International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2002)
Orlando, FL, May 13–17, 2002

MIXED-NORM FIR FILTER OPTIMIZATION USING SECOND-ORDER CONE PROGRAMMING

Dan P. Scholnik

scholnik@nrl.navy.mil
Naval Research Laboratory, Washington, DC

ABSTRACT

It has been shown that second-order cone programming (SOCP) can be used to efficiently optimize FIR filter frequency responses in the sense of L_1, L_2, and L_∞. Although SOCP cannot be used for the more general case of L_p minimization, several techniques for combining norms provide the flexibility to minimize a continuum of norms from L_1 to L_∞. Burdakov's vector ϵ-norm is here extended to complex functions on an interval of the real line, and is compared with a simple convex combination of norms. The dual norms of both are found as well. Discretized approximations of these norms can be used for filter design with SOCP, as shown in the examples.

1. INTRODUCTION

Thanks to ever expanding computational resources, numerical optimization is the approach of choice for digital filter design. For large FIR filters especially, whenever possible we desire a convex problem formulation, as this guarantees a single global solution if a solution exists. Even better is to convert the problem into one of several standard convex program classes, so that modern high-efficiency interior-point solvers can be used. One such class is *second-order cone programming* (SOCP). A superset of linear-programming, SOCP allows both linear and convex quadratic constraints. Even though SOCP is very structured (and thus efficient) it is actually quite flexible, and a wide array of functional constraints can be formed by combining basic SOC constraints. Examples include constraining the L_∞, L_2, and L_1 norms of error functions that are affine in the underlying optimization variables.

Methods abound in the literature for combining L_∞ (minimax, Chebychev) and L_2 (least-squares) optimization methods for filter design. The classic constrained optimization approach is to constrain one error measure while minimizing the other. Peak Constrained Least Squares (PCLS) [1] is one such approach. An equivalent unconstrained method is to minimize a combined norm, formed as the convex combination of the L_∞ and L_2 norms [2]. A different combined norm, the ϵ-norm, was proposed in [3] and is extended here. All of these techniques seek to continuously bridge the gap between L_∞ and L_2 without using the general L_p family of norms and losing the efficient structures that exist for solving L_∞ and L_2 problems. This paper will consider not just combinations of L_2 and L_∞, but show that their dual norms are also combinations involving L_1 and L_2 norms. While the traditional approach for mixed L_2/L_∞ optimization has been to implement custom algorithms, here we show how all the combined norms can be constrained using SOCP, taking advantage of both the flexibility of the method and the efficiency of the existing solvers [4, 5].

This work was supported by the Office of Naval Research.

The basic problem formulation used in the sequel is to minimize some norm $\|E\|$ of a weighted frequency response error $E(f;\mathbf{x}) = W(f)[G(f;\mathbf{x}) - D(f)]$, where W is a real, nonnegative *weighting function*, $G(f;\mathbf{x})$ is the frequency response of the filter being designed, and $D(f)$ is the desired frequency response. Both E and G are possibly complex-valued functions of the real optimization variables $\mathbf{x}$. The focus here is on FIR filters, where $\mathbf{x}$ represents the real and imaginary parts of the filter coefficients, G is an affine function of $\mathbf{x}$, and G is periodic in f with period one:

$$G(f;\mathbf{x}) = \sum_n (x_{2n} + jx_{2n+1})e^{-j2\pi fn}.$$

By definition FIR (and stable IIR) frequency responses have bounded magnitudes and are in the class of functions $L_\infty[0,1]$, which implies that they belong to $L_p[0,1]$ for any $p \geq 1$. This largely removes the need to consider pathological functions in the following analysis.

2. SECOND-ORDER CONE PROGRAMMING

A powerful special case of convex optimization is *second-order cone programming*, or SOCP [6]. SOCP allows the minimization of a linear objective, subject to second-order cone (SOC) constraints, which encompass linear and convex quadratic constraints. One standard formulation of a SOCP is

$$\begin{aligned} \text{min.} \quad & \mathbf{f}^T\mathbf{x} \\ \text{s.t.} \quad & \|\mathbf{A}_k\mathbf{x} + \mathbf{b}_k\| \leq \mathbf{c}_k^T\mathbf{x} + d_k, \qquad k = 1,\ldots,K. \end{aligned}$$

An alternate notation for a SOC constraint is

$$\begin{aligned} \|\mathbf{A}\mathbf{x} + \mathbf{b}\|^2 &\leq (\mathbf{c}^T\mathbf{x} + d)^2 \\ 0 &\leq \mathbf{c}^T\mathbf{x} + d \end{aligned}$$

where the nonnegativity condition, while necessary for convexity, is often not explicitly given. Deceptively simple yet powerful, SOCP can be used to constrain l_1, l_2, and l_∞ norms of linear error vectors, as well as to approximately (via discretization) constrain L_1, L_2, and L_∞ norms of linear error functions [6, 7]. Exact L_2 formulations also exist [8]. Any constraint can also be used as the objective via an auxiliary variable. The following sections briefly review how SOCP can be used to constrain the various norms of a function $E : [0,1] \times \mathbb{R}^N \to \mathbb{C}$ that is affine in the optimization variables, the dependence on which is hereafter implicit.

2.1 L_∞-Norm Constraints

For a general function the L_∞ or Chebychev norm is defined as

$$\|E\|_\infty \triangleq \operatorname*{ess\,sup}_f |E(f)|.$$

Since the functions of interest (FIR transfer functions) are continuous, this reduces to

$$\|E\|_\infty \triangleq \max_f |E(f)|.$$

For a function E that is affine in the optimization variables, the L_∞-norm can be (approximately) constrained using SOCP by discretizing the interval $[0,1]$. Letting $\{f_0, f_1, \ldots, f_K\}$ be a set of points in $[0,1]$ (choosing $f_k = k/K$ is common), the set of SOC constraints

$$|E(f_k)| \leq \delta, \qquad k = 0, \ldots, K$$

bounds the maximum value of $|E(f_k)|$ on the grid by δ, which can be a constant or an auxiliary optimization variable. The spacing Δf between the frequency points is chosen so that deviation between the points is negligible. For an FIR filter of length N a good rule of thumb is $\Delta f \approx 1/(20N)$.

2.2 L_2-Norm Constraints

In contrast to the L_∞ norm, which requires one constraint per discrete frequency, the L_2 norm requires only a single SOC constraint. The L_2 norm is defined as

$$\|E\|_2 \triangleq \left(\int_0^1 |E(f)|^2 \, df \right)^{\frac{1}{2}}.$$

Since E was assumed affine in the optimization variables $\mathbf{x}$, this expression can be written as a positive semidefinite quadratic form in $\mathbf{x}$, and thus as a SOC. The integral can be calculated in any number of ways, perhaps the simplest being a Riemann sum over a uniform partition $\{f_k\}$ of $[0,1]$. The resulting SOC constraint can be written

$$\sum_k |E(f_k)|^2 \Delta f \leq \delta^2,$$

where δ is a nonnegative constant or auxiliary variable. A somewhat smaller value, $\Delta f \approx 1/(40N)$, is required for negligible deviation from the true integral.

2.3 L_1-Norm Constraints

The L_1 norm is the trickiest of the norms to constrain using SOCP as well as the least common by far. The L_1 norm is defined as

$$\|E\|_1 \triangleq \int_0^1 |E(f)| \, df.$$

Again a uniform partition $\{f_k\}$ of $[0,1]$ is chosen, but now the auxiliary optimization variables $\{\beta_k\}$ are introduced as well. The following set of constraints then approximates the integral:

$$\begin{aligned} |E(f_k)| &\leq \beta_k, \quad k = 0, \ldots, K \\ \sum_k \beta_k \Delta f &\leq \delta. \end{aligned}$$

It is straightforward to see that at optimality all the constraints are met with equality, and so the second line represents the desired Riemann sum. Although this method introduces many new variables as well as constraints, the resulting problem is quite sparse and can be efficiently handled by most available solvers. A yet-smaller $\Delta f \approx 1/(60N)$ is required.

3. COMBINED NORMS

So what if we want to constrain a norm other than those listed in the last section? The general L_p norm, for example, is defined as

$$\|E\|_p \triangleq \left(\int_0^1 |E(f)|^p \, df \right)^{\frac{1}{p}}, \qquad p \geq 1,$$

and provides a continuum of norms. Although SOCP can handle only $p = \{1, 2, \infty\}$, there are two related methods that can be used to combine two norms to form a new norm. A single parameter allows the combined norm to provide a family similar to (but distinct from) the L_p norm. Starting with the most common method for combining the L_2 and L_∞ norms, this section then introduces an extension to Burdakov's ϵ-norm [3], and shows how the combined norms and their dual norms are related.

3.1 Convex Norm Combinations - The α-Norm

The most straightforward way to combine the L_2 and L_∞ norms is through a simple convex combination. Letting $\alpha \in [0,1]$, the *α-norm* is defined here as

$$\|E\|_\alpha \triangleq \alpha \|E\|_2 + (1-\alpha)\|E\|_\infty \tag{1}$$

This is essentially equivalent to the norm introduced in [2]. It is straightforward to verify the properties of a norm for the α-norm, as they all follow from the properties of the L_2 and L_∞ norms. As noted in [2], the unconstrained minimization

$$\min_{\mathbf{x}} \|E\|_\alpha \tag{2}$$

and the constrained minimization

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|E\|_2 \\ \text{s.t.} \quad & \|E\|_\infty \leq \delta \end{aligned} \tag{3}$$

have the same solution space and are thus functionally equivalent. An advantage of the unconstrained approach is that a solution is always obtained, as without constraints there can be no infeasible region. The constrained approach, on the other hand, allows the peak error to be explicitly fixed to meet given specifications. In any practical design there will be many iterations, and so the choice is really between iterating α and iterating the peak-error bound δ. Using SOCP allows either approach.

Some properties of the α-norm can be derived from the definition and the properties of the L_p norms. Clearly L_2 and L_∞ result as special cases of the α-norm with $\alpha = 1, 0$ respectively. The norm is monotonically nonincreasing in α, that is $\|E\|_{\alpha_1} \leq \|E\|_{\alpha_2}$ for $\alpha_1 \geq \alpha_2$, and is bounded by its component norms: $\|E\|_2 \leq \|E\|_\alpha \leq \|E\|_\infty$. Both these follow from the general $L_p[0,1]$ relation $\|E\|_p \leq \|E\|_q$ for $p \leq q$.

3.2 The ϵ-Norm

The ϵ-norm defined in [3] is actually a family of finite, real vector norms parameterized on $\epsilon \in [0,1]$, with l_∞ and l_2 as special cases when $\epsilon = 0$ and $\epsilon = 1$. The extension to complex-valued functions on an interval is straightforward and will be summarized following the framework of [3]. The ϵ-norm (and its dual) is of interest as it more closely approximates the L_p family of norms than does a simple convex combination of L_2 and L_∞ norms, yet the ϵ-norm can be efficiently optimized using SOCP (unlike L_p).

For a function $E : [0,1] \rightarrow \mathbb{C}$, we define the ϵ-norm of E, denoted $\|E\|_\epsilon$, as the solution $\nu^*(E)$ of

$$h(\nu, E) \triangleq \left\| \left[|E| - (1-\epsilon)\nu \right]_+ \right\|_2 - \epsilon\nu = 0. \qquad (4)$$

Clearly, for $\epsilon = 1$ we have $\nu^*(E) = \|E\|_2$. For $\epsilon = 0$ the solution is not unique, and the minimum solution $\nu^*(E) = \|E\|_\infty$ is chosen for convenience. (It can also be shown that $v^*(E) \rightarrow \|E\|_\infty$ for $\epsilon \rightarrow 0^+$.) For the general case $\epsilon \in (0,1)$, we must show that $\nu^*(E)$ is unique and possesses the properties of a norm. To show uniqueness, first note that $h(\nu, E)$ is a continuous, monotonically decreasing function of ν. For any $\nu > \|E\|_\infty/(1-\epsilon)$, we see that $h(\nu, E) = -\epsilon\nu < 0$, and since $h(0, E) = \|E\|_2 \geq 0$ as well, there must exist a unique solution $\|E\|_\epsilon \triangleq v^*(E)$. The positivity of $v^*(E)$ was guaranteed by definition. Positive definiteness (in the sense of L_p) follows directly from (4) and the positive definiteness of the L_2 norm. Homogeneity follows from the relation $h(|\alpha|\nu, \alpha E) = |\alpha| h(\nu, E)$, which implies that $h(|\alpha|\nu^*(E), \alpha E) = 0$ and $\nu^*(\alpha E) = |\alpha|\nu^*(E)$. It remains to show the triangle inequality, for which we use the fact that $[A+B]_+ \leq [B]_+ + [B]_+$ (in a pointwise sense):

$$\begin{aligned}
&h(\nu^*(E) + \nu^*(F), E+F) + \epsilon(\nu^*(E) + \nu^*(F)) \\
&= \left\| \left[|E+F| - (1-\epsilon)(\nu^*(E) + \nu^*(F)) \right]_+ \right\|_2 \\
&\leq \left\| \left[|E| - (1-\epsilon)\nu^*(E) \right]_+ \right\|_2 + \left\| \left[|F| - (1-\epsilon)\nu^*(F) \right]_+ \right\|_2 \\
&= h(\nu^*(E), E) + h(\nu^*(F), F) + \epsilon(\nu^*(E) + \nu^*(F)) \\
&= \epsilon(\nu^*(E) + \nu^*(F)).
\end{aligned}$$

Thus, $h(\nu^*(E) + \nu^*(F), E+F) \leq 0$, and from the monotonicity of $h(\nu, E+F)$ in ν we have $\nu^*(E+F) \leq \nu^*(E) + \nu^*(F)$ as desired. The ϵ-norm shares the monotonicity properties of the α-norm and the L_p norms: $\|E\|_2 \leq \|E\|_{\epsilon_1} \leq \|E\|_{\epsilon_2} \leq \|E\|_\infty$ for $\epsilon_1 \geq \epsilon_2$. This follows from the monotonicity of (4) in ϵ.

The ϵ-norm has a convenient geometric interpretation, which ultimately leads to a SOCP solution. The norm induces a unique decomposition $E = E^\epsilon + E^{1-\epsilon}$, where $\|E^\epsilon\|_2 = \epsilon\|E\|_\epsilon$ and $\|E^{1-\epsilon}\|_\infty = (1-\epsilon)\|E\|_\epsilon$. This can be seen by considering Fig. 1, the plot of a representative function $|E|$. The dashed line represents the value of $(1-\epsilon)\nu$ in (4). The portion of $|E|$ below the line is a flat-topped function with L_∞ norm $(1-e)\nu$, while the portion above the line is the function inside the L_2 norm in (4). When $\nu = \|E\|_\epsilon$ the value of the L_2 norm is $\epsilon\|E\|_\epsilon$, the function above the line is $|E^\epsilon|$, and the function below the line is $|E^{1-\epsilon}|$. Both E^ϵ and $E^{1-\epsilon}$ have the same phase as E. This decomposition leads to the following characterization of the unit ball

$$\{E : \|E\|_\epsilon \leq \nu\} = \{U + V : \|U\|_2 \leq \epsilon\nu,\ \|V\|_\infty \leq (1-\epsilon)\nu\} \qquad (5)$$

of the ϵ-norm as the convex combination of the unit balls of the L_2 and L_∞ norms. This leads to the following set of constraints that together effectively bound $\|E\|_\epsilon$ by ν:

$$\begin{aligned}
\|U\|_2 &\leq \epsilon\nu \\
\|E - U\|_\infty &\leq (1-\epsilon)\nu
\end{aligned}$$

where the individual norm constraints can be implemented using SOCP as previously discussed, and U represents an auxiliary function implemented as a set of complex optimization variables on the discretized set of frequencies.

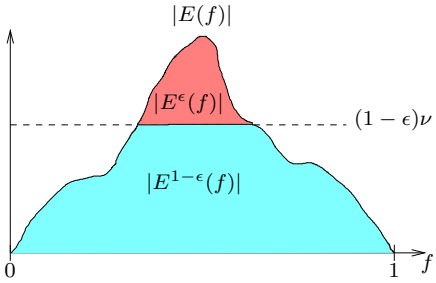

Figure 1: Graphical representation of the ϵ-decomposition.

3.3 The Dual Norms of the α- and ϵ-Norms

So far we have only considered combined norms that span from L_2 to L_∞. But, just as L_p for $p \geq 2$ has as its dual norm L_q with $q = p/(p-1) \leq 2$, the two different combined L_∞/L_2 norms have dual norms involving L_1 and L_2. The ϵ-norm dual is

$$\begin{aligned}
\|F\|_\epsilon^D &= \max_{\|E\|_\epsilon \leq 1} (F, E) \\
&= \max_{\substack{\|U\|_2 \leq \epsilon \\ \|V\|_\epsilon \leq (1-\epsilon)}} (F, U) + (F, V) \\
&= \epsilon \max_{\|U\|_2 \leq 1} (F, U) + (1-\epsilon) \max_{\|V\|_\infty \leq 1} (F, V) \\
&= \epsilon\|F\|_2 + (1-\epsilon)\|F\|_1,
\end{aligned}$$

a convex combination of the L_2 and L_1 norms. This suggests a rather satisfying symmetry that is completed with the dual of the α-norm, whose unit ball is a convex combination of the unit balls of the L_2 and L_1 norms. This norm cannot be defined in the form of (4), which depends on the properties of the L_∞ norm, but we can constrain it by

$$\begin{aligned}
\|U\|_2 &\leq \alpha\nu \\
\|E - U\|_1 &\leq (1-\alpha)\nu.
\end{aligned}$$

So far two combined norm/dual norm pairs have been formed from three special cases of the L_p norm and two methods of combination, but one pair remains. Revisiting (4), it is clear that if instead we define

$$h(\nu, E) \triangleq \left\| \left[|E| - (1-\epsilon)\nu \right]_+ \right\|_1 - \epsilon\nu = 0. \qquad (6)$$

then the analysis proceeds essentially unchanged, except that the result is a family of norms that spans the whole range from L_1 to L_∞ as does its dual, $\epsilon\|E\|_\infty + (1-\epsilon)\|E\|_1$. Although these norms do not resemble L_p in general, they have the advantage that they can be implemented using only linear programming when E is real-valued.

4. EXAMPLES

In this section a simple lowpass filter is optimized using the various combined-norm SOCP techniques presented in the previous section and compared to the general L_p solution. The filter has 31 real coefficients with no symmetry and a desired delay of 10 samples that ensures a complex frequency response. The passband and stopband cover the intervals $[0, 0.225]$ and $[0.3, 0.5]$, with the remainder of the frequency-response period defined by symmetry. An error weighting of one in the passband and five in the stopband

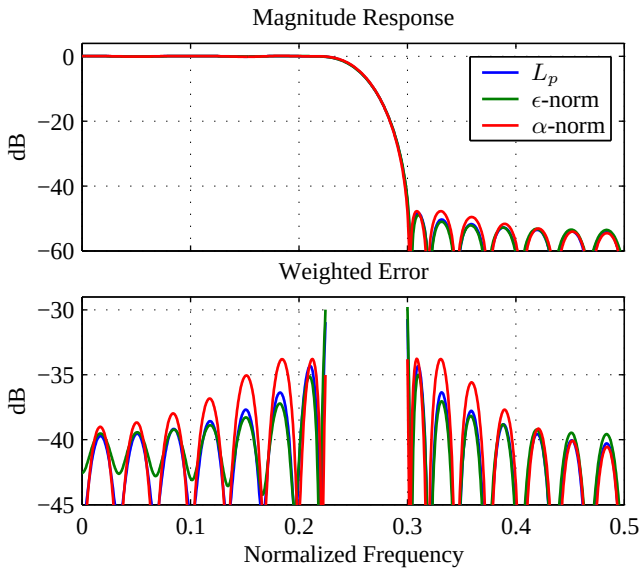

Figure 2: Mixed L_2/L_∞-optimized FIR filter responses.

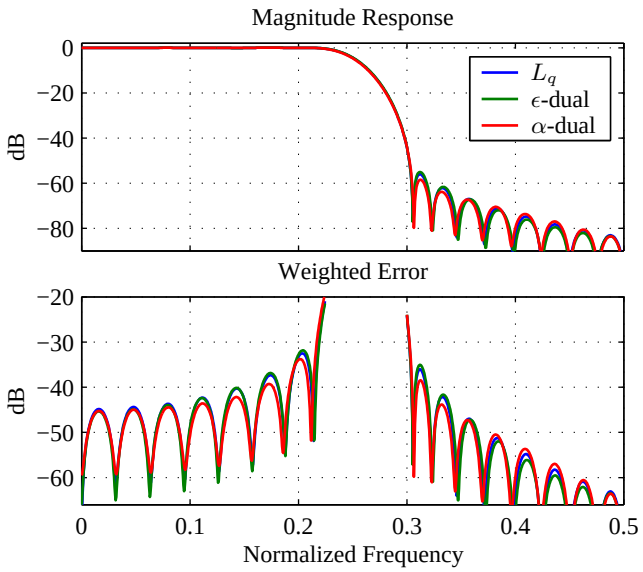

Figure 3: Mixed L_1/L_2-optimized FIR filter responses.

was applied. A general-purpose solver was used to minimize the L_4 norm of the error, and SOCPs were used to minimize the α- and ϵ-norms with $\alpha = 0.75$ and $\epsilon = 0.18$ determined empirically to give the best visual approximation to L_4 (an admittedly subjective measure). The resulting responses are shown in Fig. 2.

Qualitatively, of the two mixed norms the minimum ϵ-norm response more resembles the minimum L_4 response. A characteristic feature of the α-norm is that even for $\alpha > 0$, the optimized error still has a L_∞-like equiripple response near the transition region. The equivalent constrained formulation of (3) suggests why: since the largest error ripples typically occur near the transition, the peak-error constraint is active in this region. Away from the transition, the peak constraint is inactive, and the response resembles a pure L_2 response. Decreasing α will grow the equiripple error region. Thus the effect of combining the L_2 and L_∞ norms in this way is to partition the frequency axis into L_2 and L_∞ regions. The ϵ-norm, on the other hand, like the L_p norm, smoothly transitions from equiripple errors to L_2-type errors everywhere on the frequency axis, and generally has no equiripple errors for $\epsilon > 0$. This is because the L_2 and L_∞ constraints in the SOCP are not applied to the entire error response, but to components of it.

The same design was also optimized using the duals of the norms from the previous example. The result is shown in Fig. 3. The ϵ-dual response better approximates the L_q response ($q = 4/3$), although the difference is less easily characterized.

5. CONCLUSIONS

The goal of this paper was to demonstrate the ability of SOCP to constrain more interesting FIR filter error norms than the usual subset of L_p. In particular, Burdakov's vector ϵ-norm was extended and shown to be related to a more familiar norm combination, and it was shown that these combining techniques in fact generate three different norm/dual norm pairs that can be constrained with SOCP. The examples provide a qualitative comparison of the resulting filter responses. The real strength of SOCP is not its ability to solve these particular unconstrained optimizations, although it can do so efficiently. Rather, it lies in the tremendous flexibility it provides, allowing an arbitrary number of constraints of the sort presented here, while at the same time enforcing a uniform structure that allows for efficient solution.

REFERENCES

[1] J. Adams and J. Sullivan, "Peak-constrained least-squares optimization," *IEEE Trans. Signal Processing*, vol. 46, no. 2, pp. 306–321, Feb. 1998.

[2] W. Lertniphonphun and J. McClellan, "Unified design algorithm for complex FIR and IIR filters," in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing*, Salt Lake City, UT, May 2001.

[3] O. Burdakov and B. Merkulov, "On a new norm for data fitting and optimization problems," Tech. Rep., Linköping University, 2000.

[4] R.J. Vanderbei, "LOQO: An interior point code for quadratic programming," *Optimization Methods and Software*, vol. 12, pp. 451–484, 1999.

[5] Jos F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," in *Optimization Methods and Software*, vol. 11–12, pp. 625–653. 1999, Special issue on Interior Point Methods.

[6] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and its Applications*, vol. 284, pp. 193–228, Nov. 1998.

[7] J. O. Coleman and D. P. Scholnik, "Design of nonlinear-phase FIR filters with second-order cone programming," in *Proc. 1999 Midwest Symp. on Circuits and Systems (MWSCAS '99)*, Las Cruces NM, Aug. 1999.

[8] J. O. Coleman, "Systematic mapping of quadratic constraints on embedded FIR filters to linear matrix inequalities," in *Proc. 1998 Conf. on Information Sciences and Systems (CISS '98)*, Princeton, NJ, Mar. 1998.