INTERIOR-POINT ALGORITHMS FOR SUM-OF-SQUARES OPTIMIZATION OF MULTIDIMENSIONAL TRIGONOMETRIC POLYNOMIALS

Tae Roh[*]

Electrical Engineering Department
University of California, Los Angeles
Los Angeles, CA, U.S.A.

Bogdan Dumitrescu[†]

International Center for Signal Processing
Tampere University of Technology
Tampere, Finland

Lieven Vandenberghe[*]

Electrical Engineering Department
University of California, Los Angeles
Los Angeles, CA, U.S.A.

ABSTRACT

A wide variety of optimization problems involving nonnegative polynomials or trigonometric polynomials can be formulated as convex optimization problems by expressing (or relaxing) the constraints using sum-of-squares representations. The semidefinite programming problems that result from this formulation are often difficult to solve due to the presence of large auxiliary matrix variables. In this paper we extend a recent technique for exploiting structure in semidefinite programs derived from sum-of-squares expressions to multivariate trigonometric polynomials. The technique is based on an equivalent formulation using discrete Fourier transforms and leads to a very substantial reduction in the computational complexity. Numerical results are presented and a comparison is made with general-purpose semidefinite programming algorithms. As an application, we consider a two-dimensional FIR filter design problem.

Index Terms— Optimization methods, Multidimensional digital filters, Discrete transforms

1. INTRODUCTION

Recently, there has been a great deal of interest in semidefinite programming (SDP) for optimization problems over polynomials or pseudo-polynomials (*e.g.*, trigonometric polynomials). The basic idea is to replace the constraint that a polynomial is nonnegative on (a subset of) its domain by the constraint that it is a sum-of-squares (SOS). The nonnegativity of the polynomial and the SOS condition are equivalent for univariate polynomials. In the multivariate case useful sufficient conditions for nonnegativity are obtained. An optimization problem with SOS constraints is equivalent to an SDP, a convex problem which can be solved efficiently (in polynomial time) using interior-point solvers.

A wide variety of applications can be found in signal processing, control, combinatorial and global optimization [1, 2, 3, 4, 5]. At the moment, however, these applications are often limited by the large size of the SDPs that result from the SOS formulation. This is due to the need to introduce large matrix variables with dimensions that grow rapidly with the number of variables in the multivariate polynomials. Hence there is a need for specialized SDP algorithms that exploit structure in multivariate SOS optimization problems.

Most research on exploiting structure in SDP has focused on sparsity of the coefficient matrices [6]. Another approach based on exploiting (dense) rank-one structure was studied in [7, 8, 9] and found to be very well-suited for SOS optimization. In this paper we extend the techniques proposed in [9] to multivariate trigonometric polynomials. Our focus on trigonometric polynomials is motivated by applications in signal processing [2, 3], and by the theoretical advantages of trigonometric basis functions in SOS optimization [10].

2. SOS RELAXATION OF POSITIVE POLYNOMIALS

For the sake of simplicity, we will limit the discussion to bivariate trigonometric polynomials. (However, all the results extend to multivariate trigonometric polynomials.)

Let R be a bivariate trigonometric polynomial of degree $\mathbf{n} = (n_1, n_2) \in \mathbf{Z}^2$, with real symmetric coefficients $x_{\mathbf{k}} = x_{-\mathbf{k}}$:

$$R(\omega) = \sum_{\mathbf{k}=-\mathbf{n}}^{\mathbf{n}} x_{\mathbf{k}} e^{-j\mathbf{k}^T\omega}. \tag{1}$$

If we collect the independent coefficients $x_{\mathbf{k}}$ (in some order) in a vector x, this can be expressed as $R(\omega) = x^T f(\omega)$ where f is a vector of basis functions. A fundamental result states that if R is positive, then it can be expressed as an SOS of trigonometric polynomials,

$$x^T f(\omega) = \sum_{k=1}^{r} |a_k^T v(\omega) + j b_k^T w(\omega)|^2 \tag{2}$$

where v and w are vectors of two-dimensional cosine and sine basis functions (see [3, 10]). Equivalently, if R is positive, it

[*] roh@ee.ucla.edu, vandenbe@ee.ucla.edu. Research supported by NSF under grant ECS-0524663.

[†] bogdand@cs.tut.fi. Work supported by the Academy of Finland, project No. 213462, Finnish Centre of Excellence program (2006 - 2011). B. Dumitrescu is on leave from the Department of Automatic Control and Computers, "Politehnica" University of Bucharest, Romania.

can be expressed as

$$x^T f(\omega) = v(\omega)^T X v(\omega) + w(\omega)^T Y w(\omega), \qquad (3)$$

for some $X = \sum_k a_k a_k^T \succeq 0$ and $Y = \sum_k b_k b_k^T \succeq 0$. The dimensions of v and w can be arbitrarily high, but to obtain a sufficient condition for the positivity of R, they can be limited to a finite value given, for example, by the degree of R. In that case, the dimensions of v and w are roughly $N/2$, where $N = (n_1+1)(n_2+1)$. The "parity" of the degree $\mathbf{n}$ determines the exact dimensions [3].

Equation (3) is a set of linear equations in the coefficients x and the matrices X and Y. By equating coefficients of the same terms on both sides, it can be written in the form

$$x_{\mathbf{k}} = \mathbf{tr}(T_{\mathbf{k}} X) + \mathbf{tr}(H_{\mathbf{k}} Y), \qquad (4)$$

where $T_{\mathbf{k}}$ and $H_{\mathbf{k}}$ are sparse symmetric matrices (see [3] for details). This observation allows us to formulate the (bounded-degree) SOS constraint (2), which is a sufficient condition for positivity of R, as a semidefinite programming constraint [2, 3, 10]. Similar techniques are used in recent SDP relaxations of multivariate nonnegative polynomials [4, 5, 11]. A parametrization similar to (4) is used in these works, and general-purpose SDP software such as SeDuMi [12] is used to solve the resulting SDPs. Unfortunately, although the SDP data matrices associated with (4) are very sparse, this sparsity is only exploited to a limited extent by current solvers. For $n = \max\{n_1, n_2\}$, the complexity of solving an SDP with constraints (4) using existing general-purpose software is typically close to $O(n^8)$.

An alternative formulation based on discrete transforms was recently proposed in [9, 8], and shown to be very effective for single-variable SOS optimization problems. The technique also applies to multivariate SOS expressions. We first note that (3) can be replaced by a finite set of linear equations, by sampling both sides on an appropriately defined and sufficiently dense grid of M points,

$$x^T f(\omega_i) = v(\omega_i)^T X v(\omega_i) + w(\omega_i)^T Y w(\omega_i)$$

for $i = 1, \ldots, M$. In matrix form, this is equivalent to

$$Fx = \mathbf{diag}(VXV^T + WYW^T).$$

The matrices F, V, and W represent discrete transforms that map the coefficients of (pseudo-)polynomials to their sample values. (In our application, they are two-dimensional DFT, DCT, and DST matrices, respectively.) From the sample values $y = Fx$, the coefficient vector x can be obtained via the corresponding inverse discrete transform $x = Gy$. This leads to the following alternative to (4):

$$x = G\,\mathbf{diag}(VXV^T + WYW^T), \quad X, Y \succeq 0. \qquad (5)$$

The formulation (5) involves dense matrices. However, as we will see in section 4, simple properties of the $\mathbf{diag}$ operator can be exploited, leading to a substantial reduction in computational complexity. For typical problems with bivariate trigonometric polynomials and $n = \max\{n_1, n_2\}$, we will obtain a complexity of roughly $O(n^6)$.

3. APPLICATIONS

In the previous section, we reviewed how nonnegativity constraints on polynomials can be reformulated or relaxed as linear matrix inequalities (LMI), via SOS expressions. As a consequence, we can formulate a wide variety of optimization problems involving nonnegative polynomials as SDPs. As an illustration, we discuss a two-dimensional FIR filter design problem [2].

To represent the spectral mask constraints involved in the filter design we refer to the following observation. We are interested in sufficient conditions that guarantee that a trigonometric polynomial R is positive on a set of the form

$$\mathcal{D} = \{\omega \in [-\pi, \pi]^2 \mid D_i(\omega) \geq 0,\ i = 1, \ldots, l\}, \qquad (6)$$

where D_i is a trigonometric polynomial. An obvious sufficient condition is that it can be expressed as

$$R(\omega) = S_0(\omega) + \sum_{i=1}^{l} D_i(\omega) S_i(\omega), \qquad (7)$$

where S_i, $i = 0, \ldots, l$, are sums of squares of trigonometric polynomials. The condition is also necessary, but the degrees of the SOS may be arbitrarily high. By expressing the coefficients of S_i in the form (5), we can write (7) as a linear equation in the coefficients of R and $2(l+1)$ positive semidefinite matrices.

As a specific example we consider the problem of designing a 2-D zero-phase FIR filter

$$H(\omega) = \sum_{\mathbf{k}=-\mathbf{n}}^{\mathbf{n}} h_{\mathbf{k}} e^{-j\mathbf{k}^T\omega}$$

with maximum attenuation $\delta_{\rm s}$ in the stopband $\mathcal{D}_{\rm s}$, and subject to a maximum allowable ripple $\delta_{\rm p}$ in the passband $\mathcal{D}_{\rm p}$. The passband and stopband are both parameterized using expressions of the form (6). The optimization problem is

$$\begin{array}{ll} \text{minimize} & \delta_{\rm s} \\ \text{subject to} & |1 - H(\omega)| \leq \delta_{\rm p}, \quad \omega \in \mathcal{D}_{\rm p} \\ & |H(\omega)| \leq \delta_{\rm s}, \quad \omega \in \mathcal{D}_{\rm s}, \end{array} \qquad (8)$$

where $\delta_{\rm s}$ and the filter coefficients of H are the optimization variables. To solve the problem (8), we expand the constraints as

$$\begin{aligned} R_1(\omega) &= H(\omega) - 1 + \delta_{\rm p} \geq 0, \quad \omega \in \mathcal{D}_{\rm p} \\ R_2(\omega) &= 1 - H(\omega) + \delta_{\rm p} \geq 0, \quad \omega \in \mathcal{D}_{\rm p} \\ R_3(\omega) &= H(\omega) + \delta_{\rm s} \geq 0, \quad \omega \in \mathcal{D}_{\rm s} \\ R_4(\omega) &= H(\omega) - \delta_{\rm s} \geq 0, \quad \omega \in \mathcal{D}_{\rm s}. \end{aligned}$$

Each positive polynomial R_i can now be represented in terms of SOS polynomials as in (7). Using the LMI characterization (5), we arrive at an SDP of the form

$$\begin{array}{ll} \text{minimize} & q^T y \\ \text{subject to} & A\,\mathbf{diag}(CXC^T) + By = b \\ & X \succeq 0, \end{array} \tag{9}$$

with a matrix variable X and a vector variable y. The problem parameters A, B, C, as well as the variable X are block matrices with a small number ($8(l+1)$ or $4(l+1)$) of diagonal blocks of order $O(n^2)$, where $n = \max\{n_1, n_2\}$.

We refer to [2] for an overview of other applications, such as nonlinear-phase magnitude filter design.

4. SDP ALGORITHM

The most well-known class of SDP algorithms is called the primal-dual *interior-point method* (PD-IPM). Typically, PD-IPMs take roughly 10 to 50 iterations to reach a solution with a high accuracy. Their key feature is a set of nonlinear equations known as *central path* equations. At each iteration of the algorithm the central path equations are linearized to form a large system of linear equations referred to as *Newton equations*. The computation time spent to construct and solve the Newton equations dominates the overall computing time.

The algorithm used in this paper is the extension of the method proposed in [9], which is based on the interior-point algorithm described in [13]. We restrict the discussion to the solution of the Newton equations.

The Newton equations for (9) are

$$\begin{aligned} -T^{-1}\Delta X T^{-1} + C^T\,\mathbf{diag}(A^T\Delta z)C &= R, & (10)\\ A\,\mathbf{diag}(C\Delta X C^T) + B\Delta y &= r_1, & (11)\\ B^T\Delta z &= r_2, & (12) \end{aligned}$$

where the *scaling matrix* T is positive definite. The values of T and the righthand sides change at each iteration. By eliminating the variable ΔX from the first equation and applying the identity $\mathbf{diag}(P\,\mathbf{diag}(u)Q^T) = (P \circ Q)u$, the set of equations (10) through (12) is reduced to

$$\begin{bmatrix} H & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta y \end{bmatrix} = \begin{bmatrix} r_3 \\ r_2 \end{bmatrix}, \tag{13}$$

where

$$H = A\left((CTC^T) \circ (CTC^T)\right) A^T.$$

In this equation '$\circ$' denotes Hadamard (component-wise) product, so the cost of constructing H grows cubically with the matrix dimensions. It is therefore of the same order as the cost of solving the system. This is an improvement by an order of magnitude over general-purpose implementations, which do not exploit the specific structure in the equality constraints of (9), and for which computing H is more expensive than solving the reduced Newton system (13). We also note that the matrix-matrix products in the definition of H correspond to two-dimensional discrete transforms. Exploiting this fact, we can reduce the complexity of computing H even further by employing fast transforms such as FFT.

n	SeDuMi + [3]	DT SDP
5	0.07	0.10
7	0.21	0.37
9	1.03	1.15
11	3.15	2.97
13	9.16	6.78
15	24.4	14.1
17	49.1	26.2
19		47.2
21		80.6
23		132
25		212

Table 1. Solve time per iteration (in seconds) for problem (8).

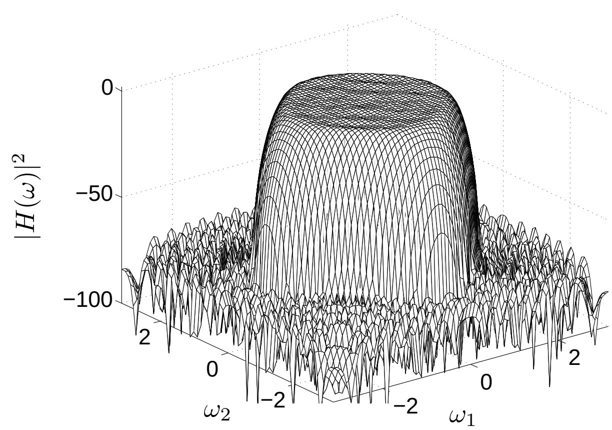

Fig. 1. Solution of the 2D filter design problem (8).

5. RESULTS

We revisit the filter problem (8) with the design parameters $\delta_\mathrm{p} = 0.05$,

$$\begin{aligned} \mathcal{D}_\mathrm{p} &= \{\omega \in [-\pi,\pi]^2 \mid D_\mathrm{p} = \cos\omega_1 + \cos\omega_2 - c_\mathrm{p} \geq 0\} \\ \mathcal{D}_\mathrm{s} &= \{\omega \in [-\pi,\pi]^2 \mid D_\mathrm{s} = c_\mathrm{s} - \cos\omega_1 - \cos\omega_2 \geq 0\}. \end{aligned}$$

The specification produces a lowpass filter, and the choice for the values c_p and c_s determines the "steepness" of the transition band. With values $c_\mathrm{p} = 1$, $c_\mathrm{s} = 0.3$, and $n_1 = n_2 = 11$, we obtain the filter shown in figure 1. Its optimal attenuation is approximately 69 dB.

For the same problem, but with varying filter lengths ($n = n_1 = n_2$), we compare the computational complexity per iteration of the discrete-transform-based SDP formulation to that

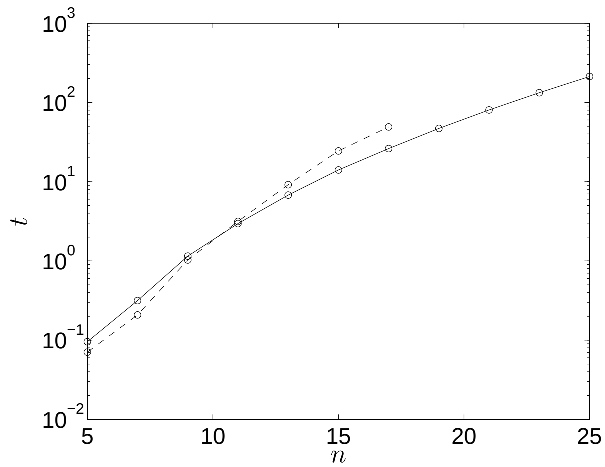

Fig. 2. Plot of the results in table 1.

of the general-purpose SDP solver (SeDuMi) with the SDP formulation discussed in [3]. There are 16 matrix variables of size roughly $(n/2)^2$.

The CPU times per iteration for the two methods are summarized in table 1 and figure 5. Both solvers reached the solutions in 16 to 26 iterations for all problem instances. It can be verified that the complexity of SeDuMi is between $O(n^7)$ and $O(n^8)$, while the discrete-transform-based SDP algorithm is between $O(n^5)$ and $O(n^6)$. For problems with $n \geq 19$ the computing times by SeDuMi could not be recorded due to "out-of-memory" error by Matlab under the computing environment.

The experiments were conducted in Matlab 7.1 on a 3.0-Ghz Pentium-4 PC with 3 GB of memory.

6. CONCLUSION

We have derived a discrete-transform-based SDP formulation of a convex optimization problem over positive multivariate trigonometric polynomials. The SDP formulation has the advantage of leading to a customized interior-point algorithm implementation that reduces the computational complexity significantly compared to general-purpose SDP solvers. We applied the results to a two-dimensional lowpass filter design problem and benchmarked an interior-point implementation based on the new formulation against a popular general-purpose SDP solver.

7. REFERENCES

[1] B. Alkire and L. Vandenberghe, "Convex optimization problems involving finite autocorrelation sequences," *Mathematical Programming Series A*, vol. 93, pp. 331–359, 2002.

[2] B. Dumitrescu, "Trigonometric polynomials positive on frequency domains and applications to 2-D FIR filter design," *IEEE Trans. Signal Proc.*, vol. 54, no. 11, pp. 4282–4292, Nov. 2006.

[3] B. Dumitrescu, "Gram pair parameterization of multivariate sum-of-squares trigonometric polynomials," *EUSIPCO*, Florence, Italy, Sept. 2006.

[4] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[5] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Mathematical Programming Series B*, vol. 96, pp. 293–320, 2003.

[6] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity," *SIAM J. on Optimization*, vol. 17, no. 1, pp. 218–242, 2006.

[7] S. J. Benson, Y. Ye, and X. Zhang, "Solving large-scale sparse semidefinite programs for combinatorial optimization," *SIAM J. on Optimization*, vol. 10, no. 2, pp. 443–461, 2000.

[8] J. Löfberg and P. A. Parrilo, "From coefficients to samples: a new approach to SOS optimization," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2004, pp. 3154–3159.

[9] T. Roh and L. Vandenberghe, "Discrete transforms, semidefinite programming and sum-of-squares representations of nonnegative polynomials," *SIAM J. on Optimization*, vol. 16, no. 4, pp. 939–964, 2006.

[10] A. Megretski, "Positivity of trigonometric polynomials," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2003, pp. 3814–3817.

[11] Y. Nesterov, "Squared functional systems and optimization problems," in *High Performance Optimization Techniques*, J. Frenk, C. Roos, T. Terlaky, and S. Zhang, Eds., pp. 405–440. Kluwer Academic Publishers, 2000.

[12] J.F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11–12, pp. 625–653, 1999, Version 1.1 available from `http://sedumi.mcmaster.ca`.

[13] M. J. Todd, K. C. Toh, and R. H. Tütüncü, "On the Nesterov-Todd direction in semidefinite programming," *SIAM J. on Optimization*, vol. 8, no. 3, pp. 769–796, 1998.