \cdot \mathbf{B}_0(z)$ with positive powers have to be set to zero in order to obtain causal filters, where it is assumed that $n_a \geq n_s$. This can always be done since $0 < n_a \leq N$. The system delay is

$$n_d = \mu \cdot 2N + N - 1 - n_a - n_s$$

As can be seen the minimum possible delay is the blocking delay of $N - 1$ samples. It is obtained with $\mu = 1$, $n_a = n_s = N$ and is independent of the filter length. The length of the non-zero part of the analysis and synthesis filters is $(\mu + \nu)N + N/2 - \max(N/2, n_a)$ for $\nu > 0$ and $\mu N + N - \max(N/2, n_a)$ for $\nu = 0$, where max(.,.) is the maximum of the two values. $h(n)$ and $h'(n)$ have $N/2$ leading zeros if n_a and n_s is zero. That is why the system delay in this case can be reduced by N without reducing the filter quality by increasing n_a and n_s to $N/2$. Observe that in this case filter banks with a standard delay are obtained if $\nu = \mu$. An efficient implementation of the filter bank can be obtained by implementing the the simple matrices and the shift matrix, and to take an efficient algorithm for the $N \times N$ DCT. The number of multiplications necessary is the number of elements in the simple matrices which are not 0 or 1, plus the number for the fast transform. This number, without the transform, is less or equal to $1.5 \cdot N - \max(N/2, n_a) + (\mu+\nu) \cdot N/2$ for each, the analysis and the synthesis. This implementation is more general and also more efficient than previous approaches, e.g. like Malvars ELT [2]. Also note that the coefficients for the synthesis matrices result from sign flipping, and that the input for the multipliers is the same as for the analysis (except for the matrix $\mathbf{B}_0$) which means that they provide perfect reconstruction even if they are implemented with low precision arithmetic, as long as the sign flipping is exact. The coefficients of the simple matrices determine the frequency responses of the filter bank. They can be obtained e.g. with the optimization described in [5, 6, 7, 8].

4. Completeness

This section is a proof to show that all FIR cosine modulated filter banks with perfect reconstruction where $h(n)$ and $h'(n)$ have one contiguous nonzero part can be represented by the given factorization. It presents an iterative algorithm for the extraction of the $\mathbf{L}_i$ and $\mathbf{H}_i$ from the filter matrices of a given filter bank. The filter matrices can be written as a polynomial of matrices,

$$\mathbf{F_a}(z) = \sum_{m=0}^{L-1} \mathbf{f_a}(m) \cdot z^{-m} \ , \ \mathbf{F_s}(z) = \sum_{m=0}^{L-1} \mathbf{f_s}(m) \cdot z^{-m}$$

Perfect reconstruction results in

$$z^{-d} \cdot \mathbf{I} = \mathbf{F_a}(z) \cdot \mathbf{F_s}(z) = \sum_{m=0}^{2L-2} z^{-m} \sum_{i+j=m} \mathbf{f_a}(i) \cdot \mathbf{f_s}(j)$$

Now consider the matrices for certain exponents m. If $d < 2L-3$, then for $m = 2L-2$ it follows

$$\mathbf{0} = \textstyle\sum_{i+j=m} \mathbf{f_a}(i) \cdot \mathbf{f_s}(j) = \mathbf{f_a}(L-1) \cdot \mathbf{f_s}(L-1) \quad (3)$$

and for $m = 2L-3$

$$\mathbf{f_a}(L-1) \cdot \mathbf{f_s}(L-2) + \mathbf{f_a}(L-2) \cdot \mathbf{f_s}(L-1) = \mathbf{0} \quad (4)$$

Since $h(n)$ is a contiguous nonzero filter $\mathbf{f_a}(L-2)$ and $\mathbf{f_s}(L-2)$ are diagonal or anti-diagonal matrices with full rank, since $\mathbf{f_a}(L-1)$ and $\mathbf{f_s}(L-1)$ contain the end of the nonzero part of the baseband impulse response (also compare with eqn.1, 2). Equation 3 means $\mathrm{rank}(\mathbf{f_a}(L-1)) + \mathrm{rank}(\mathbf{f_s}(\mathrm{L}-1)) \leq \mathrm{N}$, and equation 4 means that $\mathrm{rank}(\mathbf{f_a}(L-1)) = \mathrm{rank}(\mathbf{f_s}(\mathrm{L}-1))$. It follows that $\mathrm{rank}(\mathbf{f_a}(L-1)) = \mathrm{rank}(\mathbf{f_s}(\mathrm{L}-1)) \leq \mathrm{N}/2$. Since $\mathbf{f_a}(L-1)$ and $\mathbf{f_s}(L-1)$ are diagonal or anti-diagonal matrices, the number of their non-zero elements is less than or equal to $N/2$. Since the baseband impulse response is contiguous these non-zero elements must also be contiguous on the diagonal or anti-diagonal, bordering on the right or left side of the matrix. The right side of equations 3 and 4 is still zero if they are multiplied by $(\mathbf{f_s}(L-2))^{-1} \cdot \mathbf{J}$ from the right side or $\mathbf{J} \cdot (\mathbf{f_a}(L-2))^{-1}$ from the left. If we define

$$\mathbf{L}_i(z) := \mathbf{J} \cdot (\mathbf{f_a}(L-2))^{-1} \cdot (\mathbf{f_a}(L-2) + \mathbf{f_a}(L-1) \cdot z^{-1})$$

$$\mathbf{L}_i^{-1}(z) = (\mathbf{f_s}(L-2) + \mathbf{f_s}(L-1) \cdot z^{-1}) \cdot (\mathbf{f_s}(L-2))^{-1} \cdot \mathbf{J}$$

then

$$\mathbf{F_a}(z) \cdot \mathbf{L}_i^{-1}(z) \ , \ \mathbf{L}_i(z) \cdot \mathbf{F_s}(z)$$

have a length factor L reduced by 1, d is unchanged. Here i is the iteration index. It has the reverse order of the analysis cascade, i.e. it starts with $i = \nu$ and is reduced for each step of the iteration. The matrix $\mathbf{L}_i(z)$ has the form of $\mathbf{E}_i(z)$ or $\mathbf{G}_i(z)$, depending on whether the non-zero elements of $\mathbf{f_a}(L-1)$ are on the right or left side. The reduced $\mathbf{F_a}$ and $\mathbf{F_s}$ are again a filter matrix of a cosine modulated filter bank. They have the same form as $\mathbf{F_a}$ and $\mathbf{F_s}$ since $\mathbf{L}_i$ has a bi-diagonal form. They again result in an FIR filter bank with FIR inverse since $\mathbf{L}_i$ has an FIR inverse. The condition for a further reduction of length L is that the reduced $\mathbf{F_a}$ and $\mathbf{F_s}$ lead again to contiguous baseband impulse responses, in order to obtain the next $\mathbf{f_a}(L-2)$ and $\mathbf{f_s}(L-2)$ invertible. This is usually the case. If not, the objectionable zeros in it can be replaced by some small ϵ. This process of reducing L of $\mathbf{F_a}$ and $\mathbf{F_s}$ can be continued until $2L-3 \leq d$. This way the zero delay matrices are obtained.

Then the same can be done for the other side, the beginning of the impulse response. Observe that $n_a > 0$, $n_s > 0$, and causality results in $\mathrm{rank}(\mathbf{f_a}(0)) < N$ and $\mathrm{rank}(\mathbf{f_s}(0)) < N$, so that $d > 1$ at the start of the iteration. If $d > 1$ then for $m = 0$

$$\mathbf{f_a}(0) \cdot \mathbf{f_s}(0) = \mathbf{0} \quad (5)$$

and for $m = 1$

$$\mathbf{f_a}(0) \cdot \mathbf{f_s}(1) + \mathbf{f_a}(1) \cdot \mathbf{f_s}(0) = \mathbf{0} \quad (6)$$

Here it can be concluded that $\mathrm{rank}(\mathbf{f_a}(0))$= $\mathrm{rank}(\mathbf{f_a}(0)) \leq N/2$ and $\mathbf{f_a}(1), \mathbf{f_a}(1)$ have full rank. Define

$$\mathbf{H}_i(z) := \mathbf{J} \cdot (\mathbf{f_a}(1))^{-1}(\mathbf{f_a}(0) + \mathbf{f_a}(1) \cdot z^{-1}),$$

$$\mathbf{H}_i^{-1}(z) = (\mathbf{f_s}(0) + \mathbf{f_s}(1) \cdot z^{-1}) \cdot (\mathbf{f_s}(1))^{-1} \cdot \mathbf{J} \cdot z^2,$$

then

$$\mathbf{F_a}(z) \cdot \mathbf{H}_i^{-1}(z) \ , \ z^2 \cdot \mathbf{H}_i(z) \cdot \mathbf{F_s}(z)$$

are causal, with length L reduced by 1, d reduced by 2. The iteration starts with $i = \mu - 1$. The matrix $\mathbf{H}_i(z)$ has the form of $\mathbf{A}_i(z)$ or $\mathbf{B}_i(z)$, depending on whether the non-zero elements of $\mathbf{f_a}(0)$ are on the right or left side. As above the resulting matrices are again filter matrices of an FIR cosine modulated filter bank. This process of reducing the length of $\mathbf{F_a}$ and $\mathbf{F_s}$ is continued until $L = 2$. This way the maximum delay matrices are obtained. The matrices which are left are $\mathbf{B}_0(z)$ and $\mathbf{B}_0^{-1}(z)$.

5 Symmetries

In many applications it is desirable to have identical magnitude responses for the analysis and synthesis filters. This is always the case for orthogonal filter banks, where analysis and synthesis filters are time reversed versions of each other. It is in general not the case for bi-orthogonal filters. But it is shown that the presented filter bank can be designed such that it has this property by imposing some constraints, even in the case of a low system delay. Identical magnitude responses are obtained if the baseband impulse responses for analysis and synthesis are identical, except for the sign, $h'(n) = s \cdot h(n)$ where $s = 1$ or $s = -1$. Now $\mathbf{F_s}(z)$ is the synthesis filter matrix for perfect reconstruction if

$$P_i'(z) = \frac{s \cdot z^{-d} P_i(z)}{z^{-2} P_{N+i}(z) P_{2N-1-i}(z) - P_i(z) P_{N-1-i}(z)} \tag{7}$$

$$P_{N+i}'(z) = \frac{s \cdot z^{-d} P_{N+i}(z)}{z^{-2} P_{N+i}(z) P_{2N-1-i}(z) - P_i(z) P_{N-1-i}(z)} \tag{8}$$

for $i = 0, \ldots, N-1$.

It follows that $h'(n) = s \cdot h(n)$ if $P_i'(z) = s \cdot P_i(z)$ for $i = 0, \ldots, 2N-1$. Eq. 7, 8 show that this is the case if

$$z^{-2} P_{N+i}(z) P_{2N-1-i}(z) - P_i(z) P_{N-1-i}(z) = s \cdot z^{-d}$$

for $i = 0, \ldots, N-1$ and some integer d. The left side of this condition is like the determinant of a matrix of the elements at these 4 positions. It is multiplicative, i.e. if it is true for two matrices it is also true for their product. If this condition is fulfilled by the simple matrices by which $\mathbf{F_a}$ is constructed, it is true for $\mathbf{F_a}$. It is easy to see that the matrices $\mathbf{H}_i(z)$ and $\mathbf{L}_i(z)$ already fulfill this condition. $\mathbf{B}_0(z)$ can be designed such that it fulfills the condition if $b^0_{N+i} = s/b^0_{2N-1-i}$ for $i = 0, \ldots, N-1$. This condition can also be used to simplify the optimization since e.g. only the analysis filter matrix needs to be optimized.

Example

Figure 1 shows an example of a filter bank with a low system delay, compared with an orthogonal filter bank with a standard system delay. The parameters of the low delay filter bank are $n_a = n_s = N/2$, $\nu = 3$, $\mu = 1$, and the symmetry condition for identical magnitude responses for analysis and synthesis was imposed. The resulting cascade or structure was optimized with the optimization algorithm described in [5, 6, 7]. Both filter banks have 128 bands and a system delay of 255 samples, but the orthogonal filter bank is restricted to a filter length of 256 taps due to the given system delay. The low delay filter bank has a filter length of 512 taps, and as a result has a much higher stopband attenuation, as can be seen.

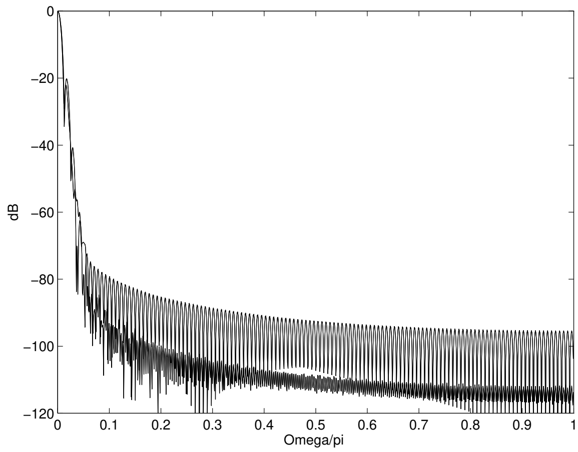

Figure 1. Magnitude responses of the baseband low delay prototype (the lower curve), compared with an orthogonal filter bank (upper curve)

References

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

[2] H. S. Malvar: "Extended Lapped Transforms: Properties, Applications, and Fast Algorithms", IEEE Transactions on Signal Processing, Vol.40, NO.11, Nov. 1992.

[3] K. Nayebi, T.P. Barnwell,III, M.J.T. Smith: "Low Delay FIR Filter Banks: Design and Evaluation", Trans. on Signal Processing, January 1994, pp.24-31.

[4] G. Schuller and M. J. T. Smith,"Efficient Low Delay Filter Banks", DSP Workshop, Oct. 1994.

[5] G. Schuller and M. J. T. Smith, "A New Algorithm for Efficient Low Delay Filter Bank Design" ICASSP 95, Detroit, MI, May 1995.

[6] G. Schuller: "A Low Delay Filter Bank for Audio Coding with Reduced Pre-Echoes", 99th AES Convention, New York, New York, October 6-9, 1995

[7] G.D.T. Schuller and M. J. T. Smith, "New Framework for Modulated Perfect Reconstruction Filter Banks", IEEE Transactions on Signal Processing, Vol.44, NO.8, August 1996

[8] G. Schuller, "Low Delay Filter Banks with Perfect Reconstruction", Frequenz, Sep.-Oct. 1996