Conference on Information Sciences and Systems, Princeton, March 1998,
Special Session on LMI Methods in Signal Processing

SYSTEMATIC MAPPING OF QUADRATIC CONSTRAINTS ON EMBEDDED FIR FILTERS TO LINEAR MATRIX INEQUALITIES

Jeffrey O. Coleman

Naval Research Laboratory, Code 5341
Washington, DC 20375
jeffc@alum.mit.edu

ABSTRACT

This paper develops a flexible operator-based framework for the natural expression of application-specific quadratic constraints on the individual and joint design of FIR filters, along with an associated systematic mapping of those constraints to the linear-matrix-inequality (LMI) form compatible with a rapidly growing family of ultra-efficient numerical-optimization engines. This framework encourages, in particular, the constraint or minimization of the power gains of various embedding system configurations (possibly multirate) in response to specified input spectra—these gains are just application-specific squared L_2 norms—and offers a straightforward, systematic approach to the design of real or complex embedded filters with linear or nonlinear phase or N-th band or other structural properties. An example design illustrates.

1. INTRODUCTION

The test-input paradigm [1–4] offers an often-simple way to think about and formulate squared L_2 norms for the quadratic optimization of FIR filters. One might, for example, ask that an optimization force the two output powers in the Fig. 1 "test systems" to be small. 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 $S_s(\nu)$, and likewise a small P_p creates a passband by forcing the upper and lower sum inputs to be similar for input spectrum $S_p(\nu)$. (Here ν is normalized frequency.) Similarly, in Fig. 2 three test systems measure individual-filter stopband performance while a fourth measures passband performance in the synthesis operation.

Frequency-domain expression of these powers would show they are weighted (squared) L_2 approximation errors, and the next section shows that time-domain equivalents lead to analytically and computationally convenient formulation of quadratic constraints for numerical optimization. Computational operators are next defined to systematize extension to a wide variety of test systems containing FIR filters of many structural types. Software based on such an operator system allows nonexperts to target powerful numerical optimization engines at a wide variety of practical design problems involving various L_2 norms in time and frequency domains.

This application-adaptible test-input approach appears compatible with most optimization paradigms supporting quadratic optimization, including techniques based on eigenfilters [5–7], generalized eigenfilters [2, 3], second-order cone programs [8], and LMIs [9, 10]. The last are more general than required for quadratic optimization but an LMI method, semidefinite programming [11], is used here to enable later generalization to include H_∞ constraints [12, 13]. The idea of optimizing filters to extremize filter output power in response to an input process has a long history in the context of analyzing the statistics of the input process [14–18], but its use with filter design as the goal appears recent [1–4].

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. Graduate-student involvement in this work was supported by National Science Foundation grant MIP-9896034 to the University of Maryland Baltimore County.

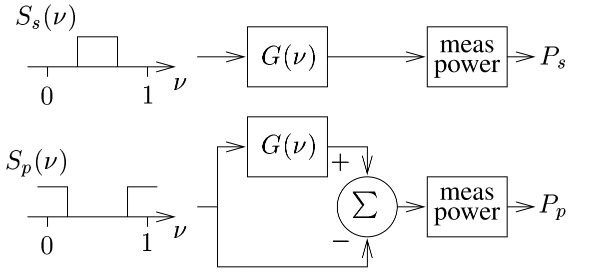

Figure 1: Quadratic optimization to make both output powers small makes $G(\nu)$ lowpass.

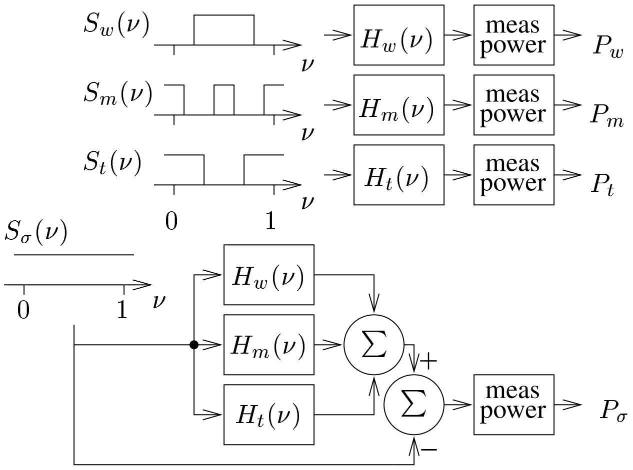

Figure 2: Performance measures in a simple spectral decomposition using woofer, midrange, and tweeter filters.

The paper ends with a design example and an appendix that presents a little LMI mathematics. A companion paper [4] better demonstrates the approach's flexibility: It presents an embedded FIR data-communication receive filter whose design is constrained by LMI norms representing mean-square intersymbol interference, mean-squared adjacent-channel interference, and noise power.

2. TEST-SYSTEM MATHEMATICS

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

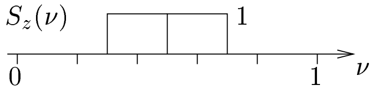

Figure 3: A highpass spectrum in five basis functions per period.

2.1. Vector Representation Sequences

Let vector $\mathbf{x} = (1, x_1, \ldots, x_M)^T$ contain a unit constant and some number M of real optimization variables. Any deterministic scalar complex discrete-time signal or impulse response $h(n)$ affine in optimization variables $\{x_1, \ldots, x_M\}$ can then be expressed as

$$h(n) = \mathbf{x}^T \mathbf{h}(n), \tag{1}$$

where the complex vector sequence $\mathbf{h}(n)$ serves as the *vector representation sequence* for the signal or system $h(n)$, which is then termed the *corresponding scalar sequence* equivalent to $\mathbf{h}(n)$. For example, allowing the nonzero portion of $\mathbf{h}(n)$ to comprise the sequence of three vectors

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ -j \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ j \end{pmatrix}$$

will structurally ensure that $h(n)$ is the complex linear-phase FIR sequence $(x_2 - jx_3), x_1, (x_2 + jx_3)$, optimizable indirectly through optimization of $\{x_1, x_2, x_3\}$. In this paper the support of a vector representation sequence is assumed finite.

Standard form (1) for the test-system impulse response next simplifies the derivation of output power, and the operators then developed aid construction of the needed vector representation sequences.

2.2. Output Power

Though the generalization to vector signals and matrix impulse responses is not difficult, the development here is limited to scalar-input, scalar-output test systems. An intermediate form of the development is initially convenient: Suppose $\mathbf{z}(n)$ is a wide-sense-stationary, zero-mean complex (column-)vector process, and suppose $\mathbf{v}(n)$ is a complex row-vector impulse response. Then $u(n) = (\mathbf{v} * \mathbf{z})(n)$ is a wide-sense stationary zero-mean complex (scalar) process, and the power P_u in u is[1] $E|u|^2 = R_u(0) = (\mathbf{v} * \mathbf{R_z} * \mathbf{v}')(0)$. Further suppose that $\mathbf{v}(n) = \mathbf{x}^T \mathbf{V}(n)$ in terms of complex matrix sequence $\mathbf{V}$. Then[2] $P_u = \mathbf{x}^T \mathrm{Re}\{(\mathbf{V} * \mathbf{R_z} * \mathbf{V}')(0)\}\mathbf{x}$. The technically unnecessary taking of the real part saves storage.

Here matrix $\mathbf{V}(n)$ and vectors $\mathbf{z}(n)$ and $\mathbf{v}(n)$ become column vector $\mathbf{g}(n)$ and scalars $z(n)$ and $g(n) = \mathbf{x}^T \mathbf{g}(n)$ respectively, so the convolution with the scalar autocorrelation in $P_u = \mathbf{x}^T \mathrm{Re}\{(\mathbf{g} * R_z * \mathbf{g}')(0)\}\mathbf{x}$ commutes. The resulting computationally convenient form[3] clearly uses only a finite number of autocorrelation samples, as the finite length of $\mathbf{g}(n)$ results in a finite sum:

$$E|u|^2 = \mathbf{x}^T \left(\sum_k \mathrm{Re}\{(\mathbf{g} * \mathbf{g}')(k) R_z^*(k)\} \right) \mathbf{x} \tag{2}$$

[1]Definition $\mathbf{R_z}(n) = E\left[\mathbf{z}(k)\,\mathbf{z}'(n-k)\right]$ and the identity $(\mathbf{v} * \mathbf{h})' = \mathbf{h}' * \mathbf{v}'$ give $R_u(n) = E\left[(\mathbf{v} * \mathbf{z})(k)(\mathbf{z}' * \mathbf{v}')(n-k)\right] = \sum_{u,v} \mathbf{v}(u) E\left[\mathbf{z}(k-u)\mathbf{z}'(n-k-v)\right] \mathbf{v}'(v) = \sum_{u,v} \mathbf{v}(u) \mathbf{R_z}(n-u-v)\mathbf{v}'(v) = (\mathbf{v} * \mathbf{R_z} * \mathbf{v}')(n)$.

[2]Autocorrelation property $\mathbf{R_z'} = \mathbf{R_z}$ implies $(\mathbf{V} * \mathbf{R_z} * \mathbf{V}')' = \mathbf{V}'' * \mathbf{R_z'} * \mathbf{V}' = \mathbf{V} * \mathbf{R_z} * \mathbf{V}'$, so matrix $(\mathbf{V} * \mathbf{R_z} * \mathbf{V}')(0)$ is Hermitian.

[3]If convolution $R_z * \mathbf{g}'$ is commutative, then $(\mathbf{g} * R_z * \mathbf{g}')(0) = ((\mathbf{g} * \mathbf{g}') * R_z)(0) = ((\mathbf{g} * \mathbf{g}') * R_z')(0) = \sum_k (\mathbf{g} * \mathbf{g}')(k) R_z'(0-k) = \sum_k (\mathbf{g} * \mathbf{g}')(k) R_z^H(k)$.

If the spectrum inverted to obtain $R_z(n)$ is specified in a particular form, inversion requires an inverse DFT instead of a numerical integral. Even approximating with that spectral form benefits: implicit approximation in the numerical integral is replaced with explicit spectral approximation whose ultimate effects are intuitive.

Suppose wide-sense-stationary complex random process $z(n)$ has spectral density $S_z(\nu)$ in terms of normalized frequency variable ν or $TS_z(fT)$ in terms of unnormalized frequency f, and suppose $S_z(fT)$, of period T^{-1}, is expressed in a basis comprising frequency-shifted copies of a function $\Phi(NfT)$:

$$S_z(fT) = \sum_k A_k\, \Phi(NfT - k), \tag{3}$$

where sequence A_k has period N. In Fig. 3, a $\{0,1\}$-valued function $\Phi(\nu)$ supported on $[-0.5, 0.5]$ is used with a sequence A_k of period $N = 5$ to represent a highpass spectrum. A symmetric triangle $\Phi(\nu)$ of width two would yield a piecewise-linear spectrum, a sinc $\Phi(\nu)$ would yield a bandlimited spectrum, etc., and larger N would permit approximation of actual spectra to arbitrary accuracy.

Using a convolution with a periodic impulse train,

$$S_z(fT) = \Phi(NfT) * \sum_k A_k\, \delta\left(f - \frac{k}{NT}\right).$$

Since $NT\,\Phi(NfT)$ and $(NT)^{-1} \sum_k A_k\, \delta(f - k/(NT))$ have inverse continuous-time transforms $\phi(\tau/(NT))$ and[4] $\sum_n a_n\, \delta(\tau - nT)$ respectively, where sequence a_n has period N and is the inverse DFT of sequence A_k, the inverse continuous-time transform of this particular $S_z(fT)$ is $R(\tau) = \phi\left(\frac{\tau}{NT}\right) \sum_n a_n\, \delta(\tau - nT)$. But discrete-time Fourier-transform relationship $R_z(n) \leftrightarrow S_z(\nu)$ ensures that $S_z(fT) = \sum_n R_z(n) e^{-j2\pi f nT}$, implying $R(\tau) = \sum_n R_z(n)\, \delta(\tau - nT)$ because each of the exponentials represents a time-domain delay. Equating these two expressions for $R(\tau)$,

$$R_z(n) = a_n \phi(n/N). \tag{4}$$

3. A SYSTEM OF OPERATORS

This section describes operators on vector representations of optimizable scalar sequences (and scalars not in sequences, in a simple extension) after first introducing functions for creating those representations. Here let $\mathbf{e}_i$ denote the i-th standard unit vector.

3.1. Creating Optimizable Sequences and Variables

Define operation $\mathbf{g} = \mathbf{convSequence}(w)$ on complex scalar sequence $w(n)$ to set $\mathbf{g}(n) = w(n)\mathbf{e}_1$. Corresponding scalar sequence $g(n)$ is then $\mathbf{x}^T \mathbf{g}(n) = \mathbf{x}^T w(n) \mathbf{e}_1 = w(n)$, so this converts from an ordinary "fixed" (nonoptimized) sequence w to its associated vector representation sequence, here $\mathbf{g}$. This is useful when

[4]Write $A(f) = \frac{1}{NT} \sum_i A_i\, \delta\left(f - \frac{i}{NT}\right)$ as

$$A(f) = \frac{1}{NT} \sum_{\substack{k \in [\mathbb{Z}/N\mathbb{Z}] \\ m \in N\mathbb{Z}}} A_k\, \delta\left(f - \frac{k+m}{NT}\right)$$

using a coset decomposition of the integers and using $A_i = A_{k+m} = A_k$. Since $A(f)$ has support in $(NT)^{-1}\mathbb{Z}$ and period T^{-1}, inverse Fourier transform $a(\tau)$ has period NT and support in $T\mathbb{Z}$. For some sequence a_ℓ of period N then, $a(\tau) = \sum_\ell a_\ell\, \delta(\tau - \ell T)$. The area a_n of the time-zero impulse in $a(\tau + nT)$ is the average of its transform $A(f)\, e^{j2\pi f T n}$ over period T^{-1}: $a_n = T \int_{\text{period}} A(f)\, e^{j2\pi f T n}\, df = \frac{1}{N} \sum_{k \in [\mathbb{Z}/N\mathbb{Z}]} A_k \int_{\text{period}} e^{j2\pi f T n} \sum_{m \in N\mathbb{Z}} \delta\left(f - \frac{k+m}{NT}\right) df$. Just one impulse is integrated, so $a_n = \frac{1}{N} \sum_{k \in [\mathbb{Z}/N\mathbb{Z}]} A_k\, e^{j2\pi \frac{kn}{N}}$, a periodic sequence and the inverse DFT of sequence A_k.

test systems include known impulse responses. Further define $\mathbf{h} = \textbf{newSequence}(L)$ to create a sequence $\mathbf{h}(n)$ on $n = 0, \ldots, L-1$ with each element set to a distinct standard unit vector, excluding $\mathbf{e}_1$ and unit vectors previously so "allocated." Then

$$\mathbf{h} = \textbf{newSequence}(2)$$
$$\mathbf{g} = \textbf{newSequence}(3)$$

might establish (other elements implicitly $\mathbf{0}$)

n	$\mathbf{h}(n)$	$\mathbf{g}(n)$
0	$\mathbf{e}_2$	$\mathbf{e}_4$
1	$\mathbf{e}_3$	$\mathbf{e}_5$
2		$\mathbf{e}_6$

so that $h(n) = \mathbf{x}^T\mathbf{h}(n)$ and $g(n) = \mathbf{x}^T\mathbf{g}(n)$ implicitly become causal sequences x_1, x_2 and x_3, x_4, x_5 respectively. Any number of real, separately optimizable scalar sequences can be so defined.

Occasionally an auxiliary optimization variable (not part of a sequence) is needed, and so $\alpha = \textbf{newVariable}(\,)$ sets α to a new unit vector so that $\mathbf{x}^T\alpha$, the corresponding optimizable scalar, is here the desired new optimization variable. Sum, difference, and scaling operations on such vector representations correspond to similar operations on the corresponding scalar equivalents, and here $|\alpha|^2$ (or α^2 if α is real) will further refer to rank-one matrix $\alpha\alpha^H$, because $\mathbf{x}^T(\alpha\alpha^H)\mathbf{x}$ is the corresponding square of the corresponding optimizable scalar. (The Fourier-transform operator discussed later also creates vector representations of optimizable scalars.)

The more-extensive array of operations available to manipulate sequences is the topic of the next section.

3.2. Operators to Manipulate Sequences

Various operations for manipulating vector representation sequences are defined here so that the effect on the corresponding scalar equivalents is reasonably intuitive. First, the ordinary linear arithmetic operations of sum, difference, and scaling by a constant can be used in the usual way with the obvious effect. For example, setting $\mathbf{w} = \mathbf{h} + \mathbf{g}$ results in $w(n) = h(n) + g(n)$.

Second, the first four operations in Table 1 are also linear and so operate on vector representation sequences in a way paralleling the corresponding familiar operations on scalar sequences. For example, defining $(\mathbf{h} \downarrow k)(n) = \mathbf{h}(kn)$ imparts a parallel relation between the corresponding scalar sequences.

Third, with the pointwise-product and convolution operations at least one of $\mathbf{h}$ or $\mathbf{g}$ must represent a fixed sequence. This fixed sequence and the operator together operate linearly on the other argument sequence, and so the result remains affine in the optimization variables. For example, if $\mathbf{h}$ represents a fixed sequence (vectors collinear with $\mathbf{e}_1$), then a formal definition of $\mathbf{h} \times \mathbf{g}$ might use the scalar $h(n)$ and compute the result as the scalar-vector product $h(n)\mathbf{g}(n)$. Convolution is similar: Element n of $\mathbf{h} * \mathbf{g}$ is defined as $(h * \mathbf{g})(n)$ using an ordinary scalar-vector convolution.

Finally, the last two operations in the table are a little different. If $h(n) = \mathbf{x}^T\mathbf{h}(n)$, then $h'(n) = h^*(-n) = \mathbf{x}^T\mathbf{h}^*(-n)$, so $\mathbf{h}^\dagger(n)$ must be defined as $\mathbf{h}^*(-n)$. Notice this is *not* equivalent to the obvious parallel $\mathbf{h}'(n)$. The Fourier Transform operator is more interesting. Mathematically, the Fourier transform of scalar sequence $h(n) = \mathbf{x}^T\mathbf{h}(n)$ is $H(\nu) = \sum_n \mathbf{x}^T\mathbf{h}(n)e^{-j2\pi\nu n} = \mathbf{x}^T\mathbf{H}(\nu)$, where $\mathbf{H}(\nu) = \sum_n \mathbf{h}(n)e^{-j2\pi\nu n}$ is just the element-wise Fourier transform of vector representation sequence $\mathbf{h}$. (Uppercase $\mathbf{H}(\nu)$ refers here to a vector, not a matrix.) So "operation" $\mathbf{h} \leftrightarrow \mathbf{H}$ on argument $\mathbf{h}$ "returns" a function of frequency $\mathbf{H}$ for which $\mathbf{H}(\nu)$ is a vector representation of an optimizable scalar.

The simple operators just defined permit construction of vector representation sequences for the impulse responses of many specialized filter types. For example,

notation	*represents*	*scalar equivalent*
$\mathbf{h} \gg k$	delay	$h(n-k)$
$\mathbf{h} \ll k$	advance	$h(n+k)$
$\mathbf{h} \downarrow k$	decimation	$h(kn)$
$\mathbf{h} \uparrow k$	interpolation	$h(n/k)$ when $n/k \in \mathbb{Z}$, else 0
$\mathbf{h} \times \mathbf{g}$	pointwise product	$h(n)g(n)$
$\mathbf{h} * \mathbf{g}$	convolution	$(h * g)(n)$
$\mathbf{h}^\dagger$	match	$h'(n)$
$\mathbf{h} \leftrightarrow \mathbf{H}$	Fourier Transform	$h(n) \leftrightarrow \mathbf{x}^T\mathbf{H}(\nu)$

Table 1: Vector representation sequences can be manipulated with sums, differences, scaling, and these defined operations.

notation	*means*
u has spectrum $S_u(\nu)$	u is appropriate triple $(\Phi, \mathbf{A}, \delta)$
$u * \mathbf{h}$	$(\Phi, \mathbf{A}, \mathbf{h} * \mathbf{g})$
$\textbf{power}(u)$	kernel matrix $\mathrm{Re}\{(\mathbf{g} * R_z * \mathbf{g}')(0)\}$ or element (1,1) of that matrix if other elements zero

Table 2: Operations on random process u represented by triple $(\Phi, \mathbf{A}, \mathbf{g})$. $\textbf{power}(\cdot)$ returns an ordinary scalar when possible.

$$\mathbf{x} = \textbf{newSequence}(L)$$
$$\mathbf{y} = \textbf{newSequence}(L)$$
$$\mathbf{r} = \mathbf{x} + j\mathbf{y}$$

defines $\mathbf{r}$ to represent the impulse response of an L-tap complex, nonlinear-phase FIR filter. The further steps

$$\mathbf{c} = \textbf{newSequence}(1)$$
$$\mathbf{s} = \mathbf{r} \gg 1$$
$$\mathbf{h} = \mathbf{s}^\dagger + c + \mathbf{s}$$

result in a representation $\mathbf{h}$ of a $2L+1$ tap complex linear-phase FIR filter. Finally, replacing the middle line with

$$\mathbf{s} = (\mathbf{r} \uparrow 2) \gg 1$$

would result instead in $\mathbf{h}$ representing a complex linear-phase half-band FIR filter of length $4L-1$.

3.3. Linear Equality Constraints

If α and β are vector representations of optimizable scalars $\mathbf{x}^T\alpha$ and $\mathbf{x}^T\beta$, operation $\alpha = \beta$ is defined to create the pair of linear equalities $\mathbf{x}^T\mathrm{Re}\{\alpha - \beta\} = 0$ and $\mathbf{x}^T\mathrm{Im}\{\alpha - \beta\} = 0$ to pass to the optimization engine. Such an equality constraint is used in [4], for example, to fix a system gain.

The extension to sequences is straightforward. If $\mathbf{h}(n)$ and $\mathbf{g}(n)$ are vector representation sequences for $h(n)$ and $g(n)$, then operation $\mathbf{h} = \mathbf{g}$ creates linear constraints $\mathbf{x}^T\mathrm{Re}\{\mathbf{h}(n) - \mathbf{g}(n)\} = 0$ and $\mathbf{x}^T\mathrm{Im}\{\mathbf{h}(n) - \mathbf{g}(n)\} = 0$ for n ranging over indices for which $\mathbf{h}(n) - \mathbf{g}(n)$ is nonzero.

3.4. Random Processes and LMI Power Constraints

In a test system a random process of form $u(n) = (z * g)(n)$ commonly results from passing some zero-mean, wide-sense-stationary input process $z(n)$ through some LTI system with an impulse response $g(n) = \mathbf{x}^T\mathbf{g}(n)$. (A unit impulse for $g(n)$ is possible.) Equation (2) shows power P_u to be a real quadratic form in $\mathbf{x}$ with a kernel fully characterized by $\mathbf{g}$ and autocorrelation $R_z(n)$. But $R_z(n)$ can be characterized spectrally as in (3) with a pair $(\Phi, \mathbf{A})$, where vector $\mathbf{A}$ is one period of sequence A_k, so triple $(\Phi, \mathbf{A}, \mathbf{g})$ here represents random process $u(n)$. Table 2 summarizes.

When making several test-system output powers small there is typically one power, call it $\mathbf{x}^T\mathbf{Q}\mathbf{x}$, that is minimized while the others are individually upperbounded. Semidefinite-programming optimization engines require an objective linear in vector $\mathbf{x}$, so minimization of power, a quadratic in $\mathbf{x}$, is accomplished through an auxiliary optimization variable, minimizing α under the constraint $\mathbf{x}^T\mathbf{Q}\mathbf{x} < \alpha$ or the often numerically preferable $\mathbf{x}^T\mathbf{Q}\mathbf{x} < \alpha^2 = \mathbf{x}^T(\mathbf{b}\mathbf{b}^H)\mathbf{x}$, for example, using description $\alpha = \mathbf{b}^T\mathbf{x}$. As a special case, $\mathbf{x}^T\mathbf{Q}\mathbf{x}$ is upperbounded by constant when $\mathbf{b}$ is collinear with $\mathbf{e}_1$.

So define a bounding operator: If P is a quadratic expression represented by matrix kernel $\mathbf{Q}$ with $\mathbf{Q} > 0$, such as returned say by[5] **power**$(\cdot)$, and if γ^2 is a quadratic expression represented by a nonnegative-definite, rank-one, real matrix kernel $\mathbf{B}$, such as returned say by the squaring of an optimizable scalar, then let constraint notation $P < \gamma^2$ mean $\mathbf{x}^T\mathbf{Q}\mathbf{x} < \mathbf{x}^T\mathbf{B}\mathbf{x}$ or equivalently $\mathbf{x}^T\mathbf{Q}\mathbf{x} < (\mathbf{b}^T\mathbf{x})^2$ for some real vector $\mathbf{b}$. This constraint can be represented by LMI

$$\begin{pmatrix} \mathbf{b}^T\mathbf{x} & (\mathbf{S}\mathbf{x})^T \\ \mathbf{S}\mathbf{x} & (\mathbf{b}^T\mathbf{x})I \end{pmatrix} > 0, \tag{5}$$

where $\mathbf{Q} = \mathbf{S}^T\mathbf{S}$ is assumed.[6] This LMI is equivalent to $\mathbf{x}^T\mathbf{Q}\mathbf{x} < (\mathbf{b}^T\mathbf{x})^2$ and $\mathbf{b}^T\mathbf{x} > 0$ together through $\mathbf{x}^T\mathbf{S}^T(\mathbf{b}^T\mathbf{x})^{-1}\mathbf{S}\mathbf{x} < \mathbf{b}^T\mathbf{x}$ and the Schur-complement equivalence (Appendix, Proposition 1). In this representational form, the constraint is ready to hand to an optimization engine.

Alternative factorizations of $\mathbf{Q}$ lead to a variety of technically equivalent LMI forms through the Schur-complement equivalence. The form in (5) was chosen for numerical reasons. All of the eigenvalues of the left side (Appendix, last section) but two are of value $\mathbf{b}^T\mathbf{x}$, and those two are $\mathbf{b}^T\mathbf{x} \pm \|\mathbf{S}\mathbf{x}\|$. In an interior-point optimization, the vector $\mathbf{x}$ will begin in the interior of the feasible region where all these eigenvalues are positive, and if this constraint is active, $\mathbf{x}$ will iteratively move to the edge of the feasible region where the smaller of these latter two eigenvalues approaches zero. The other eigenvalues are never spread widely enough to present numerical difficulties (as, for example, in [1]). Further, the value $\mathbf{b}^T\mathbf{x}$ often has an intuitive meaning in the context of a design problem—an rms error is typical—so setting up the problem to keep eigenvalues in a manageable range is often quite straightforward.

Some care is needed with the rank-one right side of the inequality. If the α^2 in $\mathbf{x}^T\mathbf{Q}\mathbf{x} < \alpha^2$ is naïvely transformed numerically into $\mathbf{x}^T\mathbf{B}\mathbf{x}$ and then factored into $(\mathbf{b}^T\mathbf{x})^2$, either $\alpha = \mathbf{b}^T\mathbf{x}$ or $\alpha = -\mathbf{b}^T\mathbf{x}$ could result. The former case results in the LMI

$$\begin{pmatrix} \alpha & (\mathbf{S}\mathbf{x})^T \\ \mathbf{S}\mathbf{x} & \alpha I \end{pmatrix} > 0,$$

but the latter results in

$$\begin{pmatrix} -\alpha & (\mathbf{S}\mathbf{x})^T \\ \mathbf{S}\mathbf{x} & -\alpha I \end{pmatrix} > 0.$$

Both imply $\mathbf{x}^T\mathbf{Q}\mathbf{x} < \alpha^2$ as desired, but the first also implies $\alpha > 0$ while the second implies $\alpha < 0$. In the first case, α should be minimized, but it is $-\alpha$ that should really be minimized in the second case. Attempting instead to minimize α will send $\alpha \to -\infty$ without forcing the quadratic constraint to become active at all. It is not yet clear what the cleanest way is to circumvent this minor technicality. In practice it has yet to present a problem.

4. A DESIGN EXAMPLE

The design of a family of three bandsplitting (analysis) filters is specified procedurally in Table 3 using the operators of Section 3 to describe the test systems of Fig. 2 and the constraints on their outputs. The first three specification segments give the filter impulse responses $h_w(n)$, $h_m(n)$, and $h_t(n)$, each a linear-phase structure with 21 taps. After the random processes $z_w(n)$, $z_m(n)$, and $z_t(n)$ that drive the test systems measuring stopband performance are defined, those performance levels are characterized as mean-squared frequency-response errors, each an input-specific power gain. The passband error, $\mathbf{MSE}_\sigma$, is defined similarly, with $\mathbf{w}$ formally representing the $\delta(n)$ impulse response of the "wire" driving the negative difference input at the bottom of the large Fig. 2 test system. The final segment formally specifies an optimization with four LMIs. The mean-square stopband errors are set to -50 dB.

$\mathbf{c} = \textbf{newSequence}(1)$
$\mathbf{r} = \textbf{newSequence}(10)$
$\mathbf{h_w} = (\mathbf{r} \gg 1)^\dagger + \mathbf{c} + (\mathbf{r} \gg 1)$

$\mathbf{c} = \textbf{newSequence}(1)$
$\mathbf{r} = \textbf{newSequence}(10)$
$\mathbf{h_m} = (\mathbf{r} \gg 1)^\dagger + \mathbf{c} + (\mathbf{r} \gg 1)$

$\mathbf{c} = \textbf{newSequence}(1)$
$\mathbf{r} = \textbf{newSequence}(10)$
$\mathbf{h_t} = (\mathbf{r} \gg 1)^\dagger + \mathbf{c} + (\mathbf{r} \gg 1)$

z_w has spectrum $S_w(\nu)$
z_m has spectrum $S_m(\nu)$
z_t has spectrum $S_t(\nu)$

$\mathbf{MSE}_w = \textbf{power}(z_w * \mathbf{h_w})/\textbf{power}(z_w)$
$\mathbf{MSE}_m = \textbf{power}(z_m * \mathbf{h_m})/\textbf{power}(z_m)$
$\mathbf{MSE}_t = \textbf{power}(z_t * \mathbf{h_t})/\textbf{power}(z_t)$

z_σ has spectrum $S_\sigma(\nu)$
$\mathbf{w} = \textbf{convSequence}(\delta)$
$\mathbf{h}_\sigma = \mathbf{h_w} + \mathbf{h_m} + \mathbf{h_t}$
$\mathbf{MSE}_\sigma = \textbf{power}(z_\sigma * (\mathbf{h}_\sigma - \mathbf{w}))/\textbf{power}(z_\sigma)$

$\rho = \textbf{newVariable}(\,)$
optimize $x_1, x_2, \ldots$ to minimize ρ
subject to
$$\begin{array}{lcl} \mathbf{MSE}_w & < & 10^{-5} \\ \mathbf{MSE}_m & < & 10^{-5} \\ \mathbf{MSE}_t & < & 10^{-5} \\ \mathbf{MSE}_\sigma & < & \rho^2 \end{array}$$

Table 3: Optimization specification for the three filters of Fig. 2.

After the optimization fixes $\mathbf{x}$, the impulse responses $h_w(n) = \mathbf{x}^T\mathbf{h_w}(n)$, $h_m(n) = \mathbf{x}^T\mathbf{h_m}(n)$, and $h_t(n) = \mathbf{x}^T\mathbf{h_t}(n)$ are known and the frequency reponses can be computed. The optimized magnitude responses are plotted in Fig. 4.

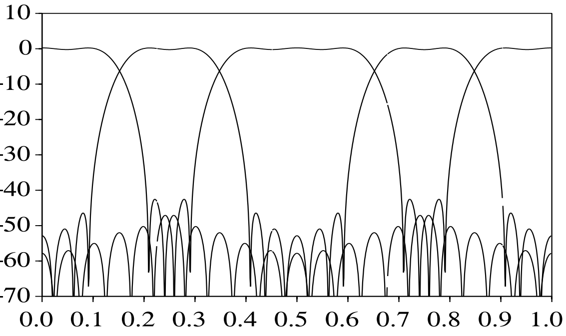

Figure 4: Magnitude responses of bandsplitting filters.

[5]FIR coefficients not all zero result in a nonzero frequency response almost everywhere, so an input power spectrum with support of nonzero measure must result in positive output power. This often ensures $\mathbf{Q} > 0$.

[6]Use singular value decomposition to obtain $\mathbf{Q} = \mathbf{U}^T\mathbf{\Lambda}\mathbf{U}$, set the negligibly small elements of diagonal matrix $\mathbf{\Lambda} > 0$ to zero, and set $\mathbf{S} = \mathbf{\Lambda}^{\frac{1}{2}}\mathbf{U}$ after discarding rows of orthogonal matrix $\mathbf{U}$ and corresponding diagonal elements of $\mathbf{\Lambda}$ wherever the latter are zero. This generally makes $\mathbf{S}$ significantly smaller than square and saves significant storage, important in nontoy problems. It may increase numerical robustness as well.

5. SUMMARY

Using numerical optimization with linear matrix inequalities, many FIR filter types are easily designed so that embedding systems will meet flexibly chosen quadratic performance criteria. This reduces filter design to the specification of embedding "test systems" whose structures reflect the filter's intended application and whose output powers characterize performance quantities of true interest.

Sequence representation is key. Relation $h(n) = \mathbf{x}^T\mathbf{h}(n)$ uses vector representation sequence $\mathbf{h}(n)$ to express the affine dependence of corresponding scalar sequence $h(n)$ on the real optimization variables in vector $\mathbf{x} = (1, x_1, \ldots, x_M)^T$. The representation of a test system's impulse response is then computed by mirroring the system's signal-processing operations on internal signals with appropriately defined and notationally parallel operations on the representations of those signals. The system output power and the LMI that constrains it become straightforward functions of $\mathbf{x}$. In this way a family of specifications as test systems becomes a family of LMIs in a semidefinite program specifying numerical optimization using some readily available engine. The optimal filter can be straightforwardly computed from the optimal $\mathbf{x}$ and the vector representation sequence of its impulse response.

This approach invites extension to broaden the compatible family of performance measures. But the ultimate goal is always the design of filters to perform in the application-specific context in which they will actually operate.

6. REFERENCES

[1] R. R. Venkatraman and J. 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.

[2] 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, accepted.

[3] J. O. Coleman, "Quadratic FIR-filter design as a generalized eigenproblem," in *Proc. 1998 Eureopean Signal Processing Conference (EUSIPCO '98)*, Sept. 1998, accepted.

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

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

[6] S.-C. Pei and J.-J. Shyu, "Complex eigenfilter design of arbitrary complex coefficient FIR digital filters," *IEEE Transactions on Circuits and Systems II*, vol. CAS-40, no. 1, pp. 32–39, Jan. 1993.

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

[8] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Second-order cone programming," *Linear Algebra and its Applications*, accepted for the special issue on linear algebra in control, signals, and image processing.

[9] T. Settle, M. Zoltowski, and V. Balakrishnan, "Optimal multichannel equalizers based on linear matrix inequalities," Invited presentation at the First Annual UCSD Conference on Wireless Communications, Center for Wireless Communications, UC San Diego, March 1998.

[10] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15 of *Studies in Applied Mathematics*, SIAM, Philadelphia, PA, June 1994.

[11] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, Mar. 1996.

[12] M. Fu, S. Dasgupta, and G. A. Williamson, "A general framework for optimal multirate filter bank design," in *Proc. Fourth Int'l Symp. on Signal Processing and its Applications (ISSPA '96)*, Gold Coast, Australia, Aug. 1996.

[13] S.-P. Wu, S. Boyd, and L. Vandenberghe, "FIR filter design via semidefinite programming and spectral factorization," in *Proc. 1996 IEEE Conf. on Decision and Control (CDC '96)*, Dec. 1996, Kobe, Japan.

[14] E. Robinson, *Statistical Communication and Detection*, Griffin, London, 1967.

[15] G. Carayannis and C. Gueguen, "The factorial linear modelling: A karhunen-loeve approach to speech analysis," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP '76)*, Apr. 1976, pp. 489–492.

[16] John Makhoul, "On the eigenvectors of symmetric toeplitz matrices," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 4, pp. 868–872, Aug. 1981.

[17] B. Farhang-Boroujeny, "On eigenvalues distribution of correlation matrices," in *Proc. IEEE Int'l Symposium on Circuits and Systems*, June 1991, pp. 1251–1254.

[18] B. Farhang-Boroujeny and S. Gazor, "Selection of orthonormal transforms for improving the performance of the transform domain normalised lms algorithm," *IEE Proceedings, Part F*, vol. 139, no. 5, pp. 327–335, Oct. 1992.

APPENDIX: SOME LMI BASICS

Positive and Nonnegative Definiteness

In this discussion of definitions, the word "positive" can be replaced everywhere with "nonnegative" to arrive at parallel definitions.

Definition 1 *The phrase "$\mathbf{A}$ is positive definite" here means that* $\mathrm{Re}\{\mathbf{z}^H\mathbf{A}\mathbf{z}\}$ *is positive for all nonzero complex vectors* $\mathbf{z}$.

But for complex matrix $\mathbf{A}$,

$$\begin{aligned} \mathrm{Re}\{\mathbf{z}^H\mathbf{A}\mathbf{z}\} &= \mathbf{z}^H\left(\frac{\mathbf{A}+\mathbf{A}^H}{2}\right)\mathbf{z} \\ i\,\mathrm{Im}\{\mathbf{z}^H\mathbf{A}\mathbf{z}\} &= \mathbf{z}^H\left(\frac{\mathbf{A}-\mathbf{A}^H}{2}\right)\mathbf{z}, \end{aligned}$$

so a complex matrix $\mathbf{A}$ is positive definite if and only if its Hermitian component $\frac{1}{2}(\mathbf{A}+\mathbf{A}^H)$ is positive definite (and indeed if any matrix with the same Hermitian component is positive definite). If $\mathbf{A}$ is Hermitian, the quadratic form is real, so this definition reduces to the conventional notion that a Hermitian matrix $\mathbf{A}$ is positive definite if $\mathbf{z}^H\mathbf{A}\mathbf{z}$ is positive for all nonzero complex vectors $\mathbf{z}$. If $\mathbf{A}$ is singular, then some nonzero complex $\mathbf{z}$ leads to $\mathbf{A}\mathbf{z} = 0 = \mathrm{Re}\{\mathbf{z}^H\mathbf{A}\mathbf{z}\}$, so positive definiteness implies invertibility (but nonnegative definiteness does not). Further, the inverse $\mathbf{A}^{-1}$ of a positive-definite Hermitian matrix $\mathbf{A}$ is itself positive definite, because $\mathbf{z}^H\mathbf{A}^{-1}\mathbf{z} = (\mathbf{A}^{-1}\mathbf{z})^H\mathbf{A}(\mathbf{A}^{-1}\mathbf{z})$ and because the nonsingularity of $\mathbf{A}^{-1}$ means $\mathbf{A}^{-1}\mathbf{z} \neq 0$ for $\mathbf{z} \neq 0$.

Now suppose matrix $\mathbf{A}$ is real. If $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ with $\mathbf{x}$ and $\mathbf{y}$ real, then $\mathrm{Re}\{\mathbf{z}^H\mathbf{A}\mathbf{z}\} = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{y}^T\mathbf{A}\mathbf{y}$, so $\mathbf{A}$ is positive definite if and only if $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is positive for all nonzero real vectors $\mathbf{x}$. Further, since $\mathbf{A}$ and $\mathbf{x}$ are real here, quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$ depends only on the symmetric component of $\mathbf{A}$ because $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\frac{1}{2}(\mathbf{A} +$

$\mathbf{A}^T)\mathbf{x}$, which is why the positive definiteness of real matrices is often defined only for symmetric matrices.

Compact notation: When matrix $\mathbf{A}$ is real, $\mathbf{A} \geq 0$ and $\mathbf{A} > 0$ mean that $\mathbf{A}$ is symmetric and either nonnegative definite or positive definite, respectively. When $\mathbf{A}$ is an affine (linear plus a constant) function of some real vector $\mathbf{x}$, then either $\mathbf{A}(\mathbf{x}) \geq 0$ or $\mathbf{A}(\mathbf{x}) > 0$ is called[7] a *linear matrix inequality* (LMI) in $\mathbf{x}$.

The Schur-Complement Equivalence

The LMIs used in this paper express design requirements through the Schur-Complement equivalence. The subscript H below indicates the Hermitian component, so that $\mathbf{R}_H = \frac{1}{2}(\mathbf{R} + \mathbf{R}^H)$.

Proposition 1 *If $\mathbf{R}$ is positive definite, then matrix $\mathbf{Q} - \mathbf{S}^H\mathbf{R}\mathbf{S}$ is positive (nonnegative) definite if and only if*

$$\begin{pmatrix} \mathbf{Q} & \mathbf{S}^H \\ \mathbf{S} & \mathbf{R}_H^{-1} \end{pmatrix} \tag{6}$$

is positive (nonnegative) definite.

PROOF The proof of the positive-definite result is presented here. Changing the word "positive" to "nonnegative," except that $\mathbf{R}$ and its inverse should everywhere be left positive definite, results in a proof of the nonnegative-definite result.

The positive definiteness of each matrix $\mathbf{R}$, $\mathbf{Q} - \mathbf{S}^H\mathbf{R}\mathbf{S}$, and

$$\begin{pmatrix} \mathbf{Q} & \mathbf{S}^H \\ \mathbf{S} & \mathbf{R}_H^{-1} \end{pmatrix}$$

is equivalent respectively to that of its Hermitian component $\mathbf{R}_H$, $\mathbf{Q}_H - \mathbf{S}^H\mathbf{R}_H\mathbf{S}$ and

$$\begin{pmatrix} \mathbf{Q}_H & \mathbf{S}^H \\ \mathbf{S} & \mathbf{R}_H^{-1} \end{pmatrix},$$

so it is enough to prove the special case of the proposition in which $\mathbf{Q}$ and $\mathbf{R}$ are both Hermitian. Assume then that this is so.

Suppose the matrix (6) is positive definite. Then

$$\mathrm{Re}\left\{ \begin{pmatrix} \mathbf{u}^H & \mathbf{v}^H \end{pmatrix} \begin{pmatrix} \mathbf{Q} & \mathbf{S}^H \\ \mathbf{S} & \mathbf{R}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \right\} \tag{7}$$

is positive for every choice of complex vectors $\mathbf{u}$ and $\mathbf{v}$ not both zero. Expanding and using the invertibility of the matrix given here as $\mathbf{R}^{-1}$, this is seen to be equivalent to the statement that

$$\mathrm{Re}\{(\mathbf{v} + \mathbf{R}\mathbf{S}\mathbf{u})^H\mathbf{R}^{-1}(\mathbf{v} + \mathbf{R}\mathbf{S}\mathbf{u})\} + \mathrm{Re}\{\mathbf{u}^H\left(\mathbf{Q} - \mathbf{S}^H\mathbf{R}\mathbf{S}\right)\mathbf{u}\} \tag{8}$$

is positive for every choice of complex vectors $\mathbf{u}$ and $\mathbf{v}$ not both zero, even when the first term is zeroed by an arbitrary nonzero $\mathbf{u}$ and $\mathbf{v} = -\mathbf{R}\mathbf{S}\mathbf{u}$. So the positive definiteness of (6) and the invertibility of $\mathbf{R}$ together imply that $\mathbf{Q} - \mathbf{S}^H\mathbf{R}\mathbf{S}$ is positive definite.

Conversely, (8) and hence (7) will be positive if matrices $\mathbf{Q} - \mathbf{S}^H\mathbf{R}\mathbf{S}$ and $\mathbf{R}^{-1}$ are positive definite and if $\mathbf{u}$ and $\mathbf{v}$ are not both zero. Positive-definite Hermitian matrix $\mathbf{R}$ necessarily has a positive-definite inverse, so (6) is positive definite if $\mathbf{Q} - \mathbf{S}^H\mathbf{R}\mathbf{S}$ is. □

In applications matrix $\mathbf{Q}$ is often naturally Hermitian, but there is no requirement here that it be so. Also, the details of the proof reveal that the "if" part does not actually require the positive definiteness of $\mathbf{R}$ but only its invertibility.

The Schur-complement equivalence is typically used for a kind of linearization. For example, the requirement that the scalar $q - \mathbf{z}^H\mathbf{R}\mathbf{z}$ be positive is equivalent by this proposition to the positive definiteness of matrix

$$\begin{pmatrix} \mathbf{q} & \mathbf{z}^H \\ \mathbf{z} & \mathbf{R}_H^{-1} \end{pmatrix}.$$

A requirement on an expression quadratic in $\mathbf{z}$ is thereby transformed to one on an expression affine in $\mathbf{z}$. If the quantities involved are all real, an LMI results. In applications, however, it is common for $\mathbf{z}$ and $\mathbf{R}$ to be complex. The next section shows that this obstacle can be straightforwardly overcome.

Constraining a Complex Matrix with an LMI

Complex matrix $\mathbf{M}$ is positive (nonnegative) definite if $\mathrm{Re}\{\mathbf{w}^H\mathbf{M}\mathbf{w}\}$ is positive (nonnegative) for every nonzero complex vector $\mathbf{w}$. To find a real matrix whose positive (nonnegative) definiteness is equivalent, define conversion matrix $\mathbf{C}$ like an identity matrix but with each column e_n replaced with the pair of columns $(e_n \; ie_n)$:

$$\mathbf{C} = \begin{pmatrix} 1 & i & 0 & 0 & & \\ 0 & 0 & 1 & i & & \\ & & & & \ddots & \ddots \end{pmatrix}.$$

Using such a matrix $\mathbf{C}$ of appropriate size, complex vector $\mathbf{w}$ can be written $\mathbf{w} = \mathbf{C}\mathbf{v}$ in terms of a vector $\mathbf{v}$ of the real and imaginary components of its elements, from which it follows that $\mathrm{Re}\{\mathbf{w}^H\mathbf{M}\mathbf{w}\}$ is positive (nonnegative) for every nonzero complex vector $\mathbf{w}$ if and only if $\mathbf{v}^T\mathrm{Re}\{\mathbf{C}^H\mathbf{M}\mathbf{C}\}\mathbf{v}$ is positive (nonnegative) for every nonzero real vector $\mathbf{v}$. The positive (nonnegative) definiteness of complex matrix $\mathbf{M}$ is therefore equivalent to that of real matrix $\mathrm{Re}\{\mathbf{C}^H\mathbf{M}\mathbf{C}\}$ or to that of its symmetric component $\mathrm{Re}\{\mathbf{C}^H\mathbf{M}_H\mathbf{C}\}$.

Now further suppose that $\mathbf{M}$ is an affine function of a complex vector $\mathbf{z}$. By writing $\mathbf{z} = \mathbf{C}_2\mathbf{x}$, constraint $\mathrm{Re}\{\mathbf{C}_1^H\mathbf{M}_H(\mathbf{z})\mathbf{C}_1\} \geq 0$ can be expressed as $\mathrm{Re}\{\mathbf{C}_1^H\mathbf{M}_H(\mathbf{C}_2\mathbf{x})\mathbf{C}_1\} \geq 0$, an LMI in real vector $\mathbf{x}$. (Subscripts distinguish the two real-to-complex conversion matrices because in general they are of different sizes.) Thus an LMI can always be found to require the positive definiteness of a complex matrix affine in a complex vector. In applications, however, this general formulation is generally bypassed in favor of application-specific equivalents that result in smaller vector and matrix sizes.

Numerical Conditioning

The numerical conditioning of an LMI in the form of (5) is not difficult to determine. The eigenvector equation

$$\begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & \eta I \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_* \end{pmatrix} = \lambda \begin{pmatrix} \phi_1 \\ \phi_* \end{pmatrix}$$

is just

$$\begin{aligned} \gamma\phi_1 + \mathbf{c}^T\phi_* &= \lambda\phi_1 \\ \mathbf{c}\phi_1 + \eta\phi_* &= \lambda\phi_*, \end{aligned}$$

solved by $\lambda = \eta$, $\phi_1 = 0$, and any ϕ_* for which $\phi_* \perp \mathbf{c}$. If the matrix is of order N, then vector $\mathbf{c} \neq 0$ is of length $N-1$ and there are necessarily exactly $N-2$ orthogonal choices of ϕ_* for which $\phi_* \perp \mathbf{c}$, so $N-2$ eigenvalues are equal to η.

The other two solutions require $\phi_1 \neq 0$. Choosing $\phi_1 = \lambda - \eta$ and $\phi_* = \mathbf{c}$ satisfies the second equation and transforms the first to $\gamma(\lambda - \eta) + \|\mathbf{c}\|^2 = \lambda(\lambda - \eta)$, a quadratic in λ solved by

$$\lambda = \frac{\gamma + \eta}{2} \pm \sqrt{\left(\frac{\gamma - \eta}{2}\right)^2 + \|\mathbf{c}\|^2}.$$

If $\eta = \gamma$, then $\lambda = \gamma \pm \|\mathbf{c}\|$.

[7] Most authors admit only the "$\geq$" version as an LMI, but the usual interior-point numerical algorithms do not actually bring the two sides of an "active" inequality quite together, though close enough for practical purposes. This makes fussing over the definition seem pointless.