J. Indian Inst. Sci., Mar.–Apr. 1999, 79, 125–132.

Projection on convex sets approach to design of near-equiripple NPR banks

S. JAYASIMHA
Signion Systems Ltd, 20 Rockdale Compound, Somajiguda, Hyderabad 500 082.
http://www.signion.com

Abstract

A projection on convex sets (POCS) approach to design an M-channel near-perfect reconstruction (NPR) banks is described. The method is conceptually simpler than previously reported methods and allows the design of high-stopband filterbanks.

Keywords: Approximation methods, channel bank filters, convergence of numerical methods, equiripple filters, iterative methods, Nyquist filters, quadrature mirror filters, signal sampling/reconstruction.

1. Introduction

The theory of pseudo-QMF banks, their applications, and an optimization procedure to design a prototype filter is described in Vaidyanathan.[1] The main disadvantage of the spectral factorization approach to design of equiripple prototypes,[2] large reconstruction errors in the neighborhood of $\omega = 0$ and $\omega = \pi$, is overcome by the method of Jayasimha and Hiremath.[3] Nguyen[4] describes a constrained nonlinear minimization algorithm to optimize a quadratic objective function to obtain a prototype that is very close to being a spectral factor of a $2M$th-band filter. Such designs are termed near-perfect reconstruction (NPR) designs, where the reconstruction error is of the order of $\sqrt{M}$ times the stopband error. However, in this NPR design procedure,[4] the maximum stopband ripple is in the immediate vicinity of the transition band. In many applications, where the maximum stopband ripple is an important measure of a filter's selectivity (in particular, those applications that require cascaded sub-band decompositions), the procedure described here yields better results. The method is based on projection on convex sets (POCS);[6] it does not use the Remez-exchange equiripple design procedure used in the earlier method.[3] The POCS method[6] has two important advantages compared to the Parks–McClellan procedure.[5]

1. The NPR constraint can be incorporated into the POCS scheme, provided its projection into the chosen set (time or frequency) is convex.

2. The method can be extended to design multidimensional (non-separable) filters.

An added advantage of the POCS method (compared to the method earlier reported[3]) is that both odd- and even-length prototypes (corresponding to DCT-II and -IV modulations) can be designed, whereas the previous design method[3] obtains only odd-length prototypes. Another method[7] to obtain weighted (in the stopband) minimax M-channel filterbanks is more complex than the proposed method.

A method to incorporate the constraint that the prototype be a spectral factor of a $2M$th-band filter into the POCS iteration scheme is described. In doing so, equiripple NPR designs are obtained (rather than the pseudo-QMF designs[3]). The design method is illustrated through an example.

2. Frequency domain specification of initial linear-phase prototype

A filter whose impulse response, $g(n)$, satisfies:

$$g(n)=\begin{cases}0, & n=(2N-2)/2+2pM \quad (p=\pm 1,\pm 2,\ldots)\\ \dfrac{1}{2M}, & n=(2N-2)/2\end{cases} \tag{1}$$

is called a $2M$th-band filter. Its frequency response, $G(e^{j\omega})$, satisfies:[1]

$$\sum_{k=0}^{2M-1} G(e^{j(\omega+\pi k/M)})=2Mg(0)=1. \tag{2}$$

Let $H(e^{j\omega})$ be a equiripple zero-phase low-pass spectral factor of a $2M$th-band FIR filter of length N, i.e. $G(e^{j\omega})=H^2(e^{j\omega})$ is also an equiripple zero-phase low-pass $2M$th-band FIR filter of length $2N$-1, whose impulse response satisfies (1). Since $H(e^{j\omega})$ is zero-phase, it takes only real values for $0 \le \omega \le \pi$, and $G(e^{j\omega})$ takes only non-negative real values.

Define $H(e^{j\omega})$ as an equiripple low-pass filter with passband ripple δ_p and stopband ripple δ_s. Then (see, for example, Fig. 1, where $M=2$), $G'(e^{j\omega})=H'^2(e^{j\omega})-\delta_s^2/2$, with passband ripple $=2\delta_p$ and stopband ripple $=\delta_s^2/2$ is a zero-phase equiripple filter that can be made approximately $2M$th-band by selecting:

(3)

$$\omega_p+\omega_s=\frac{\quad}{2M}. \tag{4}$$

Now,

$$G(e^{j\omega})=\left[G'(e^{j\omega})+\frac{\delta_s^2}{2}\right]\frac{1/2M}{(1/2M)+\left(\delta_s^2/2\right)}=H'^2(e^{j\omega})\frac{1/2M}{(1/2M)+\left(\delta_s^2/2\right)} \tag{5}$$

is a $2M$th-band filter with non-negative stopband ripple. Therefore,

$$H(e^{j\omega})=H'(e^{j\omega})\sqrt{\frac{1/2M}{(1/2M)+\left(\delta_s^2/2\right)}} \tag{6}$$

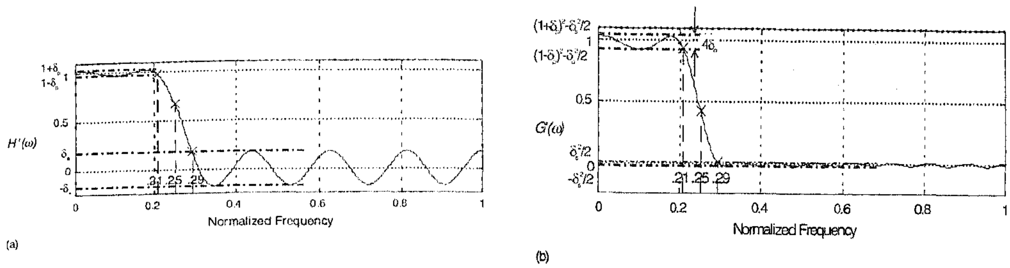

FIG. 1. Specification of (a) $H'(\omega)$ such that (b) $G'(\omega)$ is approximately an equiripple 4th-band filter.

is a linear-phase spectral factor of $G(e^{j\omega})$. The NPR design problem is to obtain an $H'(e^{j\omega})$ such that (2) is satisfied (to within a small error).

3. Iterative design by POCS

In the 1-dimensional case, the initial filter's time-domain coefficients can be obtained using Parks–McClellan procedure[5] employing pass- and stopband weights according to (3) (described earlier in detail[3]). In the multidimensional case, the initial filter is the inverse Fourier transform of the desired frequency response. Note from (3) that the error in meeting (2) for $0 < \omega < \omega_p$ is of the order of δ_s^2. If δ_s is small, then most of the error in meeting (2) is due to transition band mismatch. The idea is to correct this mismatch using, ideally, a band (to the transition band) and a time-limited (to the support of the initial filter) correction sequence applied in frequency domain. However, this ideal cannot be met by the first iteration's correction sequence; during subsequent iterations this ideal is met to an increasing degree of accuracy and the desired filter results. The frequency response, $H'(e^{j\omega})$, of the zero-phase filter is required to be within prescribed upper and lower bounds in its pass- and stopbands as follows:

$$H'_{id}(e^{j\omega}) - E_d(\omega) \le H'(e^{j\omega}) \le H'_{id}(e^{j\omega}) - E_d(\omega) \quad \omega \,\varepsilon\, F_r \tag{8}$$

where $H'_{id}(e^{j\omega})$ is the ideal filter response, $E_d(\omega)$a positive function of ω, which takes values in F_r, where F_r is the union of the pass- and stopbands. Usually, for the NPR prototype,

$$H_{id}(\omega) = \begin{cases} 1, \text{ if } \omega \in F_p \\ 0, \text{ if } \omega \in F_s \end{cases} \tag{9}$$

where F_p and F_s are the pass- and stopbands, respectively, and

$$E_d(\omega) = \begin{cases} \delta_p, \text{ if } \omega \in F_p \\ \delta_s, \text{ if } \omega \in F_s \end{cases} \tag{10}$$

where δ_p and δ_s satisfy (3).

The iterative method begins by the initial prototype, $h'_0(n)$, which is the inverse Fourier transform of $H_{id}(\omega)$, $h_{id}(n) = \mathcal{F}^{-1}[H_{id}(\omega)]$ truncated to support I:

$$h'_0(n) = \begin{cases} h_{id}(n), & \text{if } n \in I \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Each iteration consists of applying successive temporal and frequency-domain constraints to the current iterate. The kth iterate consists of the following steps:

- Compute the zero-phase frequency response of the kth iterate $h_k(n)$ on a suitably dense grid of frequencies using FFT algorithm.
- Impose frequency-domain constraints as follows:

  (a) the equiripple constraint in the pass- and stopbands consistent with (3). However, if the initial filter already conforms to (3) (using Parks–McClellan procedure[5] of Jayasimha and Hiremath[3]), the equiripple constraint need not be applied to the passband as, in many applications, the requirement is to obtain NPR prototypes with a prescribed stopband ripple.

  (b) the $2M$th-band spectral factor constraint in the transition band using a 'smooth' error allocation function for symmetrically paired (about $\pi/2M$) frequency components. One choice of allocation is the complement of the standard sigmoid, $f(z) = 1 - (1/1 + e^{-az})$, where z is a distance measured from the center of the transition band ($\pi/2M$) and a a parameter that determines the rise rate of the sigmoid. Another function that behaves similarly is the complement of the inverse tangent function, $g(z) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1} az$. The choice of such functions is predicated by the desire to allocate more error (of the order of δ_s^2) to frequency components near the passband edge (where the ripple is of the order of δ_s^2 to begin with) and the desire not to disturb the relative frequency responses of adjacent frequency bins near the transition band edges.
- Compute the inverse Fourier transform of $H'_{k+1}(e^{j\omega})$.
- Zero out the inverse Fourier transform outside the support I to obtain $h'_{k+1}(n)$.
- Exit if a convergence criterion between the iterates $h'_k(n)$ and $h'_{k+1}(n)$ is met. The choice of a suitable convergence criterion is provided in the next section.

The flow diagram of this method is shown in Fig. 2. It can be proven that the iterative FIR filter design program is globally convergent. The proof is based on POCS.[8, 9] The frequency and time-domain constraints define two convex sets in the set of square summable sequences and the imposition of frequency and time-domain constraints are orthogonal projections on to these sets. If the sets intersect, then the iterates converge to a member in the intersection set. If the specifications are too tight, then the two sets do not intersect and the algorithm does not converge. In such a case, either the filter support is progressively enlarged or the stopband specification is progressively relaxed. In one-dimensional cases, an empirical formula[10] relates the filter support to pass- and stopband specifications, without the spectral factor of $2M$th-band filter constraint. The latter constraint does not appreciably change the formulae.[3] The iterative algorithm is implemented using a power-of-two size FFT that is greater than 10 times half the chosen support.

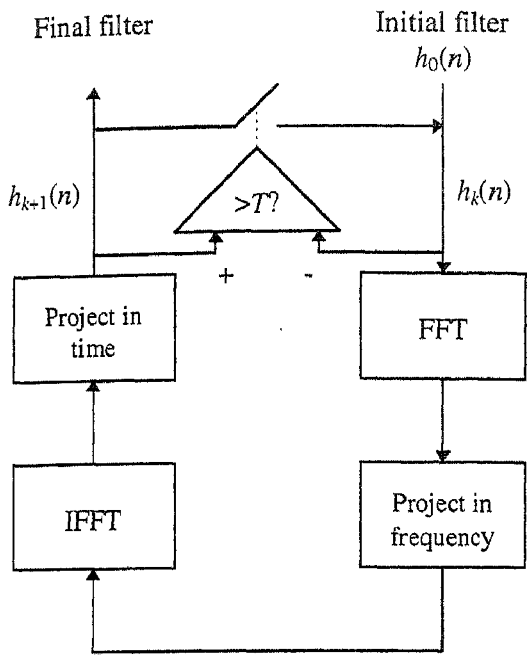

FIG. 2. Flow diagram of the iterative design.

4. Filter design example

The design procedure is illustrated for an NPR 32-band example prototype filter $H'(z)$ with the following specifications (where normalized frequency is defined as ω/π):

$M = 32$

Passband cutoff frequency (normalized) $= \dfrac{1}{64} - \dfrac{3}{256}$

Stopband cutoff frequency (normalized) $= \dfrac{1}{64} + \dfrac{3}{256}$

Stopband ripple $\delta_s = -80\text{dB}$

Number of taps = 511

The search procedure described in Section 2 of the previous method[3] was used to find the optimum relative stop-to-passband weight, K, to obtain an initial approximate spectral factor of a $2M$th-band filter. K is $12.6624\delta_s$, and the peak overall distortion[a] is –39.903dB, while the actual stopband ripple is –77.15dB. The evolution of the designed parameters, the maximum stopband gain, the peak overall distortion and the maximum deviation from the $2M$th-band condition, using the procedure shown in Fig. 2 (with the convergence criterion disabled) are graphed in Fig. 3.

[a]The peak overall distortion is the maximum deviation of $T_0(e^{i\omega})$ from unity.

$T_0(z) = \dfrac{z^{-(N-1)}}{M}\sum_{0}^{M-1} H_k(z)\tilde{H}_k(z)$. $H_k(z)$ are M cosine modulated filters obtained from the prototype $H(z)$. $\tilde{H}_k(z)$ are time reversals of the real $H_k(z)$.[4]

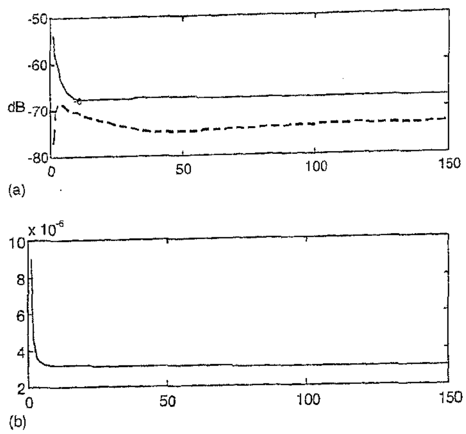

FIG. 3(a). Evolution of maximum overall distortion (solid line) and stopband ripple (dashed line) with iteration index, and (b) Evolution of deviation from $2M$th-band condition with iteration index.

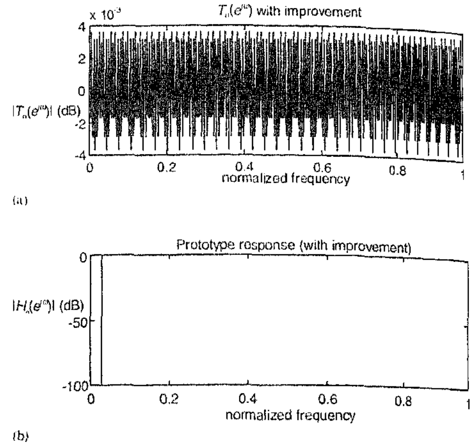

FIG. 4. Overall distortion and prototype response for design example (the vertical line in (b) indicates the stopband edge).

It can be observed from Fig. 3 that convergence is quite rapid for all three design parameters initially. However, since the constraint convex sets do not intersect[b], there will be a trade-off between stopband ripple and deviation from the $2M$th-band condition. In Fig. 3, with a (in the standard sigmoid) chosen as 4, the trade-off favors meeting the $2M$th-band condition at the expense of stopband performance. Thus, a suitable convergence criterion is met when stopband ripple no longer decreases with iteration index and reaches a constant value.

Figure 4 shows overall distortion and magnitude response of a prototype filter designed using this stopping criterion (see appendix for MALTAB listing). The maximum stopband ripple is -74.87dB, only 2.25 dB worse than the original prototype filter. However, the worst case reconstruction error has improved from -39.903dB to -67.45dB.

5. Conclusion

A simple method of designing filterbanks using the FFT algorithm and projection on to the constraint convex sets is presented. One contribution consists of modifying the frequency domain projection[6] to meet the $2M$th-band spectral factor constraint. The other contribution consists of applying constraint corrections to the prototype in such a manner that they are largely band-limited to the transition band.

References

1. VAIDYANATHAN, P. P. *Multirate systems and filterbanks*, Prentice–Hall, 1993, pp 353–370.
2. KOILPILLAI, R. D. AND VAIDYANATHAN, P. P. A spectral factorization approach to pseudo-QMF design, *IEEE Trans.*, 1993, **SP-41**, 82–92.

[b]This is because the stopband performance constraint of the initial equiripple filter (which only approximately embodies the $2M$th-band constraint) is applied (while the POCS iteration scheme imposes the $2M$th-band constraint exactly). The penalty in stopband performance due to this additional constraint is not known.

3. JAYASIMHA, S. AND HIREMATH, C. G. Pseudo-QMF banks with near-equiripple performance, *IEEE Trans.*, 1998, **SP-46**, 209–214.

4. NGUYEN, T. Q. Near-perfect-reconstruction of pseudo-QMF banks, *IEEE Trans.*, 1994, **SP-42**, 65–76.

5. MCCLELLAN, J. H., PARKS, T. W. AND RABINER, L. R. A computer program for designing optimum FIR linear phase digital filters, *IEEE Trans.*, 1973, **AU-21**, 506–526.

6. CETIN, A. E., GEREK, O. N. AND YARDIMCI, Y. Equiripple FIR filter design by the FFT algorithm, *IEEE Signal Processing Mag.*, Mar. 1997, 60–64.

7. GOH, C-K. AND LIM, Y-C. An efficient algorithm for the design of weighted minimax *M*-channel cosine-modulated filterbanks, *IEEE Trans.*, 1998, **SP-46**, 1426–1430.

8. COMBETTES, P. L. The foundations of set theoretic estimation, *Proc. IEEE*, 1993, **81**, 182–208.

9. CIVANLAR, M. R. AND NOBAKHT, R. A. Optimal pulse shape design using projection onto convex sets, *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, New York, NY, 1988, pp. 1874–1877.

10. RABINER, L. R. AND GOLD, B. *Theory and applications of digital signal processing*, Prentice-Hall, 1988, pp. 156–157.

Appendix I
MATLAB code for design example

```
% Design a 511-length 32-band equiripple filterbank
% using POCS to improve
clear;
hold off;
M=32;
N=510;
ds=1e-4;
delta=3/256;
fftN=16384;
f=[0 (1/(2*M))-delta (1/(2*M))+delta 1];
m=[1 1 0 0];
wt=12.6624;
b=remez(N,[0 1/(2*M)-delta 1/(2*M)+delta 1],[1 1 0 0],[1
wt*ds]);
factor=sqrt((1/(2*M))/((1/(2*M))+(ds*ds/2)));
b=b*factor;
c=conv(b,b);
disp('Maximum deviation from the 2Mth band condition');
disp(max(abs(c(N+2*M+1:2*M:2*N+1))))
figure(1)
[rerr1,mag1]=plerr(wt,2*M,N,b);
mg1=max(abs(abs(mag1)-
1))*sqrt(((1/(4*M))/((1/(4*M))+((ds^2)/2))));
rcerr1=20*log10(mg1);
disp('Initial Reconstruction error (dB)');
disp(rcerr1);
title('T0(z) without improvement');
xlabel('normalized frequency');
ylabel('|T0(z)| (dB)');
% find actual stopband ripple
theta=2*pi*(0:fftN-1)*N/(2*fftN);
x=real(fft(b,fftN).*exp(i*theta));
sb_edge=fftN/(4*M)+fftN/(2*(1/delta));
ds=max(abs(x(sb_edge:fftN-sb_edge)));
for l=1:150;
        %Calculate zero phase frequency response
        theta=2*pi*(0:fftN-1)*N/(2*fftN);
        x=real(fft(b,fftN).*exp(i*theta));
        disp('Maximum stopband ripple (dB):');
        sb_rip(l)=20*log10(max(abs(x(sb_edge:fftN-
                sb_edge))));
        disp(sb_rip(l));
        if (l>20)
                if (sb_rip(l)>sb_rip(l-1)) break;
                end;
        end;
        % impose equiripple constraint in stopband
        x(sb_edge:fftN-sb_edge)=min(x(sb_edge:fftN-
                        sb_edge),ds);
        x(sb_edge:fftN-sb_edge)=max(x(sb_edge:fftN-
                        sb_edge),-ds);
        % Impose the 2Mth band constraint on the first
        subband
        % (the passband) and its image. Note that for
        2Mth band
```

```
% condition sum_over_k H(e**(w+2*pi*k/M))=
1.
err(1:fftN/(2*M))=sum((reshape
            (x.^2,fftN/(2*M),2*M))')-1;
% dump all error into passband
pb_edge=fftN/(4*M)-fftN/(2*(1/delta));
x(1:pb_edge)=sqrt(x(1:pb_edge).^2-
err(1:pb_edge));
% split up errors in transition band
tbm=1+fftN/(4*M);
tbl=tbm-fftN/(2*(1/delta));
tbh=tbm+fftN/(2*(1/delta));
factor=(0.5-((tbl:tbh)-tbm)/(tbh-tbl)); % linear
factor
factor=1./(1+exp(-exp(4)*(factor-0.5)));
                        % sigmoidal factor
x(tbl:tbh)=sqrt(x(tbl:tbh).^2-factor.*err(tbl:tbh));
% dump all error into first subband
(pass+transition)
% x(1:1+fftN/(4*M)) =
sqrt(x(1:1+fftN/(4*M)).^2-
      err(1:1+fftN/(4*M)));
x(fftN:-1:fftN/2+2)= x(2:fftN/2);
% Calculate the new time domain impulse re-
sponse.
r=real(ifft(x.*exp(-i*theta)));
b1=r(1:N+1);
c1=conv(b1,b1);
disp('deviation from the 2Mth band condition');
max_dev(l)=max(abs(c1(N+2*M+1:2*M:2*N+1)
));
disp(max_dev(l));
subplot(2,1,1);
[rerr1,mag1]=plerr(wt,2*M,N,b1);
mg1=max(abs(abs(mag1)-
1))*sqrt(((1/(4*M))/((1/(4*M))+((ds^2)/2))));
    rcerr(l)=20*log10(mg1);
    disp('Reconstruction error with improvement
    (dB)');
    disp(rcerr(l));
    title('T0(z) with improvement');
    xlabel('normalized frequency');
    ylabel('|T0(z)| (dB)');
    b=b1;
    subplot(2,1,2);
    x1=2*(0:fftN-1)/fftN;
    y1=20*log10(abs(fft(b,fftN)));
    y1=min(y1,1);
    y1=max(y1,-100);
    plot(x1(1:fftN/2), y1(1:fftN/2));
    hold on;
    x1=(1/(2*M))+(1/(1/delta));
    x1=[x1 x1];
    y1=[-100 1];
    plot(x1,y1,'b');
    hold off;
    axis([0 1 -100 1]);
    title('Prototype response (with improvement)');
    xlabel('normalized frequency');
    ylabel('|Hp(z)| (dB)');
    end;
figure(2);
subplot(2,1,1);
plot(rcerr);
hold on;
plot(sb_rip, '--');
xlabel('Maximum distortion and stopband ripple');
ylabel('dB');
hold off;
subplot(2,1,2);
plot(max_dev);
xlabel('Deviation from 2Mth-band condition');
```