Proc. 1998 Eureopean Signal Processing Conference (EUSIPCO '98)
Island of Rhodes, Greece, September 1998

QUADRATIC FIR-FILTER DESIGN AS A GENERALIZED EIGENPROBLEM

*Jeffrey O. Coleman**
Naval Research Laboratory
Washington DC 20375, USA
jeffc@alum.mit.edu

ABSTRACT

This work suggests optimizing quadratic performance measures on (possibly multi-rate) DSP systems in order to design one or more embedded FIR filters, which might be real or complex and with or without linear-phase, N-th band, or other structure. The outlined approach replaces the eigenproblem of the eigenfilter design method with a generalized eigenproblem. A general capability results for minimizing an arbitrary nonnegative-definite (purely) quadratic system-performance measure subject to a single arbitrary equality constraint on a second such quadratic. Using this one constraint to fix the impulse-response energy in nonembedded single-filter design results in the ordinary eigenfilter method. A simple example and a geometric interpretation show that alternate equality constraints can profoundly affect the solution. Finally a multirate matched-filtering-system example design impossible with the ordinary eigenfilter method hints at the flexibility of the approach.

1 INTRODUCTION

The eigenfilter method of FIR-filter optimization of Vaidyanathan and Nguyen [1] and others [2–4] associates the desired optimum filter with the minimum-eigenvalue eigenvector of a positive-definite matrix whose structure embodies the filter specifications. Replacing the eigenproblem involved with a generalized eigenproblem provides a fundamentally more-powerful approach [5] by generalizing the equality constraint inherent in these methods. This paper reviews the key ideas, in particular the applicability of the generalized approach to embedded filters, and then examines the significance of the generalized equality constraint.

Both eigenfilter approaches derive conceptually from quadratic optimization, which is often simple to think about in a test-input framework [5–8]. Consider the two

*Graduate-student involvement in this work was supported by National Science Foundation grant MIP-9896034 to the University of Maryland Baltimore County. Early portions of this work were performed while the author was with Michigan Technological University and were partially supported by National Science Foundation grant MIP-9409686 and Naval Research Laboratory contract N00014-97-P-2030.

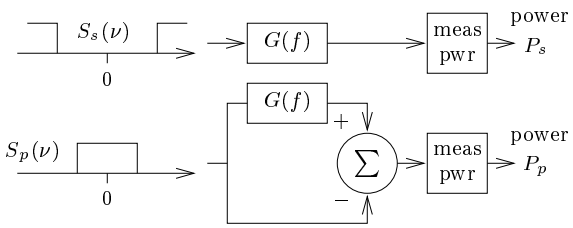

Figure 1: Optimization of lowpass filter $G(f)$ minimizes these measured average powers.

Fig. 1 "test systems" for measuring output powers that can be forced small during filter optimization. Requiring P_s to be much smaller than the power in the input random process forces the filter to reject signals in a stopband corresponding to the support of power spectrum $S_s(f)$, and likewise a small P_p creates a passband by forcing the upper and lower sum inputs to be similar for input spectrum $S_p(f)$. (These powers are weighted L_2 frequency-domain approximation errors.) The next two sections show that for a wide variety of FIR filters embedded in a wide variety of test systems such output powers can be minimized in quadratic optimizations formulated as a generalized eigenproblems.

2 TEST-SYSTEM MATHEMATICS

Concise notation simplifies the presentation. Sequence $h \star g$ is the convolution of (scalar, vector, or matrix) sequences h and g, and $(h \star g)(n)$ is the value of that sequence at time n. Prime $'$ denotes the "match" operation of matched filtering, comprising Hermitian transpose and time reversal, so that $\mathbf{w}'(n) = \mathbf{w}^H(-n)$.

Let vector $\mathbf{x}$ contain the real optimization variables, and require any test-system impulse response $h(n)$ dependent on $\mathbf{x}$ to be linear in $\mathbf{x}$, so that $h(n) = \mathbf{x}^T\mathbf{h}(n)$, where $\mathbf{h}(n)$ is the *vector representation sequence* for $h(n)$. Consider the simplest case: If $h(n)$ is FIR with every nonzero element a different real optimization variable, then each corresponding vector in sequence $\mathbf{h}(n)$ is just a different standard unit vector (one nonzero element). From such finite, real, nonlinear-phase sequences it is simple to build responses with more structure. For example, if vector sequences $\mathbf{u}(n)$ and $\mathbf{v}(n)$

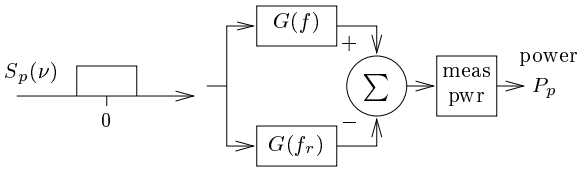

Figure 2: This reference-path gain makes the test-system impulse response linear in the filter weights.

have identical support but with distinct unit vectors, sequence $\mathbf{h}(n) = \mathbf{u}(n) + j\mathbf{v}(n)$, or the notationally simpler $\mathbf{h} = \mathbf{u} + j\mathbf{v}$, then represents an arbitrary complex scalar FIR impulse response $g(n) = \mathbf{x}^T(\mathbf{u} + j\mathbf{v})(n)$.

Using identities that relate operations on sequences to operations on their associated vector representation sequences, it is simple to construct vector representation sequences for real or complex impulse responses, either with or without linear phase, and if needed with N-th band or other structural properties. And these same identities permit straightforward construction of vector representation sequences for the impulse responses of entire test systems. Identities are available for sequence addition, scaling, delay, matching (the prime $'$ operator), interpolation, decimation, catenation, convolution with a fixed response, and Fourier transform.

The latter is exceptional in that the transform of sequence $h(n) = \mathbf{x}^T\mathbf{a}(n)$ is a scalar function of frequency $\mathbf{x}^T\mathbf{A}(f)$, where $\mathbf{A}(f) = \sum_n \mathbf{a}(n)e^{-j2\pi fn}$ is just the elementwise Fourier transform of vector representation sequence $\mathbf{a}$(n). (Uppercase $\mathbf{A}(f)$ refers here to a vector, not a matrix.) Here vector $\mathbf{A}(f)$ is not a vector representation sequence, since it is not a sequence at all.

Once impulse responses are in the standard form, output power is straightforwardly related to optimization vector $\mathbf{x}$. Suppose $z(n)$ is a wide-sense-stationary, zero-mean complex input process, and suppose $g(n)$ is a complex test-system impulse response. Further suppose that $g(n) = \mathbf{x}^T\mathbf{g}(n)$ in terms of vector representation sequence $\mathbf{g}$. Then output $u = z \star g$ is a wide-sense stationary zero-mean complex process, and the associated output power $P_u = E|u|^2 = R_u(0)$ or

$$P_u = \mathbf{x}^T \mathrm{Re}\{(\mathbf{g} \star R_z \star \mathbf{g}')(0)\}\mathbf{x}. \qquad (1)$$

Only a finite number of autocorrelation samples are required in practice, as the finite length of sequence $\mathbf{g}(n)$ results in finite convolution sums. The autocorrelation samples are easily computed from the specified power spectrum using FFT techniques [7].

3 THE TWO EIGENFILTER METHODS

Recall that ratio α/β and vector ψ are a generalized eigenvalue and the associated generalized eigenvector respectively for the pair of real symmetric matrices $(\mathbf{Q}, \mathbf{V})$ if $\beta\mathbf{Q}\psi = \alpha\mathbf{V}\psi$. If $\beta = 0$, the generalized eigenvalue is infinite. A version of the Rayleigh principle states that the vector $\mathbf{x}$ that minimizes quadratic form $\mathbf{x}^T\mathbf{Q}\mathbf{x}$ subject to the requirement that $\mathbf{x}^T\mathbf{V}\mathbf{x} = 1$ is the generalized eigenvector associated with the minimum generalized eigenvalue of nonnegative definite matrices $(\mathbf{Q}, \mathbf{V})$. Thus we have the opportunity to minimize one quadratic form while constraining another.

The quadradic objective minimized is generally some linear combination of (1) output powers in response to random-process test-system inputs and (2) output energies in response to deterministic test-system inputs. The constrained quadratic is commonly the magnitude squared of a single frequency-response sample or time-response sample in a test system. This has a singular kernel, but nonsingular kernels are also possible. For example, fixing the filter output power in response to a bandlimited input process fixes the mean-square magnitude response over some band (perhaps a passband).

The real catch is the same here as in all the classic developments of the eigenfilter method: In order to obtain a quadratic-form objective as a linear combination of output powers and energies, each test-system impulse response has to be of form $\mathbf{x}^T\mathbf{a}(n)$, linear in the elements of $\mathbf{x}$. An impulse response, such as that of the lower system of Fig. 1, that is of form $\mathbf{x}^T\mathbf{u}(n) - d(n)$ including a constant term (not a function of $\mathbf{x}$) is not permissible. To eliminate such a constant term, modify the test system to multiply $d(n)$ in the impulse response by some weighting factor $\mathbf{x}^T\mathbf{w}$. Choosing this factor according to the individual design problem requires some care. The test-system impulse response then becomes $\mathbf{x}^T(\mathbf{u}(n) - d(n)\mathbf{w})$, and so $\mathbf{a}(n) = \mathbf{u}(n) - d(n)\mathbf{w}$ in (1). Eigenfilter methods in the literature commonly include a linearizing weighting factor $G(f_r) = \mathbf{x}^T\mathbf{H}(f_r)$ in passband-design reference paths as in Fig. 2.

The ordinary eigenfilter method is as above but specifically with the equality constraint $\|h\|^2 = \|\mathbf{x}\|^2 = 1$. General strategies for embedded ordinary-eigenfilter design are not seen in the literature, as it is difficult to relate the impulse response to the optimization variables so that the first of these equalities is met.

4 EXAMPLES, SIMPLE AND NONSIMPLE

Start with a length-two optimization vector $\mathbf{x}$ and a length-three linear-phase filter impulse response $h(n) = \mathbf{x}^T(\frac{1}{\sqrt{2}}\mathbf{e}_2\delta(n+1) + \mathbf{e}_1\delta(n) + \frac{1}{\sqrt{2}}\mathbf{e}_2\delta(n-1))$ that is to be optimized, where $\mathbf{e}_i$ is the standard unit vector with unity in the i-th position. The $\frac{1}{\sqrt{2}}$ scale factors are a convenience to make the quadratic that is implicitly constrained in the ordinary eigenfilter method just $\|h\|^2 = \|\mathbf{x}\|^2$, corresponding to the circle in Fig. 3.

In this the ultimately simple, single-test-setup example, a random process with symmetric lowpass spectrum of unit-height spectral-density over $f \in [-.25, .25]$ is input through filter $h(n)$. The output power is the quadratic function of the optimization vector $\mathbf{x}$ indicated by the elliptic contours in Fig. 3 and is minimized to create an optimal $[-.25, .25]$ stopband.

Design with the ordinary eigenfilter method mini-

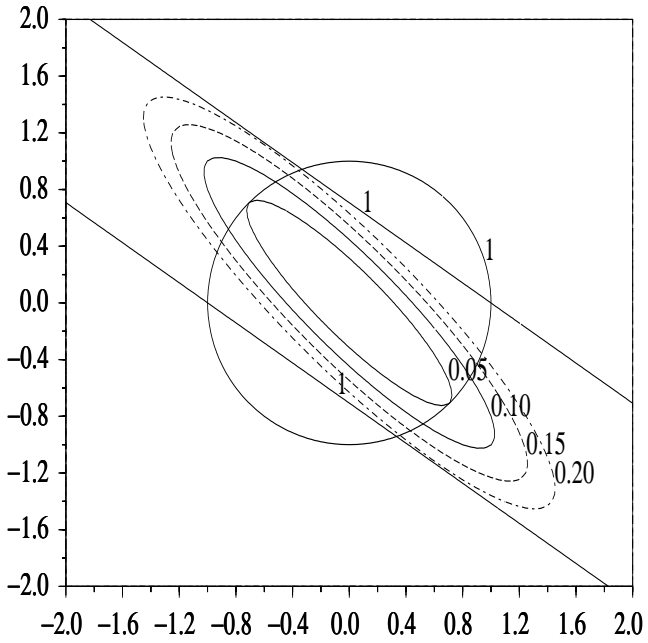

Figure 3: Loci of $(\mathbf{x}_1, \mathbf{x}_2)$ pairs for which $\|h\|^2 = 1$ (circle) and for which $|H(0)|^2 = 1$ (straight lines) in the design of $h(n)$. Elliptical contours show output power.

mizes the objective subject to the $\|h\|^2 = 1$ requirement that $\mathbf{x}$ lie on the Fig. 3 circle. This places the optimal solution at either intersection point of the circle and the semimajor axis of the ellipses. For sake of discussion, this solution is then rescaled to place DC gain $H(0)$ at unity, which moves $\mathbf{x}$ to the intersection (offscale, at $\mathbf{x}_2 = -\mathbf{x}_1 = 1 + \sqrt{2}$) of the semimajor ellipse axis with the upper straight line, which line represents the locus of $H(0) = 1$ possibilities. The optimal magnitude response after this rescaling is shown as the solid curve in Fig. 4. (The impulse response that results from this sort of stopband-only ordinary-eigenfilter design is related to the prolate-spheroidal wave function [1].)

The generalized-eigenfilter method requires choosing a quadratic to constrain. Suppose then that the objective is minimized subject to the requirement that $|H(0)|^2 = 1$ or $H(0) = \pm 1$. The dashed curve in Fig. 4 shows the resulting magnitude response. Comparison of the two responses makes clear that the constraint in these methods does more than set the scale of the solution. (The two curves would have been the same.)

Fig. 4 shows what is happening. The $H(0) = \pm 1$ condition corresponds to the two straight lines (the unit contour of the singular quadratic form representing $|H(0)|^2 = 1$ with elliptical contours stretched infinitely along their semimajor axis). The two optimal choices for $\mathbf{x}$ (at $\mathbf{x}_1 \approx \pm 0.602$, $\mathbf{x}_2 \approx \mp 1.13$) are near the points where the lines approach the outer ellipse.

Suppose the $|H(0)|^2 = 1$ constraint is generalized to $|H(f_0)|^2 = 1$. Since $H(f_0)$ has form $\mathbf{x}^T \mathbf{A}(f_0)$, an inner product, the $\mathbf{x}$ for which $H(f_0) = \pm 1$ will always lie on

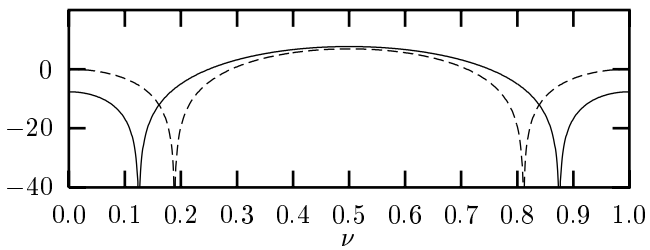

Figure 4: Magnitude responses (in dB) of the example highpass filter designed with the eigenfilter method (solid) and the generalized-eigenfilter method (dashed).

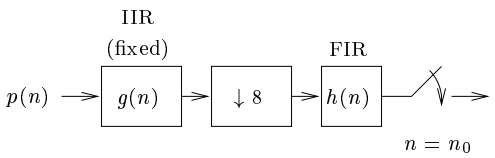

Figure 5: Matched-filter system for pulse $p(n)$.

a pair of parallel lines symmetrically placed about the origin. But the orientation of the lines is a function of constraint frequency f_0. The ordinary and generalized eigenfilter methods produce the same solution here (to within a scale factor) if these constraint lines are orthogonal to the semimajor axis of the ellipses. (This will be the case here if $f_0 = 0.5$.)

Note that the choice of f_0 is a distinct issue from the reference-frequency selection issue mentioned by some eigenvalue-method authors. The latter has to do with the choice of f_r in a weighting factor $\mathbf{x}^T \mathbf{A}(f_r) = H(f_r)$ used to linearize a test system for passband creation. Here no such passband was created.

A second design example addresses a matched-filtering system for pulse $p(n)$ in Fig. 5. An IIR filter is assumed already designed with ordinary methods to have a magnitude response that crudely approximates the general selectivity characteristics $|P(f)|$ of the ideal matched filter. A nonlinear-phase FIR filter operating at a lower sampling rate then fine tunes the response shape, including any needed phase compensation, so that the overall system acts like a matched filter. Our goal is to design this FIR filter to minimize the mismatch penalty, or equivalently, to maximize output SNR.

The basic design strategy is shown in Fig. 6. The FIR filter has been moved to the higher sample rate

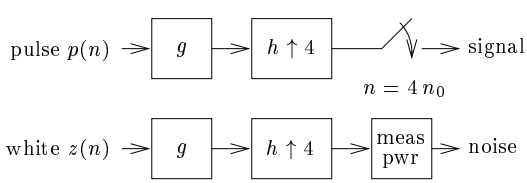

Figure 6: Design specification for $h(n)$ of Fig. 5: fix $|\,\text{signal}\,|^2$ and minimize noise power.

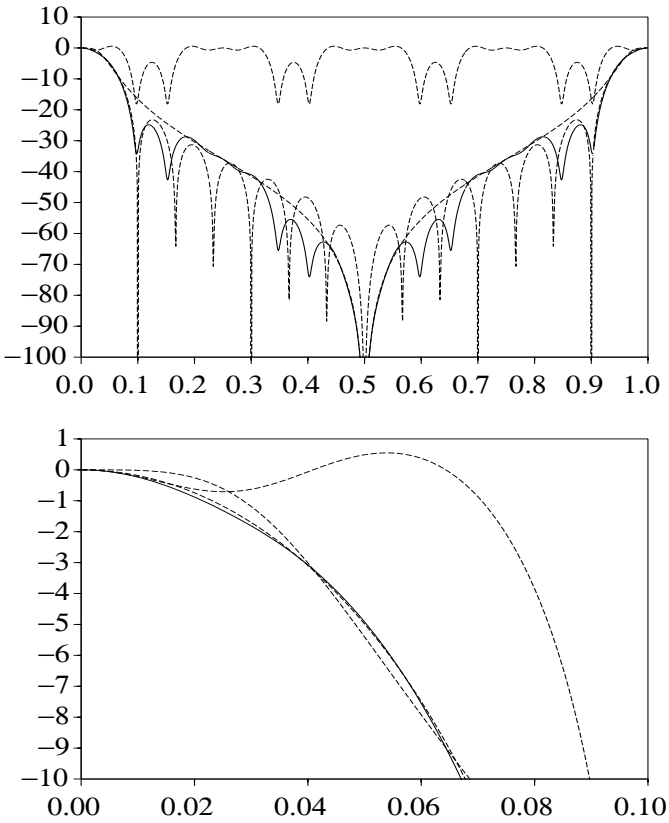

Figure 7: The key frequency responses of the numerical example on two different scales.

with a Noble identity and now appears with its impulse response zero interpolated. The upper diagram expresses that this equivalent system must respond to the input pulse $p(n)$ to some desired degree. To fix the level of this signal output, we can require that $|(p \star g \star (h \uparrow 4))(4n_0)|^2 = 1$, which, if we suppose that $h(n) = \mathbf{x}^T\mathbf{h}(n)$, becomes $\mathbf{x}^T\mathrm{Re}\{\mathbf{u}\mathbf{u}^H\}\mathbf{x} = 1$, where $\mathbf{u} = (p \star g \star (\mathbf{h} \uparrow 4))(4n_0)$, a suitable (rank-one) quadratic constraint for the generalized-eigenfilter method. Once the signal level is fixed, the degree of match is optimized by minimizing noise power out of the lower diagram in Fig. 6. Using (1) with $R_z(n) = \delta(n)$, this power is $\mathbf{x}^T\mathrm{Re}\{(g \star (\mathbf{h} \uparrow 4)) \star (\mathbf{h} \uparrow 4)' \star g')(0)\}\mathbf{x}$, suitable for minimization with the generalized-eigenfilter method.

Figure 7 shows optimal magnitude responses for a numerical example with a length-five FIR filter. The smooth curve in the upper plot represents the second-order Butterworth IIR filter $G(f)$. Its rolloff follows closely the sidelobes of the input pulse, which is 15 time-domain samples of a sinusoidal half cycle. The response of optimal FIR filter $h \uparrow 4$, scaled for unit DC gain, appears at the top. The cascade response $g \star (\mathbf{h} \uparrow 4)$, shown by the solid curve, nearly matches the input pulse in magnitude in the high-energy spectral regions, even in the enlarged-scale lower plot in Fig. 7. Finally Figure 8 shows the group delays of the two component filters and their cascade. The phase compensation resulting from the optimization is clearly evident.

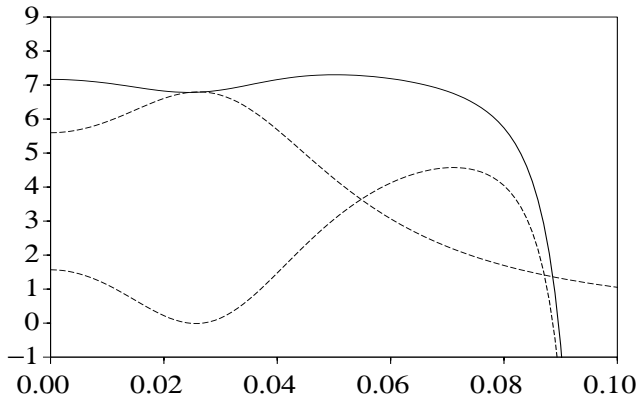

Figure 8: In-band group delays of the two component filters (dashed lines) and their cascade (solid line).

References

[1] P . P . Vaidyanathan and Truong Q. Nguyen, "Eigenfilters: A new approach to least-squares FIR filter design and applications including Nyquist filters," *IEEE Trans. Circuits and Systems*, vol. CAS-34, no. 1, pp. 11–23, Jan. 1987.

[2] Soo-Chang Pei and Jong-Jy Shyu, "Complex eigenfilter design of arbitrary complex coefficient FIR digital filters," *IEEE Trans. Circuits and Systems II*, vol. CAS-40, no. 1, pp. 32–39, Jan. 1993.

[3] Truong Q. Nguyen, "The design of arbitrary FIR digital filters using the eigenfilter method," *IEEE Trans. Sig. Proc.*, vol. 41, no. 3, pp. 1128–1139, Mar. 1993.

[4] L. Andrew, V. T. Franques, and V. K. Jain, "Eigen design of quadrature mirror filters," *IEEE Trans. Circuits and Systems II*, vol. CAS-44, no. 9, pp. 754–757, Sept. 1997.

[5] J. O. Coleman, "Embedded FIR generalized eigenfilters using test inputs," in *Proc. 1998 IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing (ICASSP '98)*, May 1998.

[6] Rajesh R. Venkatraman and Jeffrey O. Coleman, "A test-waveform view of the quadratic design of FIR filters," in *Proc. the 7th Int'l Conf. on Signal Processing Applications and Technology (ICSPAT '96)*, Oct. 1996.

[7] Jeffrey O. Coleman, "Systematic mapping of quadratic constraints on embedded FIR filters to linear matrix inequalities," in *Proc. Conf. on Information Sciences and Systems*, Princeton, Mar. 1998.

[8] D. P. Scholnik and J. O. Coleman, "Constrained quadratic design of FIR data-communication filters with linear matrix inequalities," in *Proc. 1998 Conf. on Information Sciences and Systems (CISS '98)*, Mar. 1998.