MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

A Region-Dividing Approach to Robust Semidefinite Programming and Its Error Bound

Yasuaki OISHI

METR 2006–10 February 2006

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.i.u-tokyo.ac.jp/mi/mi-e.htm

A Region-Dividing Approach to Robust Semidefinite Programming and Its Error Bound*

Yasuaki Oishi†

A new asymptotically exact approach is presented for robust semidefinite programming, where coefficient matrices polynomially depend on uncertain parameters. The parameter region is divided into subregions to make an approximate problem for a given robust semidefinite programming problem. The optimal value of the approximate problem converges to that of the original problem as the resolution of the division becomes finer. An advantage of this approach is that an upper bound on the approximation error is available before solving the approximate problem. This bound shows how the approximation error depends on the resolution of the division. Furthermore, it leads to the construction of an efficient division that attains small approximation error with low computational complexity. Numerical examples show usefulness of the present approach. In particular, an exact optimal value is often found with a division of finite resolution. Some useful properties related to polynomial optimization are also presented.

Keywords: robust semidefinite programming, polynomial optimization, linear matrix inequalities, approximation, conservatism, computational complexity.

1. Introduction

Robust semidefinite programming (robust SDP in short) is the optimization of a linear objective function subject to linear matrix inequalities (LMIs in short) whose coefficients depend on uncertain parameters. Its research in the framework of mathematical programming started in the independent works of Ben-Tal–Nemirovski [3] and El Ghaoui–Oustry–Lebret [18] though essentially the same problem had been considered in robust control [8]. Surveys are found in [2, 17, 4, 6]. Robust SDP has many applications such as robust control, robust design of antenna arrays, and robust truss design. Furthermore, polynomial optimization [23, 29], which attracts the interest of many researchers, is reduced to robust SDP.

A robust SDP problem is difficult to solve in general due to its semi-infinite nature, that is, this problem essentially has infinitely many LMI constraints corresponding to different values of the parameters. Hence, an approximate approach has been taken: A solvable SDP problem is constructed such that its feasible region is included in that of the original robust SDP problem. This approach is conservative in that there is usually a gap between the optimal values of

*February 9, 2006

†Department of Mathematical Informatics, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan; phone: +81-3-5841-6921; fax: +81-3-5841-6886; email: oishi@mist.i.u-tokyo.ac.jp

the approximate problem and the original problem. Ben-Tal–Nemirovski [3, 5, 6] proposed an approximation scheme in the case of affine parameter dependence and gave an upper bound on the approximation error. They also showed that the original robust SDP problem is NP-hard even in this special case. El Ghaoui–Oustry–Lebret [18] proposed a different approximation scheme in the case of rational parameter dependence.

Recently, asymptotically exact approaches were proposed in the case that parameter dependence is polynomial. In particular, Ohara–Sasaki [28] and Bliman [7] proposed an approach based on the Kalman–Yakubovich–Popov lemma, Scherer [31] an approach based on Pólya's theorem, and Scherer–Hol [32] an approach based on sum of squares (SOS). See also the works of Lasserre [23] and Parrilo [29] for the SOS technique. The approximate problems considered in these approaches are related to the existence of some polynomial in the parameters. As the assumed degree of the polynomial increases, the optimal value of the approximate problem converges to that of the original robust SDP problem. More interestingly, it is reported in the SOS approach that the optimal value of the original problem is often obtained with a polynomial of finite degree. See [31] for a general theory on this type of approximations and on verification of exactness.

Although these approaches are quite attractive, they have some issues to be settled. (i) How the approximation error depends on the assumed degree of a polynomial is not well understood quantitatively. Indeed, this dependence is really non-intuitive [23, 29]. (ii) It is not clear how one should increase the degree of a polynomial. When one is not satisfied with the quality of the obtained approximation, he is to increase the degree of the polynomial. This introduces many new terms into the polynomial especially when the parameter dimension is high. If possible, one would like to introduce only necessary terms to suppress the computational complexity. It is not clear, however, how to do it.

The objective of this paper is to investigate another approach proposed by Emoto–Oishi [19] with particular focus on the above mentioned issues. This approach provides an asymptotically exact sequence of approximate problems just as the existing approaches. A difference is that it is based on a division of the parameter region and improves the quality of approximation by refining the division. An advantage of this approach is that an upper bound on the approximation error is available before solving the approximate problem. Although this bound can be conservative as in Section 6, it reveals a quantitative property of the asymptotic exactness. In particular, this bound shows how the approximation error depends on the resolution of the division. Moreover, it enables us to make a good division, which attains a small approximation error with small computational load.

The approach to be discussed is a generalization of two techniques. Both of them are developed for less conservative approximation of a special class of robust SDP problems in the context of robust control. The first technique was proposed by Masubuchi–Shimemura [26] using descriptor forms. Its application to robust and gain-scheduled control is found in [25, 22]. A closely related approach is presented in [11]. The second technique is matrix dilation, which was proposed by de Oliveira–Bernussou–Geromel [12] and was extended in many directions [14, 1, 33, 34, 13, 15, 16]. It is notable that division of the parameter region was considered in the context of matrix dilation [27] though its scope was still limited to a special class of robust SDP problems. Based on these results, Emoto–Oishi [19] proposed an approach to general robust SDP problems with a proof of its asymptotic exactness. The mentioned results on an error bound are new contribution of the present paper. They seem to be new even in the special class discussed before.

This paper is constructed as follows. Section 2 introduces robust SDP with some examples. Section 3 gives the region-dividing approach, which is to be investigated. Section 4 is the main section of this paper and gives an upper bound on the approximation error. An efficient division is discussed in Section 5 while numerical examples are provided in Section 6. Section 7 focuses on a special robust SDP problem related to polynomial optimization and presents some useful properties. Section 8 concludes the paper.

The symbol $\mathbb{R}^n$ stands for the set of n-dimensional real vectors. The symbol $^{\mathrm{T}}$ denotes the transpose of a matrix or a vector. We let $O_{n\times m}$ and I_n designate the $n \times m$ zero matrix and the $n \times n$ identity matrix, respectively. The sizes of these matrices are omitted when they are clear from the context. The maximum singular value of a matrix A is written as $\overline{\sigma}(A)$. For a real symmetric matrix A, the inequality $A \succeq O$ means that A is positive semidefinite, that is, $x^{\mathrm{T}}Ax$ is nonnegative for any real vector x. Similarly, $A \succ O$ expresses that A is positive definite. For two real symmetric matrices A and B, the inequalities $A \succeq B$ and $A \succ B$ mean $A - B \succeq O$ and $A - B \succ O$, respectively. The Kronecker product of two (not necessarily symmetric) matrices $A = (a_{ij})$ and B is defined as

$$A \otimes B := \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix}.$$

2. Robust Semidefinite Programming

A robust SDP problem considered in this paper is written as follows:

$$\begin{aligned} P: \quad & \text{minimize} \quad c^{\mathrm{T}}x \\ & \text{subject to} \quad E(x) \succeq O, \quad F(x,\theta) \succeq O \quad (\forall \theta \in \Theta). \end{aligned}$$

Here, $c \in \mathbb{R}^n$ is a given nonzero vector; $x = [x_1\ x_2\ \cdots\ x_n]^{\mathrm{T}} \in \mathbb{R}^n$ is an optimization variable; $\theta = [\theta_1\ \theta_2\ \cdots\ \theta_p]^{\mathrm{T}} \in \Theta$ is an uncertain parameter; its domain Θ is a polytope, *i.e.*, a bounded polyhedron, in $\mathbb{R}^p$ with a nonempty interior; $E(x)$ is a function affine in x, whose value is an $\ell \times \ell$ real symmetric matrix; $F(x, \theta)$ is a function affine in x and polynomial in θ, whose value is an $m \times m$ real symmetric matrix. We denote by d_i the maximum degree of $F(x, \theta)$ in θ_i for $i = 1, 2, \ldots, p$. We write $D := \prod_{i=1}^{p}(d_i + 1)$.

Here are two examples for robust SDP problems.

Example 1 (Robust stability analysis [8])**.** We analyze robust stability of the system $(\mathrm{d}/\mathrm{d}t)\xi(t) = A(\theta(t))\xi(t)$. Here, $A(\theta(t))$ is a matrix whose elements are rational functions of $\theta(t)$; $\theta(t)$ is a time-varying parameter that continuously moves in a polytope Θ, which is often a multi-dimensional interval $\prod_{i=1}^{p}[\underline{\theta}_i, \overline{\theta}_i]$. This system is asymptotically stable for any time variation of $\theta(t)$ if and only if the following problem has a negative minimum value:

$$\begin{array}{ll} \text{minimize} & y \\ \text{subject to} & X - I \succeq O, \quad yI - XA(\theta) - A(\theta)^{\mathrm{T}}X \succeq O \quad (\forall \theta \in \Theta). \end{array}$$

This problem can be written in the form of P. Recall that the elements of $A(\theta)$ are rational functions of θ. If we consider the least common multiplier of their denominators and multiply its square to both sides of the second constraint of the problem, then this constraint involves only polynomials of θ. In this case, the optimization variable x in P consists of y and the independent elements of X. $\square$

Example 2 (Polynomial optimization [23, 29])**.** Consider the maximization problem:

$$\begin{array}{ll} \text{maximize} & f(\theta) \\ \text{subject to} & \theta \in \Theta, \end{array}$$

where $f(\theta)$ is a scalar polynomial in θ. It is easy to see that the maximum value of this problem is equal to the minimum value of the problem:

$$\begin{array}{ll} \text{minimize} & x \\ \text{subject to} & x - f(\theta) \geq 0 \quad (\forall \theta \in \Theta), \end{array}$$

which is in the form of P. $\square$

The robust SDP problem P is difficult to solve. Hence, an approximate approach is taken usually: to replace the constraint $F(x, \theta) \succeq O$ $(\forall \theta \in \Theta)$ by a stronger constraint expressed by a usual parameter-independent LMI. The approximate problem constructed by this replacement is a usual SDP problem, which is solvable by a standard method. Since the feasible region of the approximate problem is a subset of the original feasible region, the minimum value of the approximate problem is greater than or equal to the minimum value of the original problem P. Recently, asymptotically exact approaches are proposed [28, 7, 31, 32], in which the minimum value of the approximate problem converges to the minimum value of P as the degree of an assumed polynomial increases. The SOS approach, in particular, can deal with a problem more general than P. That is, the parameter region Θ can be any set implicitly described by polynomial inequalities. In the succeeding sections, we investigate yet another asymptotically exact approach and show that it has good properties not possessed by the existing approaches. Although this approach cannot deal with a general parameter region like the SOS approach, many important problems are formulated in the form of P as is seen in the examples.

Before proceeding, we mention an approach, lying opposite to those discussed so far. They consider the problem:

$$\begin{aligned} P(\theta^{(1)},\dots,\theta^{(Q)}): \quad & \text{minimize} \quad c^{\mathrm{T}}x \\ & \text{subject to} \quad E(x) \succeq O, \quad F(x,\theta^{(q)}) \succeq O \quad (\forall q = 1,2,\dots,Q) \end{aligned}$$

with $\{\theta^{(q)}\}$ being points sampled from Θ. Note that the feasible region of this problem is a *superset* of that of the original problem P. Hence, the minimum value of this problem is *smaller* than or equal to that of P. Calafiore–Campi [9, 10] considered the case that the points $\{\theta^{(q)}\}$ are chosen randomly and quantified the discrepancy between the two feasible regions in a probabilistic sense. See [35] for related randomized techniques.

3. A Region-Dividing Approach

3.1. Preliminary results

We will present the region-dividing approach to the robust SDP problem P, which was proposed by Emoto–Oishi [19]. We begin by presenting preliminary results.

Proposition 3. *Let G be a real symmetric matrix. Let $[M \quad H]$ be a real nonsingular matrix of the same size as G satisfying $M^{\mathrm{T}}H = O$. The following three statements are equivalent:*

(a) $M^{\mathrm{T}}GM \succ O$;

(b) There exists a positive number a such that $G + aHH^{\mathrm{T}} \succ O$;

(c) There exists a real matrix W such that $G + HW^{\mathrm{T}} + WH^{\mathrm{T}} \succ O$.

Proof. From (b), the statement (c) follows with $W := (a/2)H$. Since M is of full-column rank, premultiplication of M^{T} and postmultiplication of M to the inequality in (c) gives the statement (a).

It remains to derive (b) from (a). Since $[M \;\; H]$ is nonsingular, it suffices to show the positive definiteness of

$$\begin{bmatrix} M^{\mathrm{T}} \\ H^{\mathrm{T}} \end{bmatrix} (G + aHH^{\mathrm{T}})[M \;\; H] = \begin{bmatrix} M^{\mathrm{T}}GM & M^{\mathrm{T}}GH \\ H^{\mathrm{T}}GM & H^{\mathrm{T}}GH \end{bmatrix} + a \begin{bmatrix} O \\ H^{\mathrm{T}}H \end{bmatrix} [O \;\; H^{\mathrm{T}}H].$$

Here, we used $M^{\mathrm{T}}H = O$. Since $M^{\mathrm{T}}GM \succ O$ by (a), it suffices to show the positive definiteness of its Schur complement

$$H^{\mathrm{T}}GH + a(H^{\mathrm{T}}H)^2 - H^{\mathrm{T}}GM(M^{\mathrm{T}}GM)^{-1}M^{\mathrm{T}}GH.$$

But this is true for a sufficiently large a due to the positive definiteness of $H^{\mathrm{T}}H$. $\square$

In the control community, the equivalence between (a) and (b) is known as Finsler's lemma while the equivalence between (b) and (c) as a special case of the elimination lemma.

Our approach to the robust SDP problem P is to express $2F(x, \theta) = M(\theta)^{\mathrm{T}}G(x)M(\theta)$ and to use the condition (c) with some $H(\theta)$ affine in θ. The matrix W has to be found for each θ. We construct an approximation by requiring W to be constant in some polytope and considering the inequality only at the vertices.

In order to define $G(x)$ and $M(\theta)$, we expand $F(x, \theta)$ as a polynomial in θ:

$$F(x, \theta) = F_{00\cdots 0}(x) + F_{10\cdots 0}(x)\theta_1 + \cdots + F_{d_1 d_2 \cdots d_p}(x)\theta_1^{d_1}\theta_2^{d_2}\cdots\theta_p^{d_p}.$$

Although the terms can be arranged in any order, we consider here an inversely lexicographic order, that is, the term $\theta_1^{\alpha_1}\theta_2^{\alpha_2}\cdots\theta_p^{\alpha_p}$ precedes the term $\theta_1^{\beta_1}\theta_2^{\beta_2}\cdots\theta_p^{\beta_p}$ if

$$\begin{cases} \alpha_p < \beta_p \quad \text{or} \\ \alpha_q < \beta_q, \alpha_{q+1} = \beta_{q+1}, \ldots, \alpha_p = \beta_p \quad \text{for some } q = 1, 2, \ldots, p-1. \end{cases}$$

With this order, we define

$$\begin{aligned} F_*(x) &= [F_{10\cdots 0}(x) \;\; F_{20\cdots 0}(x) \;\; \cdots \;\; F_{d_1 d_2 \cdots d_p}(x)], \\ M(\theta) &= [I_m \;\; \theta_1 I_m \;\; \theta_2 I_m \;\; \cdots \;\; \theta_1^{d_1}\theta_2^{d_2}\cdots\theta_p^{d_p} I_m]^{\mathrm{T}} \end{aligned}$$

so that $F(x, \theta) = [F_{00\cdots 0}(x) \;\; F_*(x)]M(\theta)$. These matrices look as in (1) and (2) below in the case of $p = 2$ and $d_1 = d_2 = 2$. Define also

$$G(x) := \begin{bmatrix} 2F_{00\cdots 0}(x) & F_*(x) \\ F_*(x)^{\mathrm{T}} & O \end{bmatrix}.$$

It is easy to see that $M(\theta)^{\mathrm{T}}G(x)M(\theta) = 2F(x,\theta)$. With $D = \prod_{i=1}^{p}(d_i+1)$, the matrix $M(\theta)$ is $Dm \times m$ while $G(x)$ is $Dm \times Dm$.

We next define the matrix $H(\theta)$ such that $[M(\theta) \;\; H(\theta)]$ is nonsingular, $M(\theta)^{\mathrm{T}}H(\theta) = O$ holds, and $H(\theta)$ is affine in θ. In the case of $p=2$ and $d_1 = d_2 = 2$, this is possible with

$$F_*(x) = \left[\begin{array}{cc|ccc|ccc} F_{10}(x) & F_{20}(x) & F_{01}(x) & F_{11}(x) & F_{21}(x) & F_{02}(x) & F_{12}(x) & F_{22}(x) \end{array}\right], \tag{1}$$

$$M(\theta) = \left[\begin{array}{c} I_m \\ \theta_1 I_m \\ \theta_1^2 I_m \\ \hline \theta_2 I_m \\ \theta_1\theta_2 I_m \\ \theta_1^2\theta_2 I_m \\ \hline \theta_2^2 I_m \\ \theta_1\theta_2^2 I_m \\ \theta_1^2\theta_2^2 I_m \end{array}\right], \quad H(\theta) = \left[\begin{array}{cc|ccc|ccc} -\theta_1 I_m & & -\theta_2 I_m & & & & & \\ I_m & -\theta_1 I_m & & -\theta_2 I_m & & & & \\ & I_m & & & -\theta_2 I_m & & & \\ \hline & & I_m & & & -\theta_2 I_m & & \\ & & & I_m & & & -\theta_2 I_m & \\ & & & & I_m & & & -\theta_2 I_m \\ \hline & & & & & I_m & & \\ & & & & & & I_m & \\ & & & & & & & I_m \end{array}\right]. \tag{2}$$

For a formal definition of $H(\theta)$, we introduce $q \times q$ matrices J_q and K_q defined as

$$(J_q)_{ij} := \begin{cases} 1 & \text{if } i+1=j, \\ 0 & \text{otherwise;} \end{cases} \qquad (K_q)_{ij} := \begin{cases} 1 & \text{if } i=j=1, \\ 0 & \text{otherwise,} \end{cases}$$

and set

$$\begin{aligned} H(\theta) := \big(& -\theta_1 K_{d_p+1} \otimes \cdots \otimes K_{d_3+1} \otimes K_{d_2+1} \otimes J_{d_1+1} \otimes I_m \\ & -\theta_2 K_{d_p+1} \otimes \cdots \otimes K_{d_3+1} \otimes J_{d_2+1} \otimes I_{d_1+1} \otimes I_m - \cdots \\ & -\theta_p J_{d_p+1} \otimes I_{d_{p-1}+1} \otimes \cdots \otimes I_{d_1+1} \otimes I_m + I_{Dm}\big) \times \begin{bmatrix} O_{m\times(Dm-m)} \\ I_{Dm-m} \end{bmatrix}, \end{aligned}$$

which is a $Dm \times (Dm-m)$ matrix.

We first fix x and θ and consider the following condition: There exists $W = W(x,\theta)$ such that $G(x) + H(\theta)W^{\mathrm{T}} + WH(\theta)^{\mathrm{T}} \succeq O$. This condition implies $F(x,\theta) \succeq O$ by premultiplication of $M(\theta)^{\mathrm{T}}$ and postmultiplication of $M(\theta)$. The converse implication is not true in general. By Proposition 3, however, it is true if the inequality is replaced by a strict one. In this sense, the considered condition is almost equivalent to $F(x,\theta) \succeq O$.

Next, we consider this condition for varying θ. Although we can choose W for each θ, we impose a stronger condition that W be constant in a polytope in Θ. On one hand, this restricts the choice of W. On the other hand, this restriction enables us to conclude $F(x,\theta) \succeq O$ for all θ in that polytope by checking $G(x) + H(\theta)W^{\mathrm{T}} + WH(\theta)^{\mathrm{T}} \succeq O$ only at the vertices of the polytope. This is a consequence of affinity of $H(\theta)$. This observation is formally stated below.

Proposition 4. *Let x be any point in $\mathbb{R}^n$ and $\theta^{(1)}, \ldots, \theta^{(Q)}$ be any points in Θ. Then, $F(x, \theta) \succeq O$ holds for all θ in the convex hull of $\{\theta^{(1)}, \ldots, \theta^{(Q)}\}$, if there exists W satisfying*

$$G(x) + H(\theta^{(q)})W^{\mathrm{T}} + WH(\theta^{(q)})^{\mathrm{T}} \succeq O \tag{3}$$

for $q = 1, 2, \ldots, Q$.

Proof. Consider θ in the convex hull of $\{\theta^{(1)}, \ldots, \theta^{(Q)}\}$ and represent it as a convex combination. Convex combination of the Q inequalities (3) with the same coefficients gives $G(x) + H(\theta)W^{\mathrm{T}} + WH(\theta)^{\mathrm{T}} \succeq O$ for the considered θ. This implies $F(x, \theta) \succeq O$. $\square$

This proposition gives an idea of the region-dividing approach to be presented. That is, we divide the parameter region Θ to convex polytopes and consider the condition (3) for each polytope. Since the condition (3) is affine in x and W, minimization of $c^{\mathrm{T}}x$ with this condition is tractable as a standard SDP problem. This problem is an approximation of the original problem because the considered condition is only sufficient for $F(x, \theta) \succeq O$ ($\forall \theta \in \Theta$). The quality of the approximation is, however, improved as the resolution of the division becomes higher.

We present the approach more formally in the next subsection. We conclude this subsection by preparing terminology.

A *division* of Θ is a finite family $\{\Theta^{[j]}\}_{j=1}^{J}$ such that $\Theta = \bigcup_{j=1}^{J} \Theta^{[j]}$ holds and $\Theta^{[j]} \cap \Theta^{[k]}$ has no interior point for $j \neq k$. Each element $\Theta^{[j]}$ of a division $\{\Theta^{[j]}\}$ is called a *subregion*. We require each subregion to be a convex polytope with a nonempty interior. Since Θ is a polytope with a nonempty interior, such a division always exists. The *radius* of a subregion $\Theta^{[j]}$ is defined as $\operatorname{rad} \Theta^{[j]} := \min_{\theta \in \Theta^{[j]}} \max_{\theta' \in \Theta^{[j]}} \max_i |\theta_i - \theta'_i|$. A point θ that attains the minimum is called a *center* of $\Theta^{[j]}$. For example, when $\Theta^{[j]}$ is a multi-dimensional interval $\prod_{i=1}^{p} [\underline{\theta}_i, \overline{\theta}_i]$, its radius is $\max_i (\overline{\theta}_i - \underline{\theta}_i)/2$ and its center is $[(\underline{\theta}_1 + \overline{\theta}_1)/2 \quad (\underline{\theta}_2 + \overline{\theta}_2)/2 \quad \cdots \quad (\underline{\theta}_p + \overline{\theta}_p)/2]^{\mathrm{T}}$. The *maximum radius* of a division $\{\Theta^{[j]}\}$ is defined as $\overline{\operatorname{rad}}\, \{\Theta^{[j]}\} := \max_j \operatorname{rad} \Theta^{[j]}$. Finally, a division $\{\widetilde{\Theta}^{[k]}\}$ is called a *subdivision* of another division $\{\Theta^{[j]}\}$, if any subregion $\Theta^{[j]}$ is expressed as a union $\bigcup_k \widetilde{\Theta}^{[k]}$ of some subfamily of $\{\widetilde{\Theta}^{[k]}\}$.

3.2. A region-dividing approach

For a given robust SDP problem P, we consider a division $\{\Theta^{[j]}\}_{j=1}^{J}$ of the parameter region Θ. For each subregion $\Theta^{[j]}$, we consider the condition (3) with $\theta^{(1)}, \ldots, \theta^{(Q)}$ being the vertices of $\Theta^{[j]}$ and write the set of all (x, W) satisfying this condition as $S(\Theta^{[j]})$. We allow the matrix W to have different values in different subregions. Hence, we have the following approximate problem:

$$
\begin{aligned}
P(\{\Theta^{[j]}\}): \quad & \text{minimize} \quad c^{\mathrm{T}} x \\
& \text{subject to} \quad E(x) \succeq O, \quad (x, W^{[j]}) \in S(\Theta^{[j]}) \quad (\forall j = 1, 2, \ldots, J).
\end{aligned}
$$

The feasible region of $P(\{\Theta^{[j]}\})$ is in the space of $(x, \{W^{[j]}\}_{j=1}^{J})$. We consider its projection to the space of x and call it the *projected feasible region*. By Proposition 4, the projected feasible region of $P(\{\Theta^{[j]}\})$ is included in the feasible region of P. This implies $\min P \leq \min P(\{\Theta^{[j]}\})$.

Subdivision of $\{\Theta^{[j]}\}$ improves the quality of approximation. Indeed, if $\{\widetilde{\Theta}^{[k]}\}$ is a subdivision of $\{\Theta^{[j]}\}$, the projected feasible region of $P(\{\widetilde{\Theta}^{[k]}\})$ includes that of $P(\{\Theta^{[j]}\})$, which implies $\min P \leq \min P(\{\widetilde{\Theta}^{[k]}\}) \leq \min P(\{\Theta^{[j]}\})$. Thus, we obtain a sequence of approximations for the problem P. Qualitative reasoning shows that the approximation becomes exact as the maximum radius $\overline{\mathrm{rad}}\,\{\Theta^{[j]}\}$ goes to zero [19]. However, in this paper, we would like to make more quantitative discussion. We obtain in the next section an upper bound on the approximation error $\min P(\{\Theta^{[j]}\}) - \min P$ as a function of the maximum radius.

We would like to mention here the relationship to the branch-and-bound approach for optimization. When the region-dividing approach is used for polynomial optimization, it looks similar to the branch-and-bound approach. A difference is that the present approach can be used for a more general class of robust SDP problems. Moreover, an upper bound on the approximation error is not available in the branch-and-bound approach in general.

4. An Upper Bound on the Approximation Error

4.1. An upper bound

This is the main section of this paper. It gives an upper bound on the approximation error $\min P(\{\Theta^{[j]}\}) - \min P$ for a given division $\{\Theta^{[j]}\}$.

Our upper bound is *a priori* in the sense that one can obtain it *before* solving $P(\{\Theta^{[j]}\})$. We would like to mention that an *a posteriori* upper bound is easily obtained. Indeed, the problem $P(\theta^{(1)}, \ldots, \theta^{(Q)})$ in Section 2 satisfies $\min P(\theta^{(1)}, \ldots, \theta^{(Q)}) \leq \min P \leq \min P(\{\Theta^{[j]}\})$. The discrepancy $\min P(\{\Theta^{[j]}\}) - \min P(\theta^{(1)}, \ldots, \theta^{(Q)})$ is computable and forms an upper bound on $\min P(\{\Theta^{[j]}\}) - \min P$.

We use the following auxiliary problem to derive an upper bound:

$$
\begin{aligned}
P_{\epsilon}: \quad & \text{minimize} \quad c^{\mathrm{T}} x \\
& \text{subject to} \quad E(x) \succeq O, \quad F(x, \theta) \succeq \epsilon I \quad (\forall \theta \in \Theta),
\end{aligned}
$$

where ϵ is a nonnegative number. Obviously, the feasible region of P_ϵ is included in that of P and, thus, $\min P \le \min P_\epsilon$. We choose ϵ so that

$$\min P \le \min P(\{\Theta^{[j]}\}) \le \min P_\epsilon. \tag{4}$$

We evaluate $\min P_\epsilon - \min P$ using convexity of $\min P_\epsilon$ and obtain an upper bound on $\min P(\{\Theta^{[j]}\}) - \min P$.

For precise discussion, we make the following assumption in the rest of this paper.

Assumption 5.

(a) The robust SDP problem P is strictly feasible, that is, there exists $x \in \mathbb{R}^n$ such that

$$E(x) \succ O, \quad F(x, \theta) \succ O \quad (\forall \theta \in \Theta).$$

(b) For any v, the level set

$$\left\{ x \in \mathbb{R}^n \mid c^{\mathrm{T}} x \le v, \quad E(x) \succeq O, \quad F(x, \theta) \succeq O \quad (\forall \theta \in \Theta) \right\}$$

is bounded. □

We first focus on one subregion $\Theta^{[j]}$ in a division $\{\Theta^{[j]}\}$. Aiming at the inequality (4), we will choose ϵ so that $F(x, \theta^{\mathrm{c}}) \succeq \epsilon I$ implies the existence of W such that $(x, W) \in S(\Theta^{[j]})$, where θ^{c} is a center of this subregion.

It is convenient to expand $F(x, \theta)$ around θ^{c}. With $\theta^{\mathrm{c}} = [\theta_1^{\mathrm{c}} \ \theta_2^{\mathrm{c}} \ \cdots \ \theta_p^{\mathrm{c}}]^{\mathrm{T}}$, let us write

$$F(x, \theta) = F_{00\cdots0}^{\theta^{\mathrm{c}}}(x) + F_{10\cdots0}^{\theta^{\mathrm{c}}}(x)(\theta_1 - \theta_1^{\mathrm{c}}) + \cdots + F_{d_1 d_2 \cdots d_p}^{\theta^{\mathrm{c}}}(x)(\theta_1 - \theta_1^{\mathrm{c}})^{d_1}(\theta_2 - \theta_2^{\mathrm{c}})^{d_2} \cdots (\theta_p - \theta_p^{\mathrm{c}})^{d_p}.$$

Arranging the terms in the inversely lexicographic order, define

$$\begin{aligned} F_*^{\theta^{\mathrm{c}}}(x) &= [F_{10\cdots0}^{\theta^{\mathrm{c}}}(x) \ \ F_{20\cdots0}^{\theta^{\mathrm{c}}}(x) \ \ \cdots \ \ F_{d_1 d_2 \cdots d_p}^{\theta^{\mathrm{c}}}(x)], \\ G^{\theta^{\mathrm{c}}}(x) &= \begin{bmatrix} 2F_{00\cdots0}^{\theta^{\mathrm{c}}}(x) & F_*^{\theta^{\mathrm{c}}}(x) \\ F_*^{\theta^{\mathrm{c}}}(x)^{\mathrm{T}} & O \end{bmatrix}. \end{aligned} \tag{5}$$

With this $G^{\theta^{\mathrm{c}}}(x)$, we consider the inequality

$$G^{\theta^{\mathrm{c}}}(x) + H(\theta^{(q)} - \theta^{\mathrm{c}})(W^{\theta^{\mathrm{c}}})^{\mathrm{T}} + W^{\theta^{\mathrm{c}}} H(\theta^{(q)} - \theta^{\mathrm{c}})^{\mathrm{T}} \succeq O \tag{6}$$

for each vertex $\theta^{(q)}$ of the subregion $\Theta^{[j]}$. This forms a counterpart of the inequality (3) around $\theta = \theta^{\mathrm{c}}$. Let us write as $S^{\theta^{\mathrm{c}}}(\Theta^{[j]})$ the set of all $(x, W^{\theta^{\mathrm{c}}})$ satisfying the inequality (6) for all vertices of $\Theta^{[j]}$. Then, we see in the next lemma that the condition with $S^{\theta^{\mathrm{c}}}(\Theta^{[j]})$ is equivalent to the previous condition with $S(\Theta^{[j]})$.

Lemma 6. *Let x be any point in $\mathbb{R}^n$. There exists a matrix W satisfying $(x, W) \in S(\Theta^{[j]})$ if and only if there exists a matrix W^{θ^c} satisfying $(x, W^{\theta^c}) \in S^{\theta^c}(\Theta^{[j]})$.*

Proof. We prove the "if" part only. The proof of the "only if" part is similar.

By the definition of $M(\theta)$, there exists a constant matrix T such that $M(\theta - \theta^c) = TM(\theta)$ is an identity in θ. Since we have $M(\theta)^{\mathrm{T}} T^{\mathrm{T}} H(\theta - \theta^c) = M(\theta - \theta^c)^{\mathrm{T}} H(\theta - \theta^c) = O$ and the columns of $H(\theta)$ constitute the basis of the space $\{h \in \mathbb{R}^{Dm} \mid M(\theta)^{\mathrm{T}} h = 0\}$, there has to exist $\widetilde{T}$ such that $T^{\mathrm{T}} H(\theta - \theta^c) = H(\theta)\widetilde{T}$. This $\widetilde{T}$ does not depend on θ because the left-hand side matrix is of order one in θ.

Now, premultiply T^{T} and postmultiply T to both sides of (6). Then, $T^{\mathrm{T}} G^{\theta^c}(x) T = G(x)$ because $F(x, \theta) = [F^{\theta^c}_{00\cdots 0}(x) \ \ F^{\theta^c}_{*}(x)] M(\theta - \theta^c) = [F^{\theta^c}_{00\cdots 0}(x) \ \ F^{\theta^c}_{*}(x)] TM(\theta)$. Moreover, we have $T^{\mathrm{T}} H(\theta^{(q)} - \theta^c)(W^{\theta^c})^{\mathrm{T}} T = H(\theta^{(q)})\widetilde{T}(W^{\theta^c})^{\mathrm{T}} T$. Hence, $W^{\mathrm{T}} := \widetilde{T}(W^{\theta^c})^{\mathrm{T}} T$ gives the desired matrix. $\square$

The following lemma gives ϵ such that the desired implication holds.

Lemma 7. *Let x be any point in $\mathbb{R}^n$ and suppose* $\operatorname{rad} \Theta^{[j]} \leq 1/(p+1)$. *If* $\epsilon \geq \big(\{\overline{\sigma}[F^{\theta^c}_{*}(x)] + \sqrt{pm}\}^2/2\big) \operatorname{rad} \Theta^{[j]}$, *then the inequality $F(x, \theta^c) \succeq \epsilon I$ implies the existence of W satisfying $(x, W) \in S(\Theta^{[j]})$.*

Proof. We will show that the choice $W^{\theta^c} = (1/\operatorname{rad} \Theta^{[j]})[O_{(Dm-m)\times m} \quad I_{Dm-m}]^{\mathrm{T}}$ satisfies $(x, W^{\theta^c}) \in S^{\theta^c}(\Theta^{[j]})$. Then, application of Lemma 6 gives Lemma 7.

To show $(x, W^{\theta^c}) \in S^{\theta^c}(\Theta^{[j]})$, we pick a vertex $\theta^{(q)}$ of $\Theta^{[j]}$ and derive the inequality (6). The second term on the left-hand side, $H(\theta^{(q)} - \theta^c)(W^{\theta^c})^{\mathrm{T}}$, has the structure

$$\begin{bmatrix} O_{m\times m} & H_1 \\ O_{(Dm-m)\times m} & H_2 \end{bmatrix}.$$

By the construction of W^{θ^c} and $H(\theta^{(q)} - \theta^c)$ (see (2)), each row of H_1 contains at most p nonzero elements, each of which has the form $-(\theta^{(q)}_i - \theta^c_i)/\operatorname{rad} \Theta^{[j]}$, $i = 1, 2, \ldots, p$. Note that the magnitude of these nonzero elements is less than or equal to unity. We have, hence, $\overline{\sigma}(H_1) \leq \sqrt{pm}$. On the other hand, the matrix H_2 has the diagonal elements all equal to $1/\operatorname{rad} \Theta^{[j]}$. Each of its rows has at most p nonzero non-diagonal elements whose magnitude is less than or equal to unity. Each of its columns has at most one nonzero non-diagonal element whose magnitude is again less than or equal to unity. These facts imply that

$$H_2 + H_2^{\mathrm{T}} \succeq \Big(\frac{2}{\operatorname{rad} \Theta^{[j]}} - p - 1\Big) I.$$

Now, the left-hand side of (6) has the upper-left $m \times m$ block equal to

$$2F^{\theta^{\rm c}}_{00\cdots 0}(x) = 2F(x, \theta^{\rm c}) \succeq 2\epsilon I.$$

Its Schur complement is

$$\begin{aligned} H_2 + H_2^{\rm T} - [F^{\theta^{\rm c}}_*(x) + H_1]^{\rm T}[2F^{\theta^{\rm c}}_{00\cdots 0}(x)]^{-1}[F^{\theta^{\rm c}}_*(x) + H_1] \\ \succeq \Big(\frac{2}{\operatorname{rad} \Theta^{[j]}} - p - 1\Big)I - \big\{\overline{\sigma}[F^{\theta^{\rm c}}_*(x)] + \sqrt{pm}\big\}^2 \frac{1}{2\epsilon} I. \end{aligned}$$

The right-hand side matrix is positive semidefinite by the assumptions of the lemma. Hence, we have (6) and complete the proof. □

We next apply this lemma to each subregion in $\{\Theta^{[j]}\}$ to obtain the desired inequality $\min P \le \min P(\{\Theta^{[j]}\}) \le \min P_\epsilon$. With Assumption 5, there exists $\epsilon_0 > 0$ such that, for any $0 \le \epsilon \le \epsilon_0$, the auxiliary problem P_ϵ is strictly feasible. Let v_0 be a number with $\min P_{\epsilon_0} \le v_0$ and define the level set

$$X = \big\{x \in \mathbb{R}^n \,|\, c^{\rm T}x \le v_0, \quad E(x) \succeq O, \quad F(x, \theta) \succeq O \quad (\forall \theta \in \Theta)\big\},$$

which is nonempty and bounded. Then, for each $0 \le \epsilon \le \epsilon_0$, the minimum of P_ϵ is attained in X and only in X. Let U be a number such that

$$\max_{x \in X} \max_{\theta \in \Theta} \overline{\sigma}[F^{\theta}_*(x)] \le U, \tag{7}$$

where $F^\theta_*(x)$ is defined from the expansion of $F(x, \theta)$ around θ as in (5).

Lemma 8. *Suppose that a given division* $\{\Theta^{[j]}\}$ *satisfies*

$$\overline{\operatorname{rad}}\,\{\Theta^{[j]}\} \le \min\Big\{\frac{2\epsilon_0}{(U + \sqrt{pm})^2}, \quad \frac{1}{p+1}\Big\}.$$

Then, we have $\min P \le \min P(\{\Theta^{[j]}\}) \le \min P_\epsilon$ *for* $\epsilon = [(U + \sqrt{pm})^2/2]\overline{\operatorname{rad}}\,\{\Theta^{[j]}\}$.

Proof. It suffices to show $\min P(\{\Theta^{[j]}\}) \le \min P_\epsilon$. By the assumption, $0 \le \epsilon \le \epsilon_0$ holds and, thus, any minimizer of P_ϵ belongs to X. Let x be such a minimizer. Then, due to (7), Lemma 7 can be applied to each subregion in $\{\Theta^{[j]}\}$. Hence, x belongs to the projected feasible region of $P(\{\Theta^{[j]}\})$, which gives $\min P(\{\Theta^{[j]}\}) \le \min P_\epsilon$. □

We now take the final step. By Lemma 8, the discrepancy $\min P_\epsilon - \min P$ bounds from above the approximation error $\min P(\{\Theta^{[j]}\}) - \min P$. Note that $\min P_\epsilon$ is a convex function in ϵ. Using this fact, we evaluate $\min P_\epsilon - \min P$.

Theorem 9. *Suppose that Assumption 5 holds. Let ϵ_0 be a number such that the auxiliary problem P_ϵ is strictly feasible for any $0 \leq \epsilon \leq \epsilon_0$. Let U be a number in (7). For some $0 < \epsilon_1 < \epsilon_0$, let g be an upper bound on the left derivative of* $\min P_\epsilon$ *at $\epsilon = \epsilon_1$. Finally, suppose that a given division $\{\Theta^{[j]}\}$ satisfies*

$$\overline{\mathrm{rad}}\,\{\Theta^{[j]}\} \leq \min\left\{\frac{2\epsilon_1}{(U+\sqrt{pm})^2}, \quad \frac{1}{p+1}\right\}.$$

Then, we have

$$\min P(\{\Theta^{[j]}\}) - \min P \leq \frac{g(U+\sqrt{pm})^2}{2}\overline{\mathrm{rad}}\,\{\Theta^{[j]}\}. \tag{8}$$

Proof. By $\epsilon_1 < \epsilon_0$, we can apply Lemma 8 to have $\min P \leq \min P(\{\Theta^{[j]}\}) \leq \min P_\epsilon$ for $\epsilon = [(U+\sqrt{pm})^2/2]\overline{\mathrm{rad}}\,\{\Theta^{[j]}\}$. This ϵ satisfies $0 \leq \epsilon \leq \epsilon_1$ by the assumption.

Due to convexity of $\min P_\epsilon$, the upper bound g is greater than or equal to the left derivative of $\min P_\epsilon$ at this ϵ. Hence, we have $\min P \geq \min P_\epsilon - g\epsilon$, which gives $\min P_\epsilon - \min P \leq g\epsilon$. Substitution of the concrete form of ϵ completes the proof. □

An immediate consequence of this result is the asymptotic exactness of the present approach, that is, the approximation error $\min P(\{\Theta^{[j]}\}) - \min P$ converges to zero as the maximum radius of the division goes to zero. Moreover, the rate of the convergence is at least the first order of the maximum radius. Computation of the right-hand side of (8) is possible as we will see in the next subsection. Note that the corresponding result has not been obtained for the existing asymptotically exact approaches [28, 7, 31, 32].

Theorem 9 claims that we need to decrease $\overline{\mathrm{rad}}\,\{\Theta^{[j]}\}$ in order to make the approximation error small. This is not easy when the parameter dimension p is high. Indeed, the number of LMIs in the approximate problem $P(\{\Theta^{[j]}\})$ is at least proportional to the number of subregions, which is in the order of $(\overline{\mathrm{rad}}\,\{\Theta^{[j]}\})^{-p}$. Hence, decrease of $\overline{\mathrm{rad}}\,\{\Theta^{[j]}\}$ leads to rapid increase of the size of the approximate problem. This difficulty may be inevitable because the original problem P is NP-hard. It would be nice, however, if we could suppress the rapid increase of the number of subregions. We will discuss this issue in Section 5.

4.2. Computation of the upper bound

The upper bound (8) is computable. For its computation, we need three numbers v_0, U, and g. We will discuss computation of each in order.

We begin by computation of v_0. Let ϵ_0 be a number such that the auxiliary problem P_ϵ is strictly feasible for $0 \leq \epsilon \leq \epsilon_0$. Since P_{ϵ_0} is itself a robust SDP problem, we can construct an approximate problem for it. The minimum value of the constructed problem is larger than or equal to $\min P_{\epsilon_0}$. Hence, this number can be used as v_0.

Next, we compute an upper bound U defined in (7). For this purpose, we first obtain an n-dimensional interval $\prod_{i=1}^n [\underline{x}_i, \overline{x}_i]$ that contains the level set X. In order to compute $\underline{x}_1$, for example, we consider the problem:

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & c^{\mathrm{T}} x \leq v_0, \quad E(x) \succeq O, \quad F(x, \theta) \succeq O \quad (\forall \theta \in \Theta). \end{array}$$

A lower bound on the minimum value is easily computed by sampling θ's as in $P(\theta^{(1)}, \dots, \theta^{(Q)})$. Other ends of the intervals can be computed in a similar way. After obtaining this n-dimensional interval, we consider the problem:

$$\begin{array}{ll} \text{minimize} & y \\ \text{subject to} & \begin{bmatrix} yI & F_*^\theta(x) \\ F_*^\theta(x)^{\mathrm{T}} & yI \end{bmatrix} \succeq O \quad \left(\forall x \in \prod_{i=1}^n [\underline{x}_i, \overline{x}_i], \quad \forall \theta \in \Theta\right), \end{array}$$

with the optimization variable y. This problem is in the form of the robust SDP problem P and, thus, an upper bound on its minimum value can be computed by construction of an approximation. The computed upper bound can be used as U.

Finally, we discuss computation of g, which is an upper bound on the left derivative of $\min P_\epsilon$ at $\epsilon = \epsilon_1$. It is easy to compute a lower bound v_1 on $\min P_{\epsilon_1}$ again by sampling θ's. On the other hand, v_0 is an upper bound on $\min P_{\epsilon_0}$. With these numbers, we can put $g := (v_0 - v_1)/(\epsilon_0 - \epsilon_1)$ because

$$g = \frac{v_0 - v_1}{\epsilon_0 - \epsilon_1} \geq \frac{\min P_{\epsilon_0} - \min P_{\epsilon_1}}{\epsilon_0 - \epsilon_1} \geq (\text{the left derivative at } \epsilon_1).$$

The last inequality follows from convexity of $\min P_\epsilon$.

5. Adaptive Division

We consider in this section how to make a good division that attains a small approximation error with a small number of subregions. To this end, we improve the error bound in Theorem 9.

For a given division $\{\Theta^{[j]}\}$, consider a minimizer $(x, \{W^{[j]}\})$ of the problem $P(\{\Theta^{[j]}\})$. With this minimizer, some constraints are active while others are not. Here, the constraint (3) is said to be *active* if the left-hand side matrix has a zero eigenvalue. If a subregion $\Theta^{[j]}$ has a vertex whose corresponding constraint is active, we say that this subregion is active. Since an active subregion is an important subregion, it is reasonable to define a new index for the resolution of the division:

$$\overline{\text{a-rad}}\, \{\Theta^{[j]}\} := \max_j \operatorname{rad} \Theta^{[j]},$$

where the maximum is taken over all j such that the subregion $\Theta^{[j]}$ becomes active at a minimum. When there are multiple minimizers, we define $\overline{\text{a-rad}}\,\{\Theta^{[j]}\}$ by taking the minimum of all possible values. In this case, $\overline{\text{a-rad}}\,\{\Theta^{[j]}\}$ is not easy to compute. However, its upper bound is always computable by solving $P(\{\Theta^{[j]}\})$. We call this index the *maximum active radius*. Obviously, $\overline{\text{a-rad}}\,\{\Theta^{[j]}\} \le \overline{\text{rad}}\,\{\Theta^{[j]}\}$. The next theorem says that the approximation error can be bounded with the maximum active radius.

Theorem 10. *With the same assumptions and symbols as in Theorem 9, we have*

$$\min P(\{\Theta^{[j]}\}) - \min P \le \frac{g(U+\sqrt{pm})^2}{2}\overline{\text{a-rad}}\,\{\Theta^{[j]}\}.$$

Proof. Let $\epsilon = [(U+\sqrt{pm})^2/2]\overline{\text{rad}}\,\{\Theta^{[j]}\}$. Since $0 \le \epsilon \le \epsilon_1 < \epsilon_0$, the problem P_ϵ has a strictly feasible solution $x \in \mathbb{R}^n$, with which $E(x) \succ O$ and $F(x,\theta) \succ \epsilon I$ $(\forall \theta \in \Theta)$. Using this x and constructing $W^{[j]}$ as in the proof of Lemma 7, we can have a strictly feasible solution of $P(\{\Theta^{[j]}\})$. Note also that $\min P(\{\Theta^{[j]}\})$ is bounded below. By these facts and the duality theorem on SDP (Theorem 2.4.1 of [4]), the dual problem of $P(\{\Theta^{[j]}\})$ has an optimal solution that attains $\min P(\{\Theta^{[j]}\})$. The optimal solution consists of positive semidefinite matrices $Y^{[j](q)}$ corresponding to all the subregions $\Theta^{[j]}$ and their vertices $\theta^{(q)}$. In particular, the matrices $Y^{[j](q)}$ are equal to zero for inactive subregions $\Theta^{[j]}$.

Now, let us subdivide each inactive subregion, if necessary, so that each of the created subregion has the radius smaller than or equal to $\overline{\text{a-rad}}\,\{\Theta^{[j]}\}$. Then, the resulting new division $\{\widetilde{\Theta}^{[k]}\}$ has $\overline{\text{rad}}\,\{\widetilde{\Theta}^{[k]}\} = \overline{\text{a-rad}}\,\{\Theta^{[j]}\}$. Note also that $\min P(\{\widetilde{\Theta}^{[k]}\}) \le \min P(\{\Theta^{[j]}\})$. We then consider the SDP dual of $P(\{\widetilde{\Theta}^{[k]}\})$ and construct its solution. In particular, if a subregion $\widetilde{\Theta}^{[k]}$ is a newly created subregion, assign zero matrices to the dual variables corresponding to this subregion; If a subregion $\widetilde{\Theta}^{[k]}$ coincides with one subregion in $\{\Theta^{[j]}\}$, assign the same values as before, *i.e.*, $Y^{[j](q)}$. With these assigned values, the dual objective function takes the same value as before, *i.e.*, $\min P(\{\Theta^{[j]}\})$. By weak duality, we have $\min P(\{\Theta^{[j]}\}) \le \min P(\{\widetilde{\Theta}^{[k]}\}) \le \min P(\{\Theta^{[j]}\})$, which shows

$$\begin{aligned}\min P(\{\Theta^{[j]}\}) - \min P &= \min P(\{\widetilde{\Theta}^{[k]}\}) - \min P \\ &\le \frac{g(U+\sqrt{pm})^2}{2}\overline{\text{rad}}\,\{\widetilde{\Theta}^{[k]}\} = \frac{g(U+\sqrt{pm})^2}{2}\overline{\text{a-rad}}\,\{\Theta^{[j]}\}.\end{aligned}$$

□

This theorem motivates us to consider the following procedure: Start with a coarse division; Repeat dividing an active subregion until the estimated application error becomes small enough. For estimation of the approximation error, both *a priori* and *a posteriori* bounds can be used though the latter is more recommendable for its tightness. Details of the procedure are given below. See Figure 1 for an example.

Algorithm 11.

0. Consider a coarse division $\{\Theta^{[j]}\}$.

1. Solve $P(\{\Theta^{[j]}\})$ for the current division.

2. Estimate the approximation error with the *a priori* or the *a posteriori* error bound and stop if it is small enough.

3. Find an active subregion of the maximum radius.

4. Divide that subregion into two subregions so that they have small radii. Go back to Step 1.

$\square$

Recall that subdivision never increases the approximate error. Hence, with this algorithm, the approximation error is monotonically non-increasing though the estimated value may increase.

There is no guarantee that the computational complexity is reduced by this algorithm. Indeed, in the worst case, the maximum active radius $\overline{\text{a-rad}}\,\{\Theta^{[j]}\}$ is equal to the maximum radius $\overline{\text{rad}}\,\{\Theta^{[j]}\}$. This means that the number of subregions is in the order of $(\overline{\text{a-rad}}\,\{\Theta^{[j]}\})^{-p}$. In many cases, however, more efficient division can be expected. See the examples in Section 6.

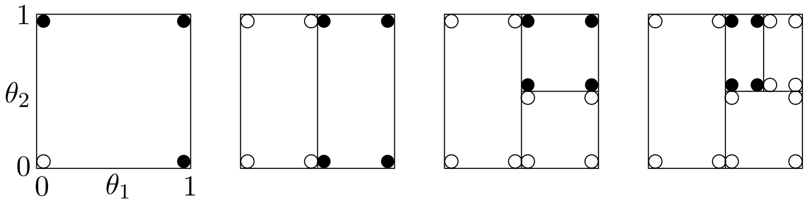
6. Examples

We applied our approach to two problems, both of which are polynomial optimization over a two-dimensional interval. See Example 2 in Section 2 for polynomial optimization. The result is satisfactory and shows some interesting aspects of the approach. We have $n = 1, p = 2, \ell = 0$, and $m = 1$ for the considered problems. All the computation was performed on YALMIP [24] with SDPA-M [20] as an SDP solver.

In the first example, we maximize

$$f(\theta) = -375\theta_1^4 + 800\theta_1^3 - 570\theta_1^2 + 144\theta_1 - 375\theta_2^4 + 800\theta_2^3 - 570\theta_2^2 + 144\theta_2$$

over the interval $\Theta = [0,1]^2$. This function has four local maxima, in particular, $f(0.2, 0.2) = 23.6$, $f(0.2, 0.8) = f(0.8, 0.2) = 18.2$, and $f(0.8, 0.8) = 12.8$. Hence, when we formulate this problem in the robust SDP problem P, its minimum $\min P$ is equal to 23.6.



minimum value:	1.090017	1.086318	1.082518	1.080000
a priori bound:	—	—	42.937136	42.937136
a posteriori bound:	0.010260	0.006560	0.002760	0.000242

Figure 1. Divisions generated for maximization of (9). A filled circle means that the corresponding constraint is active while a hollow circle inactive. Approximate minimum values as well as estimated approximation errors are also presented.

With the coarsest division $\{\Theta\}$, we constructed an approximation $P(\{\Theta\})$ and solved it. The obtained minimum value was 23.6, which is apparently equal to the exact value $\min P$. In fact, the proposed approach often gives an apparently exact result with a coarse division. This phenomenon is interesting because the SOS approach has a similar property.

Next, we maximize

$$f(\theta) = -5\theta_1^2\theta_2 - 5\theta_1\theta_2^2 + 9\theta_1\theta_2 \tag{9}$$

over $\Theta = [0,1]^2$. It has a unique local maximum $f(0.6, 0.6) = 1.08$ and, hence, the global maximum.

The corresponding robust SDP problem P has $\min P = 1.08$. With the coarsest division, our approximation gave $\min P(\{\Theta\}) = 1.090017$, which is larger than $\min P = 1.08$. In order to improve the approximation, we repeated subdivision with Algorithm 11. The result is summarized in Figure 1. In the figure, a filled circle means that the corresponding constraint is active while a hollow circle inactive. We can see that the division becomes finer around $\theta = [0.6 \;\; 0.6]^{\mathrm{T}}$, which maximizes $f(\theta)$. This is reasonable because accurate computation may be necessary around this point. We obtained a satisfactory approximation with the fourth division.

When $\overline{\mathrm{rad}}\,\{\Theta^{[j]}\} \leq 1/3$, the *a priori* error bound can be computed. The result is presented in Figure 1. We can see that the computed bound is much larger than the actual approximation error. On the other hand, the *a posteriori* error bound gives a tight estimation. It is computed from 1000 θ's randomly sampled in Θ. We should notice, however, that the *a posteriori* error bound can be computed only after the corresponding approximate problem is solved.

7. Computation of an Optimal Point in Polynomial Optimization

We saw in Sections 2 and 6 that the maximum value of a polynomial $f(\theta)$ over $\theta \in \Theta$ can be computed by the region-dividing approach. It was not clear, however, how to compute θ that attains the maximum value. We consider this issue in this section. The corresponding result is known in the SOS approach [23].

For a given polynomial optimization problem, divide Θ into $\{\Theta^{[j]}\}$ and consider the approximate problem $P(\{\Theta^{[j]}\})$. The function $F(x,\theta)$ is set as a scalar function $x - f(\theta)$ here while the objective function is x. Write as $(\widehat{x}, \{\widehat{W}^{[j]}\})$ an optimal solution of this approximate problem. Then, this $\widehat{x}$ is in general an upper bound on the maximum of $f(\theta)$, $\theta \in \Theta$. A necessary and sufficient condition for this upper bound to be exact is that $F(\widehat{x}, \widehat{\theta}) = 0$ holds for some $\widehat{\theta} \in \Theta$. In this case, this $\widehat{\theta}$ is a maximizer of $f(\theta)$. The next theorem gives a necessary and sufficient condition for $F(\widehat{x}, \widehat{\theta}) = 0$ to hold and, at the same time, gives a way to compute such $\widehat{\theta}$.

Theorem 12. *Let $\{\Theta^{[j]}\}$ be a division of Θ and $(\widehat{x}, \{\widehat{W}^{[j]}\})$ be an optimal solution of the problem $P(\{\Theta^{[j]}\})$. Let $\widehat{\theta}$ be a point in Θ and $\widehat{W}$ be the optimal solution $\widehat{W}^{[j]}$ for a subregion $\Theta^{[j]}$ containing $\widehat{\theta}$. Finally, let $\{\widehat{\theta}^{(1)}, \widehat{\theta}^{(2)}, \dots, \widehat{\theta}^{(Q')}\}$ be a subset of the vertices of the same $\Theta^{[j]}$ such that the relative interior of its convex hull contains $\widehat{\theta}$. Then, $F(\widehat{x}, \widehat{\theta}) = 0$ holds if and only if*

$$M(\widehat{\theta})^{\mathrm{T}}[G(\widehat{x}) + H(\widehat{\theta}^{(q)})\widehat{W}^{\mathrm{T}} + \widehat{W}H(\widehat{\theta}^{(q)})^{\mathrm{T}}]M(\widehat{\theta}) = 0 \tag{10}$$

holds for all $q = 1, 2, \dots, Q'$.

Proof. We show the "only if" part. Express $\widehat{\theta}$ as a convex combination of $\{\widehat{\theta}^{(1)}, \widehat{\theta}^{(2)}, \dots, \widehat{\theta}^{(Q')}\}$: $\widehat{\theta} = a^{(1)}\widehat{\theta}^{(1)} + a^{(2)}\widehat{\theta}^{(2)} + \cdots + a^{(Q')}\widehat{\theta}^{(Q')}$. Since $\widehat{\theta}$ is in the relative interior of the convex hull, the coefficients $a^{(q)}$ can be chosen all positive. Here, we have

$$0 = 2F(\widehat{x}, \widehat{\theta}) = \sum_{q=1}^{Q'} a^{(q)} M(\widehat{\theta})^{\mathrm{T}}[G(\widehat{x}) + H(\widehat{\theta}^{(q)})\widehat{W}^{\mathrm{T}} + \widehat{W}H(\widehat{\theta}^{(q)})^{\mathrm{T}}]M(\widehat{\theta})$$

due to the affinity of $H(\theta)$. Since all the terms are nonnegative, they have to be equal to zero. Division by $a^{(q)}$ shows the claim. The "if" part follows from the reversed reasoning. □

Based on Theorem 12, a maximizer $\widehat{\theta}$ can be computed as follows: Solve the approximate problem $P(\{\Theta^{[j]}\})$; Compute the kernel for each active constraint; Find $\widehat{\theta}$ such that the vector $M(\widehat{\theta})$ belongs to that kernel. Note that this theorem can be used also for verification of exactness of the approximation.

The condition (10) in Theorem 12 is nonlinear in $\widehat{\theta}$. This can be an obstacle in finding $\widehat{\theta}$. We can obtain a condition affine in $\widehat{\theta}$ by following the approach of Scherer [31], who discussed verification of exactness in a general setting.

Corollary 13. *Let the symbols be the same as in Theorem 12. Moreover, let $\widehat{Y}^{(q)}$ be a dual optimal solution of $P(\{\Theta^{[j]}\})$ corresponding to the primal constraint at $\widehat{\theta}^{(q)}$ in the subregion $\Theta^{[j]}$,* i.e., *$G(x)+H(\widehat{\theta}^{(q)})(W^{[j]})^{\mathrm{T}}+W^{[j]}H(\widehat{\theta}^{(q)})^{\mathrm{T}} \succeq O$. Then, $F(\widehat{x},\widehat{\theta})=0$ holds if $H(\widehat{\theta})^{\mathrm{T}}\widehat{Y}^{(q)}=O$ and $\widehat{Y}^{(q)} \neq O$ for all $q=1,2,\dots,Q'$.*

Proof. Recall that the kernel of $H(\widehat{\theta})^{\mathrm{T}}$ is spanned by the column vectors of $M(\widehat{\theta})$. Since $H(\widehat{\theta})^{\mathrm{T}}\widehat{Y}^{(q)}=O$ holds and $M(\widehat{\theta})$ is a column vector in the present case, each column of $\widehat{Y}^{(q)}$ is equal to $M(\widehat{\theta})$ multiplied by some real number. Here, $\widehat{Y}^{(q)}$ is a nonzero positive semidefinite matrix. Hence, we can write $\widehat{Y}^{(q)}=b^{(q)}M(\widehat{\theta})M(\widehat{\theta})^{\mathrm{T}}$ with $b^{(q)}>0$ for each $q=1,2,\dots,Q'$. Since $\widehat{Y}^{(q)}$ is a dual optimal solution, it satisfies the complementarity condition

$$\mathrm{tr}\,[G(\widehat{x})+H(\widehat{\theta}^{(q)})\widehat{W}^{\mathrm{T}}+\widehat{W}H(\widehat{\theta}^{(q)})^{\mathrm{T}}]\widehat{Y}^{(q)}=0$$

for $q=1,2,\dots,Q'$. Substitution of $\widehat{Y}^{(q)}=b^{(q)}M(\widehat{\theta})M(\widehat{\theta})^{\mathrm{T}}$ and Division by $b^{(q)}>0$ give the condition (10) in Theorem 12. $\square$

The condition in the corollary is not only affine in $\widehat{\theta}$ but also representable as an LMI. Hence, we can find $\widehat{\theta}$ by solving a standard SDP problem. This condition is not necessary for the exactness because a dual optimal solution is not unique and may not satisfy $\widehat{Y}^{(q)} \neq O$. A remedy is to use the primal-dual path-following interior-point method to solve $P(\{\Theta^{[j]}\})$. Since this method gives an optimal solution as large as possible [21], $\widehat{Y}^{(q)} \neq O$ can be expected.

We applied Theorem 12 to the two polynomial optimization problems in the previous section. We successfully obtained the optimal points $(0.2,\ 0.2)$ and $(0.6,\ 0.6)$, respectively.

8. Conclusion

An approximate approach to robust SDP is presented. The approximation error can be reduced to any degree by subdivision of the parameter region. An *a priori* upper bound on the approximation error is available, which is proportional to the maximum radius of the division. This error bound leads to an algorithm for an efficient division, which attains a small approximation error with low computational complexity. Numerical examples show that the proposed approach is promising. In particular, the exact minimum is often found with a coarse division.

In the context of the SOS approach, some techniques for efficient computation have been proposed [36, 30]. These techniques are quite different from the technique proposed in this paper. It may be interesting to consider their combination.

Acknowledgments

The author is particularly thankful to Kazuo Murota for a number of helpful comments. His thanks go to Yoshio Ebihara, Shinji Hara, Izumi Masubuchi, and Takashi Tsuchiya for their valuable comments. Part of this work is supported by the 21st Century COE Program on Information Science and Technology Strategic Core and a Grant-in-Aid of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] P. Apkarian, H. D. Tuan, and J. Bernussou, "Continuous-time analysis, eigenstructure assignment, and H^2 synthesis with enhanced linear matrix inequalities (LMI) characterizations," *IEEE Transactions on Automatic Control*, vol. 46, no. 12, pp. 1941–1946, 2001.

[2] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, "Robustness," in *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds. Boston, USA: Kluwer, 2000, pp. 139–162.

[3] A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Mathematics of Operations Research*, vol. 23, no. 4, pp. 769–805, 1998.

[4] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, USA: SIAM, 2001.

[5] A. Ben-Tal and A. Nemirovski, "On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty," *SIAM Journal on Optimization*, vol. 12, no. 3, pp. 811–833, 2002.

[6] A. Ben-Tal and A. Nemirovski, "Robust optimization: methodology and applications," *Mathematical Programming*, ser. B, vol. 92, pp. 453–480, 2002.

[7] P.-A. Bliman, "On robust semidefinite programming," in *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems (MTNS2004)*, Leuven, Belgium, July 2004.

[8] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, USA: SIAM, 1994.

[9] G. Calafiore and M. C. Campi, "A new bound on the generalization rate of sampled convex programs," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, Paradise Island, The Bahamas, December 2004, pp. 5328–5333.

[10] G. Calafiore and M. C. Campi, "Uncertain convex programs: randomized solutions and confidence levels," *Mathematical Programming*, vol. 102, no. 1, pp. 25–46, 2005.

[11] Y.-Y. Cao and Z. Lin, "A descriptor system approach to robust stability analysis and controller synthesis," *IEEE Transactions on Automatic Control*, vol. 49, no. 11, pp. 2081–2084, 2004.

[12] M. C. de Oliveira, J. Bernussou, and J. C. Geromel, "A new discrete-time robust stability condition," *Systems & Control Letters*, vol. 37, pp. 261–265, 1999.

[13] M. C. de Oliveira, J. C. Geromel, and J. Bernussou, “Extended H^2 and H^∞ norm characterizations and controller parametrizations for discrete-time systems,” *International Journal of Control*, vol. 75, no. 9, pp. 666–679, 2002.

[14] M. Dettori and C. W. Scherer, “New robust stability and performance conditions based on parameter dependent multiplies,” in *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia, December 2000, pp. 4187–4192.

[15] Y. Ebihara and T. Hagiwara, “New dilated LMI characterizations for continuous-time multiobjective controller synthesis,” *Automatica*, vol. 40, no. 11, pp. 2003–2009, 2004.

[16] Y. Ebihara and T. Hagiwara, “A dilated LMI approach to robust performance analysis of linear time-invariant uncertain systems,” *Automatica*, vol. 41, no. 11, pp. 1933–1941, 2005.

[17] L. El Ghaoui and S.-I. Niculescu, “Robust decision problems in engineering: a linear matrix inequality approach,” in *Advances in Linear Matrix Inequality Methods in Control*, L. El Ghaoui and S.-I. Niculescu, Eds. Philadelphia, USA: SIAM, 2000, pp. 3–37.

[18] L. El Ghaoui, F. Oustry, and H. Lebret, “Robust solutions to uncertain semidefinite programs,” *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 33–52, 1998.

[19] K. Emoto and Y. Oishi, “Analysis and synthesis of control systems rationally dependent on uncertain parameters” (in Japanese), *Transactions of the Society of Instrument and Control Engineers*, vol. 41, no. 4, pp. 314–321, 2005.

[20] K. Fujisawa, Y. Futakata, M. Kojima, S. Matsuyama, S. Nakamura, K. Nakata, and M. Yamashita, “SDPA-M (SemiDefinite Programming Algorithm in MATLAB) user’s manual: version 6.2.0,” Technical Report B-359, revised version, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, May 2005, available from `http://grid.r.dendai.ac.jp/sdpa/`

[21] D. Goldfarb and K. Scheinberg, “Interior-point trajectories in semidefinite programming,” *SIAM Journal on Optimization*, vol. 8, no. 4, pp. 871–886, 1998.

[22] M. Kawata, “A multiobjective gain scheduling using descriptor representation: an approach from LMI non-common solutions” (in Japanese), *Transactions of the Society of Instrument and Control Engineers*, vol. 38, no. 8, pp. 677–685, 2002.

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

[24] J. Löfberg, “YALMIP: a toolbox for modeling and optimization in MATLAB,” in *Proceedings of the CACSD Conference*, Taipei, Taiwan, September 2004, available from `http://control.ee.ethz.ch/~joloef/yalmip.php`

[25] I. Masubuchi, T. Akiyama, and M. Saeki, “Synthesis of output feedback gain-scheduling controllers based on the descriptor LPV system representation,” in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, USA, December 2003, pp. 6115–6120.

[26] I. Masubuchi and E. Shimemura, “On application of the descriptor form to design of gain scheduling systems” (in Japanese), *Transactions of the Institute of Systems, Control and Information Engineers*, vol. 12, no. 7, pp. 390–394, 1999.

[27] Y. Matsuoka, Y. Ebihara, and T. Hagiwara, “A dilated LMI approach to robust performance analysis of time-invariant systems and its improvement by sectioning of parameter regions” (in Japanese), in *Proceedings of the 47th ISCIE Annual Conference*, Kyoto, Japan, May 2003, pp. 725–726.

[28] A. Ohara and Y. Sasaki, "On solvability and numerical solutions of parameter-dependent differential matrix inequality," in *Proceedings of the First SICE Annual Conference on Control Systems*, Kyoto, Japan, May 2001, pp. 61–64.

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

[30] P. A. Parrilo, "Exploiting algebraic structure in sum of squares programs," in *Positive Polynomials in Control*, D. Henrion and A. Garulli, Eds. Berlin, Germany: Springer, 2005, pp. 181–194.

[31] C. W. Scherer, "Relaxations for robust linear matrix inequality problems with verifications for exactness," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 2, pp. 365–395, 2005.

[32] C. W. Scherer and C. W. J. Hol, "Matrix sum-of-squares relaxations for robust semi-definite programs," *Mathematical Programming*, to appear.

[33] U. Shaked, "Improved LMI representation for the analysis and the design of continuous-time systems with polytopic type uncertainty," *IEEE Transactions on Automatic Control*, vol. 46, no. 4, pp. 652–656, 2001.

[34] T. Shimomura, M. Takahashi, and T. Fujii, "Extended-space control design with parameter-dependent Lyapunov functions," in *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, USA, December 2001, pp. 2157–2162.

[35] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems*. London, UK: Springer, 2005.

[36] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity," *SIAM Journal on Optimization*, to appear.