MATHEMATICAL ENGINEERING TECHNICAL REPORTS

Exploiting Sparsity in the Matrix-Dilation Approach to Robust Semidefinite Programming

Yasuaki OISHI and Yusuke ISAKA

METR 2006–63 November 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

Exploiting Sparsity in the Matrix-Dilation Approach to Robust Semidefinite Programming*

Yasuaki Oishi† and Yusuke Isaka†

A computationally improved approach is proposed for a robust semidefinite programming problem whose constraint is polynomially dependent on uncertain parameters. By exploiting sparsity, the proposed approach gives an approximate problem smaller in size than the matrix-dilation approach formerly proposed by the group of the first author. Here, the sparsity means that the constraint of the given problem has only a small number of nonzero terms when it is expressed as a polynomial of the uncertain parameters. This sparsity is extracted with a special graph called a rectilinear Steiner arborescence, based on which a reduced-size approximate problem is constructed. The quality of the approximation can be evaluated quantitatively. This evaluation shows that the quality can be improved to any level by dividing the parameter region into small subregions.

Keywords: robust semidefinite programming, linear matrix inequalities, matrix dilation, sparsity, rectilinear Steiner arborescences, computational cost, conservatism.

1. Introduction

Robust semidefinite programming (robust SDP in short) is an important optimization problem that has many applications in nonlinear optimization and robust control. See [1, 9, 3, 5, 31] for surveys. It is an optimization of a linear objective function subject to a linear matrix inequality (LMI in short) constraint whose coefficients depend on uncertain parameters. The LMI constraint is required to be satisfied for all possible parameter values. A robust SDP is difficult to solve in general. Indeed, it is proven to be NP-hard even in a simple case that the parameter dependence of the constraint is affine and the parameter region is box-shaped [21]. Hence, approximate approaches have been considered [2, 10, 4, 5], where a usual and thus solvable SDP problem is constructed as an approximation of a given robust SDP problem.

In general, there is a nonzero approximation error between the optimal values of the original robust SDP problem and the constructed approximate problem. Recently, asymptotically exact approaches were proposed in the case of polynomial parameter dependence. In these approaches, one can reduce the approximation error to any level by allowing the increase of the

*November 30, 2006

†Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Correspondence: Yasuaki Oishi (phone: +81-3-5841-6921; fax: +81-3-5841-6886; email: oishi@mist.i.u-tokyo.ac.jp)

size of the approximate problem, that is, the increase of the number of variables and/or the increase of the size of the LMI constraints. More interestingly, it is reported that an approximate problem of a finite size often gives an almost exact result. In particular, Ohara–Sasaki [22] and Bliman [6] proposed an approach based on the Kalman–Yakubovich–Popov lemma, Scherer [30] an approach based on Pólya's theorem, Scherer–Hol [32] an approach based on the sum of squares (SOS in short; see also the works of Lasserre [16], Parrilo [24], and Kojima [13]), and the group of the first author an approach based on matrix dilation [11, 23]. On the other hand, these approaches have a computational drawback. Namely, when they are applied to a practical problem, the constructed approximate problem often becomes too large for the currently available SDP solvers. This may be a natural consequence of the difficulty of the original problem. However, still there is a possibility that a small-size approximate problem is constructed for a special class of robust SDP problems of practical importance.

In this paper, we consider reduction of the size of the approximate problem in the matrix-dilation approach. The key idea is to assume a kind of sparsity in the given robust SDP problem and to exploit it with a special graph called a rectilinear Steiner arborescence. To be specific, suppose that the LMI constraint of the given robust SDP problem polynomially depends on a p-dimensional uncertain parameter $\theta = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_p]^{\mathrm{T}}$. Let d_i denote the maximum degree of the LMI constraint as a polynomial in θ_i for each of $i = 1, 2, \ldots, p$. Then, in general, the LMI constraint has $\prod_{i=1}^{p}(d_i + 1)$ terms when it is expressed as a polynomial in θ. In the matrix-dilation approach of [23], all of these terms are supposed to have nonzero coefficients. However, this is not believed to be the case in many practical problems. We assume in this paper that only a small number of terms has nonzero coefficients. We express these nonzero terms as integral points in the p-dimensional Euclidean space and construct a rectilinear Steiner arborescence covering these points. Using the properties of the rectilinear Steiner arborescence, we can construct an approximate problem. The size of the arborescence determines the size of the resulting approximate problem, which is always smaller than or equal to that of the approximate problem in [23]. The discrepancy is especially evident when the number of nonzero terms is small and the degrees d_i are high.

Even with the above improvement, an advantage of the matrix-dilation approach is still kept. Namely, an *a priori* upper bound can be obtained on the approximation error. This upper bound can be reduced to any level by dividing the parameter region into small subregions. Hence, the asymptotic exactness of the approach is guaranteed. The upper bound is also useful in making a good approximation with a small computational cost. No corresponding result has been obtained in other asymptotically exact approaches though an attempt toward this direction is found in [7].

In the recent works [27, 29], an approach close to [11, 23] is independently taken for special robust SDP problems related to robust control. The reduction technique in the present paper is readily applicable to these problems. Exploitation of sparsity has been considered in the SOS approach [25, 12, 14, 35, 17, 15]. The techniques for the exploitation are, however, quite different from the present one. Comparison with these techniques is a future research subject.

The construction of this paper is as follows. Section 2 provides a robust SDP problem as well as the outline of the matrix-dilation approach. The succeeding two sections provide main results of this paper. In particular, Section 3 gives a reduced-size approximate problem while Section 4 provides an upper bound on the approximation error. After a numerical example is presented in Section 5, the paper is concluded in Section 6.

The symbol $\mathbb{R}^p$ stands for the set of p-dimensional real vectors while $\mathbb{Z}_+^p$ stands for the set of p-dimensional vectors of nonnegative integers. The symbol $^{\mathrm{T}}$ denotes the transpose of a matrix or a vector. For $\theta = [\theta_1 \;\; \theta_2 \;\; \cdots \;\; \theta_p]^{\mathrm{T}} \in \mathbb{R}^p$ and $\alpha = [\alpha_1 \;\; \alpha_2 \;\; \cdots \;\; \alpha_p]^{\mathrm{T}} \in \mathbb{Z}_+^p$, the symbol θ^α means the product $\theta_1^{\alpha_1}\theta_2^{\alpha_2}\cdots\theta_p^{\alpha_p}$. The symbols $O_{q\times r}$ and I_q designate the $q \times r$ zero matrix and the $q \times q$ identity matrix, respectively. The sizes of these matrices are omitted when they are obvious 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 inequality $A \succeq B$ means $A - B \succeq O$. The Kronecker product of two (not necessarily symmetric) matrices $A = (a_{ij})$ and B is defined as

$$A \otimes B := \begin{bmatrix} a_{11}B & \cdots & a_{1r}B \\ \vdots & & \vdots \\ a_{q1}B & \cdots & a_{qr}B \end{bmatrix}.$$

There hold $(A\otimes B)^{\mathrm{T}} = A^{\mathrm{T}} \otimes B^{\mathrm{T}}$ and $(A\otimes B)(C\otimes D) = (AC)\otimes(BD)$ for matrices of appropriate sizes. For a set S, the symbol $|S|$ denotes its cardinality.

2. A robust SDP problem and the matrix-dilation approach

In this section, we present a robust SDP problem as well as the outline of the matrix-dilation approach, which was proposed in [11] and analyzed in [23].

A robust SDP problem considered in this paper is the following:

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

Here, $c \in \mathbb{R}^n$ is a given 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 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 $F(x,\theta)$ as a polynomial in θ_i for $i = 1, 2, \ldots, p$. With $V_0 := \{\alpha \in \mathbb{Z}_+^p \mid 0 \le \alpha_i \le d_i \text{ for } i = 1, 2, \ldots, p\}$, we can write $F(x,\theta) = \sum_{\alpha \in V_0} F_\alpha(x)\theta^\alpha$. Note that $|V_0| = \prod_{i=1}^p (d_i + 1)$. We can assume that $\max_{i=1,2,\ldots,p} d_i \ge 1$ without loss of generality. Otherwise, $F(x,\theta) \succeq O$ is independent of θ and the problem P is easy to solve.

It is difficult to solve the problem P directly. The matrix-dilation approach of [11, 23] is an approximate approach to overcome this difficulty, which is a generalization of the robust control technique of Masubuchi–Shimemura [19] within the framework of matrix dilation [8, 26]. The difficulty of P originates from the semi-infinite constraint $F(x,\theta) \succeq O$ $(\forall \theta \in \Theta)$. In the matrix-dilation approach, this semi-infinite constraint is replaced by its sufficient condition expressed as a finite number of usual LMI constraints. Construction of the sufficient condition is based on matrix dilation and division of the parameter region. The approximate problem thus obtained has the minimum value, in general, larger than that of the original problem P. The approximation error converges to zero, however, as the maximum size of the subregions tends to zero.

In order to present the approximate problem explicitly, we define

$$\begin{aligned} F_*(x) &:= [F_{\alpha^{(2)}}(x)\quad F_{\alpha^{(3)}}(x)\quad \cdots \quad F_{\alpha^{(|V_0|)}}(x)], \\ G(x) &:= \begin{bmatrix} 2F_{\alpha^{(1)}}(x) & F_*(x) \\ F_*(x)^{\mathrm{T}} & O \end{bmatrix}, \\ M(\theta) &:= [\theta^{\alpha^{(1)}} I_m \quad \theta^{\alpha^{(2)}} I_m \quad \cdots \quad \theta^{\alpha^{(|V_0|)}} I_m]^{\mathrm{T}}, \end{aligned}$$

where $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(|V_0|)}$ are the elements of V_0 with $\alpha^{(1)}$ being the origin. The matrix $G(x)$ is $|V_0|m \times |V_0|m$ while the matrix $M(\theta)$ is $|V_0|m \times m$. They satisfy $2F(x,\theta) = M(\theta)^{\mathrm{T}} G(x) M(\theta)$. Furthermore, we consider a $|V_0|m \times (|V_0| - 1)m$ matrix $H(\theta)$ such that the square matrix $[M(\theta)\ \ H(\theta)]$ is nonsingular and the relation $M(\theta)^{\mathrm{T}} H(\theta) = O$ holds for all $\theta \in \mathbb{R}^p$. Such $H(\theta)$ is called an *orthogonal complement* of $M(\theta)$. The key fact is that the orthogonal complement $H(\theta)$ can be chosen to be affine in θ [23]. This establishes the following fact, which provides a basis for the present approach.

Lemma 1. *Let x be any point in $\mathbb{R}^n$ and $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(Q)}$ be any points in $\mathbb{R}^p$. Then, $F(x,\theta) \succeq O$ holds for all θ in the convex hull of $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(Q)}\}$ if there exists W satisfying*

$$G(x) + H(\theta^{(q)})W^{\mathrm{T}} + WH(\theta^{(q)})^{\mathrm{T}} \succeq O \quad (\forall q = 1, 2, \ldots, Q). \tag{1}$$

Proof. Choose any θ in the convex hull of $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(Q)}\}$ and express it as a convex combination. Convex combination of (1) with the same coefficients gives $G(x) + H(\theta)W^{\mathrm{T}} + WH(\theta)^{\mathrm{T}} \succeq O$ for the chosen θ because of the affinity of $H(\theta)$. Premultiplication of $M(\theta)^{\mathrm{T}}$ and postmultiplication of $M(\theta)$ to this inequality provide $M(\theta)^{\mathrm{T}}G(x)M(\theta) = 2F(x, \theta) \succeq O$. □

We now formally state the matrix-dilation approach. Let $\mathit{\Delta} = \{\mathit{\Theta}^{[j]}\}_{j=1}^{J}$ be a *division* of $\mathit{\Theta}$, *i.e.*, a family of convex polytopes with nonempty interiors such that the equality $\mathit{\Theta} = \bigcup_{j=1}^{J} \mathit{\Theta}^{[j]}$ holds and the set $\mathit{\Theta}^{[j]} \cap \mathit{\Theta}^{[k]}$ has an empty interior whenever $j \neq k$. For a convex polytope $\mathit{\Theta}^{[j]}$ with a nonempty interior, let $\operatorname{ver} \mathit{\Theta}^{[j]}$ denote the set of its vertices. For a given division $\mathit{\Delta} = \{\mathit{\Theta}^{[j]}\}_{j=1}^{J}$, we consider the following approximate problem:

$$
\begin{aligned}
P(\mathit{\Delta}): \quad & \text{minimize} \quad c^{\mathrm{T}}x \\
& \text{subject to} \quad E(x) \succeq O, \quad G(x) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]}H(\theta)^{\mathrm{T}} \succeq O \\
& \qquad\qquad\qquad (\forall \theta \in \operatorname{ver} \mathit{\Theta}^{[j]}; \quad \forall j = 1, 2, \ldots, J).
\end{aligned}
$$

This problem $P(\mathit{\Delta})$ has only a finite number of constraints and, thus, is solvable as a usual SDP problem. From Lemma 1, the relationship between this problem and the original problem follows, which is stated in the next proposition. Here, noting that the feasible region of $P(\mathit{\Delta})$ is in the space of $(x, \{W^{[j]}\}_{j=1}^{J})$, we refer to its projection to the x-space as the *projected feasible region* of $P(\mathit{\Delta})$.

Proposition 2. *Let $\mathit{\Delta}$ be a division of $\mathit{\Theta}$. Then, the projected feasible region of the approximate problem $P(\mathit{\Delta})$ is contained in the feasible region of the original problem P. In particular, $\min P \leq \min P(\mathit{\Delta})$.*

In the matrix-dilation approach, we solve the approximate problem $P(\mathit{\Delta})$ in place of the original problem P. Although there is a nonzero approximation error $\min P(\mathit{\Delta}) - \min P$ in general, we can make it smaller by subdividing each subregion in $\mathit{\Delta}$ and considering the corresponding approximate problem. Indeed, the approximation error is known to have an upper bound proportional to the maximum radius of the division [23]. Here, the *radius* of a subregion $\mathit{\Theta}^{[j]}$ is $\operatorname{rad} \mathit{\Theta}^{[j]} := \min_{\theta \in \mathit{\Theta}^{[j]}} \max_{\theta' \in \mathit{\Theta}^{[j]}} \max_{i=1,2,\ldots,p} |\theta_i - \theta'_i|$, where a θ that attains the minimum is called a *center* of $\mathit{\Theta}^{[j]}$. The *maximum radius* of a division $\mathit{\Delta} = \{\mathit{\Theta}^{[j]}\}_{j=1}^{J}$ is defined as $\overline{\operatorname{rad}}\, \mathit{\Delta} := \max_{j=1,2,\ldots,J} \operatorname{rad} \mathit{\Theta}^{[j]}$.

From a computational point of view, the approximate problem is desired to have a small size. Unfortunately, the size of the approximate problem tends to be large for a practical problem even with a coarse division. Sometimes it exceeds the capability of the currently available SDP solvers. This is partly because $|V_0| = \prod_{i=1}^{p}(d_i + 1)$ increases rapidly as the degrees d_i increase.

In the next section, we assume a kind of sparsity in the given problem P and use it for the construction of reduced-size $G(x)$, $M(\theta)$, and $H(\theta)$. This leads to a reduced-size approximate

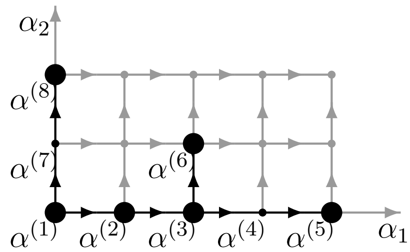

Figure 1. The directed graph (V_0, A_0) (gray) in the case of $p = 2$, $d_1 = 4$, and $d_2 = 2$ together with a rectilinear Steiner arborescence (V, A) (black) for the set S whose elements are shown by the large circles

problem. Even with this improvement, an upper bound on the approximation error remains available, whose derivation is a task of Section 4.

3. A reduced-size approximate problem

In this section, we consider reduction of the size of the approximate problem. The key idea is to exploit sparsity in the given robust SDP problem P using a special graph called a rectilinear Steiner arborescence. This graph enables us to redefine $G(x)$, $M(\theta)$, and $H(\theta)$ with smaller matrices and, then, to obtain a reduced-size approximate problem. Let us recall the requirements on these matrices. There has to hold $2F(x,\theta) = M(\theta)^{\mathrm{T}}G(x)M(\theta)$. Moreover, $H(\theta)$ has to be an orthogonal complement of $M(\theta)$ and be affine in θ.

Recall the expression $F(x,\theta) = \sum_{\alpha\in V_0} F_\alpha(x)\theta^\alpha$. We assume that the coefficient matrices $F_\alpha(x)$ are nonzero for only a small number of α's. This is a kind of sparsity of the given robust SDP problem P. Such sparsity is observed in many practical situations. We let S be the *support* of $F(x,\theta)$ defined as $S := \{\alpha \in V_0 \subset \mathbb{Z}_+^p \mid F_\alpha(x) \not\equiv O\}$. Since $F(x,\theta)$ is not independent of θ, the support S contains at least one element not being the origin.

In order to exploit this sparsity, we first embed $V_0 \in \mathbb{Z}_+^p$ into $\mathbb{R}^p$ in a natural way. We consider a directed graph (V_0, A_0) in $\mathbb{R}^p$ with the set of vertices being V_0 and with the set of arcs being $A_0 = \{(\alpha,\beta) \mid \alpha,\beta \in V_0,\ \alpha_i + 1 = \beta_i$ for some $i = 1, 2, \ldots, p$, and $\alpha_j = \beta_j$ for all $j \neq i\}$. In a word, the arcs are the line segments of length one connecting two vertices and directed away from the origin. When an arc (α,β) satisfies $\alpha_i + 1 = \beta_i$, it is said to be *parallel* to the ith axis.

For the support S, we consider a subgraph (V, A) of (V_0, A_0) having the following properties: (i) V contains any vertex in S as well as the origin; (ii) any vertex in V is reachable from the origin through a unique path in (V, A). Here, a vertex α is said to be *reachable* from a vertex β if either $\alpha = \beta$ or there is a path connecting β to α through the arcs in the directed way.

Such a graph (V, A) is referred to as a *rectilinear Steiner arborescence* for S (see [28] and the references therein). Figure 1 gives an example in the case of $p = 2$, $d_1 = 4$, and $d_2 = 2$. The directed graph (V_0, A_0) is presented in gray. For the set S whose elements are expressed by the large circles, a rectilinear Steiner arborescence is shown in black.

The rectilinear Steiner arborescence (V, A) is important for redefinition of $G(x)$, $M(\theta)$, and $H(\theta)$. In particular, Property (i) is used to have $2F(x, \theta) = M(\theta)^{\mathrm{T}} G(x) M(\theta)$ while Property (ii) is used to have an affine orthogonal complement $H(\theta)$. The rectilinear Steiner arborescence is desired to have a small $|V|$, though not necessarily the smallest, because this leads to small-size $G(x)$, $M(\theta)$, and $H(\theta)$. Note that $2 \le |V| \le |V_0|$.

With these preparations, we now redefine the required matrices. Let the vertices in V be $\alpha^{(1)}$, $\alpha^{(2)}$, ..., $\alpha^{(|V|)}$. The numbering is arbitrary as far as $\alpha^{(1)}$ is the origin. For technical convenience, however, we choose the numbering to be consistent with the partial order defined by (V, A). That is, the vertex $\alpha^{(r)}$ is reachable from $\alpha^{(q)}$ only if $q \le r$. This is consistent with $\alpha^{(1)}$ being the origin. See Figure 1 for an example. With this notation, we define

$$\begin{aligned} F_*(x) &:= [F_{\alpha^{(2)}}(x) \quad F_{\alpha^{(3)}}(x) \quad \cdots \quad F_{\alpha^{(|V|)}}(x)], \\ G(x) &:= \begin{bmatrix} 2F_{\alpha^{(1)}}(x) & F_*(x) \\ F_*(x)^{\mathrm{T}} & O \end{bmatrix}, \\ M(\theta) &:= [\theta^{\alpha^{(1)}} I_m \quad \theta^{\alpha^{(2)}} I_m \quad \cdots \quad \theta^{\alpha^{(|V|)}} I_m]^{\mathrm{T}}. \end{aligned} \tag{2}$$

See Example 4 below for an example. The new $G(x)$ is $|V|m \times |V|m$ while $M(\theta)$ is $|V|m \times m$. It is easy to see that $2F(x, \theta) = M(\theta)^{\mathrm{T}} G(x) M(\theta)$. We next define

$$H(\theta) := \widetilde{H}(\theta) \otimes I_m \tag{3}$$

with the (q, r)-element of $\widetilde{H}(\theta)$ being

$$\widetilde{H}(\theta)_{qr} = \begin{cases} -\theta_i, & \text{if the arc } (\alpha^{(q)}, \alpha^{(r+1)}) \text{ belongs to } A \text{ and is parallel to the } i\text{th axis;} \\ 1, & \text{if } q = r + 1; \\ 0, & \text{otherwise} \end{cases}$$

for $q = 1, 2, \ldots, |V|$ and $r = 1, 2, \ldots, |V| - 1$. The size of $H(\theta)$ is thus $|V|m \times (|V| - 1)m$. Obviously, $H(\theta)$ is affine in θ. The next lemma states the relationship between $M(\theta)$ and $H(\theta)$.

Lemma 3. *The matrix $H(\theta)$ is an orthogonal complement of $M(\theta)$.*

Proof. By the definition, the matrix $\widetilde{H}(\theta)$ is upper triangular in the sense that $\widetilde{H}(\theta)_{qr} \neq 0$ only if $q \le r + 1$. Moreover, $\widetilde{H}(\theta)_{qr} = 1$ if $q = r + 1$. Hence, the matrix $\widetilde{H}(\theta)$ is of column full rank and, thus, so is $H(\theta)$. Since $M(\theta)$ is clearly of column full rank, the nonsingularity of

$[M(\theta)\ \ H(\theta)]$ follows from the orthogonality $M(\theta)^{\mathrm{T}}H(\theta) = O$. Let us show the orthogonality. Write $M(\theta) = \widetilde{M}(\theta) \otimes I_m$ with $\widetilde{M}(\theta) = [\theta^{\alpha^{(1)}}\ \ \theta^{\alpha^{(2)}}\ \ \cdots\ \ \theta^{\alpha^{(|V|)}}]^{\mathrm{T}}$. Noting that $M(\theta)^{\mathrm{T}}H(\theta) = [\widetilde{M}(\theta)^{\mathrm{T}}\widetilde{H}(\theta)] \otimes I_m$, we consider the product between $\widetilde{M}(\theta)^{\mathrm{T}}$ and the rth column vector of $\widetilde{H}(\theta)$. Since (V, A) is an arborescence, there is one and only one q such that the arc $(\alpha^{(q)}, \alpha^{(r+1)})$ belongs to A. Let us write this q as $\widehat{q}$ and suppose that the arc $(\alpha^{(\widehat{q})}, \alpha^{(r+1)})$ is parallel to the ith axis. Then, $\widetilde{H}(\theta)_{qr}$ is equal to $-\theta_i$ for $q = \widehat{q}$, equal to unity for $q = r + 1$, and equal to zero otherwise. Hence, the considered product is equal to $\theta^{\alpha^{(\widehat{q})}}(-\theta_i) + \theta^{\alpha^{(r+1)}}$, which is equal to zero. Repeating this reasoning for each r, we arrive at the desired orthogonality. $\square$

The matrices defined above have the required properties and, hence, can be used in the matrix-dilation approach. Namely, with $\Delta = \{\Theta^{[j]}\}_{j=1}^{J}$ being a division of Θ, we consider a new approximate problem

$$\begin{aligned} P(\Delta):\ \ & \text{minimize} && c^{\mathrm{T}}x \\ & \text{subject to} && E(x) \succeq O, \quad G(x) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]}H(\theta)^{\mathrm{T}} \succeq O \\ & && (\forall \theta \in \operatorname{ver}\Theta^{[j]};\ \ \forall j = 1, 2, \ldots, J) \end{aligned}$$

with $G(x)$ and $H(\theta)$ in (2) and (3), respectively. Then, Proposition 2 holds on this problem. As is mentioned in Section 2, we can reduce the approximation error by subdividing Δ. It is not obvious, however, whether a quantitative relationship can be obtained in this case between the approximation error and the maximum radius of the division. Indeed, the technique in [23] cannot be used in the present setting. In the next section, we use another technique to derive an upper bound of the approximation error.

Example 4. Let us consider maximization of $f(\theta) = -100(\theta_2 - \theta_1^2)^2 - (1 - \theta_1)^2$ in $\theta \in \Theta = [0, 2]^2$. Its maximum value $f(\theta) = 0$ is attained at $\theta = [1\ \ 1]^{\mathrm{T}}$. This function is known as the Rosenbrock function and is often used as a benchmark for nonlinear optimization [20]. Maximization of this $f(\theta)$ is formulated as the following robust SDP problem:

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

In this case, $F(x, \theta) = x - f(\theta) = x + 1 - 2\theta_1 + \theta_1^2 + 100\theta_1^4 - 200\theta_1^2\theta_2 + 100\theta_2^2$ with $n = 1$, $\ell = 0$, $m = 1$, $p = 2$, $d_1 = 4$, and $d_2 = 2$. Its support S is as presented in Figure 1. Using the rectilinear Steiner arborescence in the figure, we obtain the matrices

$$F_{\alpha^{(1)}}(x) = x + 1, \qquad F_*(x) = [-2\ \ 1\ \ 0\ \ 100\ \ -200\ \ 0\ \ 100],$$

$$
M(\theta) = \begin{bmatrix} 1 \\ \theta_1 \\ \theta_1^2 \\ \theta_1^3 \\ \theta_1^4 \\ \theta_1^2\theta_2 \\ \theta_2 \\ \theta_2^2 \end{bmatrix}, \qquad H(\theta) = \begin{bmatrix} -\theta_1 & & & & & -\theta_2 & \\ 1 & -\theta_1 & & & & & \\ & 1 & -\theta_1 & & -\theta_2 & & \\ & & 1 & -\theta_1 & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & 1 & -\theta_2 \\ & & & & & & 1 \end{bmatrix}.
$$

With these matrices, we can construct an approximate problem for a division of Θ. □

We close this section by discussing the size of the new approximate problem. In the new approximate problem, the dilated LMI constraint $G(x) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]}H(\theta)^{\mathrm{T}} \succeq O$ has the size $|V|m \times |V|m$ while $|V_0|m \times |V_0|m$ in the conventional approximate problem. Since $|V| \leq |V_0|$, the new approximate problem has the size smaller than or equal to that of the conventional one. This discrepancy becomes especially evident in some cases. This is seen from the next result.

Proposition 5. *For a given support S, there exists a rectilinear Steiner arborescence (V, A) with $|V| \leq |S| \sum_{i=1}^{p} d_i + 1$.*

Proof. For each $\alpha = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_p]^{\mathrm{T}} \in S$, consider a directed path connecting $[0 \ 0 \ \cdots \ 0]^{\mathrm{T}}$, $[1 \ 0 \ \cdots \ 0]^{\mathrm{T}}, \ldots, [\alpha_1 \ 0 \ \cdots \ 0]^{\mathrm{T}}, [\alpha_1 \ 1 \ \cdots \ 0]^{\mathrm{T}}, \ldots, [\alpha_1 \ \alpha_2 \ \cdots \ 0]^{\mathrm{T}}, \ldots, [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_p]^{\mathrm{T}}$ in this order. The length of this path is less than or equal to $\sum_{i=1}^{p} d_i$. The union of these $|S|$ paths forms a rectilinear Steiner arborescence for S. It is clear that this arborescence (V, A) satisfies $|V| \leq |S| \sum_{i=1}^{p} d_i + 1$. □

Recall that $|V_0| = \prod_{i=1}^{p}(d_i + 1)$. When $|S|$ is small and d_i's are large, the $|V|$ in the proposition is much smaller than $|V_0|$.

It is NP-hard to find a rectilinear Steiner arborescence with the smallest $|V|$ for a given S [33]. Fortunately, what we need is not an arborescence with the smallest $|V|$ but an arborescence with a small $|V|$. A simple heuristic algorithm for this purpose is found in [28].

Remark 6. Suppose that the support S is equal to the whole vertex set V_0. If we construct a rectilinear Steiner arborescence for this S as in the proof of Proposition 5, we obtain the full-size $G(x)$, $M(\theta)$, and $H(\theta)$ in the previous section. In this sense, the present approach is a generalization of the conventional approach in [23]. □

Remark 7. In general, the reduced-size approximate problem has a larger approximation error $\min P(\Delta) - \min P$ than the full-size approximate problem in the previous section. Although no quantitative result has been known on this possible deterioration, computational experience tells that it is not as evident as the profit of the size reduction. See Section 5. □

4. An upper bound on the approximation error

In this section, we give an upper bound on the approximation error of the reduced-size approximate problem. This upper bound is proportional to the maximum radius $\overline{\mathrm{rad}}\,\Delta$ of the division and hence gives a quantitative relationship between the approximation error and the resolution of the division. This result is an extension of that in [23], which is on the full-size approximate problem.

The key idea is to relate the following auxiliary problem with the approximate problem $P(\Delta)$:

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

where ϵ is a nonnegative number. We will show below that ϵ can be chosen so that $\min P \le \min P(\Delta) \le \min P_\epsilon$. Note that $\min P_\epsilon$ is convex in ϵ and that P_0 is identical with the original problem P. With g being an upper bound on a subgradient of $\min P_\epsilon$, we have $\min P_\epsilon - \min P \le g\epsilon$. Since this implies $\min P(\Delta) - \min P \le g\epsilon$, we obtain the desired upper bound on the approximation error. This basic idea is the same as in the full-size case [23]. However, the technique used in [23] for the choice of ϵ is not applicable to the reduced-size approximate problem because the whole set of $F_\alpha(x)$, $\alpha \in V_0$, is not used there. We, hence, introduce a different technique with a coordinate transformation matrix $L(\theta)$.

We need the following assumption to give the result.

Assumption 8.

(a) The robust SDP problem P is strictly feasible, that is, there exists $x \in \mathbb{R}^n$ such that $E(x) \succ O$ and $F(x,\theta) \succ O$ $(\forall\theta \in \Theta)$.

(b) For any v, the level set $\{x \in \mathbb{R}^n \mid c^{\mathrm{T}}x \le v,\ E(x) \succeq O, \text{ and } F(x,\theta) \succeq O\ (\forall\theta \in \Theta)\}$ is bounded. □

We now present the main result of this section, which provides the desired upper bound.

Theorem 9. *Suppose that Assumption 8 holds. Then, the reduced-size approximate problem $P(\mathit{\Delta})$ satisfies*

$$\min P(\mathit{\Delta}) - \min P \leq C_1 \overline{\mathrm{rad}}\, \mathit{\Delta}. \tag{4}$$

for any division $\mathit{\Delta}$ with $\overline{\mathrm{rad}}\, \mathit{\Delta} \leq C_2$, where C_1 and C_2 are positive numbers independent of $\mathit{\Delta}$.

The specific forms of C_1 and C_2 will be given in (9) and (10), respectively.

A direct consequence of this theorem is the asymptotic exactness of our matrix-dilation approach. That is, the approximation error of the reduced-size approximate problem converges to zero as the maximum radius of the division goes to zero. Evaluation of C_1 and C_2 is possible as in [23], though the resulting bound is often conservative. Recall that the existing asymptotically exact approaches [22, 6, 30, 32] do not have a corresponding quantitative result. Their asymptotic exactness is proven only qualitatively.

The bound (4) also gives a relationship between the approximation error and the size of the approximate problem. Namely, in order to reduce the approximation error, we need to decrease the maximum radius, which increases the number of subregions and, then, the number of the LMI constraints. Especially when the parameter dimension is high, this increase is rapid and makes it difficult to solve the approximate problem. It is possible to partially address this issue by improvement of the bound and adaptive division of the parameter region. The discussion is completely parallel to that in [23]. The details are omitted.

The rest of this section is devoted to the proof of Theorem 9. We first prepare the $|V|m \times |V|m$ matrix $L(\theta)$, which will be used for simplification of the dilated LMI constraint $G(x) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]}H(\theta)^{\mathrm{T}} \succeq O$. This matrix is defined as

$$L(\theta) := \widetilde{L}(\theta) \otimes I_m \tag{5}$$

with

$$\widetilde{L}(\theta)_{qr} = \begin{cases} \theta^{\alpha^{(q)} - \alpha^{(r)}}, & \text{if } \alpha^{(q)} \text{ is reachable from } \alpha^{(r)} \text{ in } (V, A); \\ 0, & \text{otherwise.} \end{cases}$$

The matrix $\widetilde{L}(\theta)$ is lower triangular, *i.e.*, $\widetilde{L}(\theta)_{qr} \neq 0$ only if $q \geq r$. Moreover, its diagonal elements are all equal to unity. A consequence is nonsingularity of $\widetilde{L}(\theta)$ and also of $L(\theta)$.

The product $L(\theta)^{\mathrm{T}} G(x) L(\theta)$ has the following form.

Lemma 10. *For the matrices $G(x)$ and $L(\theta)$ in* (2) *and* (5)*, respectively, we can write*

$$L(\theta)^{\mathrm{T}} G(x) L(\theta) = \begin{bmatrix} 2F(x,\theta) & F_{**}(x,\theta) \\ F_{**}(x,\theta)^{\mathrm{T}} & O \end{bmatrix}$$

with

$$F_{**}(x,\theta) = \Big[\sum_{\alpha \in V^{(2)}} F_\alpha(x)\theta^{\alpha - \alpha^{(2)}} \quad \sum_{\alpha \in V^{(3)}} F_\alpha(x)\theta^{\alpha - \alpha^{(3)}} \quad \cdots \quad \sum_{\alpha \in V^{(|V|)}} F_\alpha(x)\theta^{\alpha - \alpha^{(|V|)}} \Big],$$

where $V^{(r)}$ is the set of vertices reachable from $\alpha^{(r)}$ in (V, A) for $r = 2, 3, \dots, |V|$.

Proof. Direct calculation gives the lemma. □

We next consider the product $L(\theta)^{\mathrm{T}}H(\theta')$ for $\theta, \theta' \in \mathbb{R}^p$.

Lemma 11. *For the matrices $H(\theta)$ and $L(\theta)$ in (3) and (5), respectively, we can write*

$$L(\theta)^{\mathrm{T}}H(\theta') = \begin{bmatrix} * & * & \cdots & * \\ 1 & * & \cdots & * \\ 0 & 1 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \otimes I_m,$$

where an element expressed by $*$ is either equal to zero or expressed as $\theta^{\alpha}(\theta_i - \theta'_i)$ for some $\alpha \in \mathbb{Z}^p_+$ and $i = 1, 2, \dots, p$. When $\theta = \theta'$ in particular, the elements expressed by $*$ are all equal to zero.

Proof. Noting that $L(\theta)^{\mathrm{T}}H(\theta') = [\widetilde{L}(\theta)^{\mathrm{T}}\widetilde{H}(\theta')] \otimes I_m$, we will evaluate the (s, r)-element of $\widetilde{L}(\theta)^{\mathrm{T}}\widetilde{H}(\theta')$, which is the inner product between the sth column of $\widetilde{L}(\theta)$ and the rth column of $\widetilde{H}(\theta')$, where $s = 1, 2, \dots, |V|$ and $r = 1, 2, \dots, |V| - 1$. As in the proof of Lemma 3, the element $\widetilde{H}(\theta')_{qr}$ is equal to $-\theta'_i$ if $q = \widehat{q}$, equal to unity if $q = r+1$, and equal to zero otherwise. Here, we suppose that the unique arc ending at $\alpha^{(r+1)}$ is $(\alpha^{(\widehat{q})}, \alpha^{(r+1)})$ and is parallel to the ith axis. The value of the considered (s, r)-element depends on the reachability of the vertices $\alpha^{(\widehat{q})}$ and $\alpha^{(r+1)}$ from $\alpha^{(s)}$ in (V, A). Namely, there are three possible cases.

(Case 1) Both $\alpha^{(\widehat{q})}$ and $\alpha^{(r+1)}$ are reachable from $\alpha^{(s)}$. This occurs only when $s \le \widehat{q} < r+1$. In this case, the considered value is equal to

$$\theta^{\alpha^{(\widehat{q})}-\alpha^{(s)}}(-\theta'_i) + \theta^{\alpha^{(r+1)}-\alpha^{(s)}} = \theta^{\alpha^{(\widehat{q})}-\alpha^{(s)}}(\theta_i - \theta'_i).$$

Note that $\alpha^{(\widehat{q})} - \alpha^{(s)} \in \mathbb{Z}^p_+$.

(Case 2) The vertex $\alpha^{(r+1)}$ coincides with $\alpha^{(s)}$, *i.e.*, $s = r + 1$. In this case, the quantity is equal to unity.

(Case 3) Neither $\alpha^{(\widehat{q})}$ nor $\alpha^{(r+1)}$ is reachable from $\alpha^{(s)}$. In this case, the quantity is equal to zero.

Summarizing these results, we obtain the statement of the lemma. □

With these preparations, we relate P_ϵ and $P(\varDelta)$. By Assumption 8, 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

$$X := \left\{x \in \mathbb{R}^n \,\middle|\, c^{\mathrm{T}}x \le v_0,\ E(x) \succeq O,\ \text{and}\ F(x, \theta) \succeq O\ (\forall \theta \in \varTheta)\right\},$$

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.

We begin by the special case that $\Theta \subseteq [-1,1]^p$. In this case, $|\theta_i| \le 1$ for any $\theta \in \Theta$ and any $i = 1, 2, \ldots, p$. Let U be a number such that

$$\max_{x \in X} \max_{\theta \in \Theta} \overline{\sigma}[F_{**}(x,\theta)] \le U.$$

Lemma 12. *Suppose that* $\Theta \subseteq [-1,1]^p$ *and*

$$\overline{\mathrm{rad}}\,\Delta \le \min\left\{\frac{2\epsilon_0}{(U+\sqrt{|V|m})^2},\ \ \frac{1}{|V|}\right\}.$$

Then, we have $\min P \le \min P(\Delta) \le \min P_\epsilon$ *for*

$$\epsilon = \frac{(U+\sqrt{|V|m})^2}{2}\overline{\mathrm{rad}}\,\Delta.$$

Proof. Since $\min P \le \min P(\Delta)$ as is noticed in Section 3, we will show $\min P(\Delta) \le \min P_\epsilon$. Let Δ be $\{\Theta^{[j]}\}_{j=1}^J$. The ϵ given in the lemma satisfies $0 \le \epsilon \le \epsilon_0$. Hence, the minimum of P_ϵ is attained at a point in X. Let such a minimizing point be $x \in X$. Note that $F(x,\theta) \succeq \epsilon I$ for any $\theta \in \Theta$. The proof is complete if this x is contained in the projected feasible region of $P(\Delta)$, that is, there exists $W^{[j]}$ for each $j = 1, 2, \ldots, J$ such that

$$G(x) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]}H(\theta)^{\mathrm{T}} \succeq O \quad (\forall \theta \in \mathrm{ver}\,\Theta^{[j]}). \tag{6}$$

We show that this inequality holds with $W^{[j]} := (1/\mathrm{rad}\,\Theta^{[j]})H(\theta^{\mathrm{c}})$ for each j, where θ^{c} is a center of $\Theta^{[j]}$.

In order to show the desired inequality (6), we premultiply $L(\theta)^{\mathrm{T}}$ and postmultiply $L(\theta)$ to it for $\theta \in \mathrm{ver}\,\Theta^{[j]}$. Lemmas 10 and 11 give the concrete forms of $L(\theta)^{\mathrm{T}}G(x)L(\theta)$ and $L(\theta)^{\mathrm{T}}H(\theta)$. By Lemma 11 again, the product $(W^{[j]})^{\mathrm{T}}L(\theta) = (1/\mathrm{rad}\,\Theta^{[j]})H(\theta^{\mathrm{c}})^{\mathrm{T}}L(\theta)$ has the form

$$\begin{bmatrix} * & 1/\mathrm{rad}\,\Theta^{[j]} & 0 & \cdots & 0 \\ * & * & 1/\mathrm{rad}\,\Theta^{[j]} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & 1/\mathrm{rad}\,\Theta^{[j]} \end{bmatrix} \otimes I_m,$$

where an element expressed by $*$ is either equal to zero or of the form

$$\frac{\theta^\alpha(\theta_i - \theta_i^{\mathrm{c}})}{\mathrm{rad}\,\Theta^{[j]}},$$

whose magnitude is at most one since $|\theta_i| \le 1$ and $|\theta_i - \theta_i^{\mathrm{c}}| \le \mathrm{rad}\,\Theta^{[j]}$ for $i = 1, 2, \ldots, p$. Let us write the product $(W^{[j]})^{\mathrm{T}}L(\theta)$ as $[H_1 \ \ H_2]$ with the $(|V|-1)m \times m$ matrix H_1 and the $(|V|-1)m \times (|V|-1)m$ matrix H_2. Then, we have

$$L(\theta)^{\mathrm{T}}H(\theta)(W^{[j]})^{\mathrm{T}}L(\theta) = \begin{bmatrix} O_{m\times m} & O_{m\times(|V|-1)m} \\ H_1 & H_2 \end{bmatrix}.$$

Since H_1 has at most $(|V|-1)m$ nonzero elements whose magnitude is at most one, we have $\overline{\sigma}(H_1) \le \sqrt{(|V|-1)m}$. On the other hand, H_2 is lower triangular and each of its columns has the diagonal element $1/\operatorname{rad}\Theta^{[j]}$ and at most $|V|-2$ nonzero off-diagonal elements, whose magnitude is at most one. Hence, we have

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

Now, the left-hand side of (6) multiplied by $L(\theta)^{\mathrm{T}}$ and $L(\theta)$ has the upper-left $m \times m$ block equal to

$$2F(x,\theta) \succeq 2\epsilon I.$$

Its Schur complement is

$$\begin{aligned} &H_2 + H_2^{\mathrm{T}} - [F_{**}(x,\theta) + H_1^{\mathrm{T}}]^{\mathrm{T}}[2F(x,\theta)]^{-1}[F_{**}(x,\theta) + H_1^{\mathrm{T}}] \\ &\quad \succeq \Big(\frac{2}{\operatorname{rad}\Theta^{[j]}} - |V| + 2\Big)I - \big\{\overline{\sigma}[F_{**}(x,\theta)] + \sqrt{(|V|-1)m}\big\}^2 \frac{1}{2\epsilon} I. \end{aligned}$$

Noting that $2/\operatorname{rad}\Theta^{[j]} - |V| + 2 \ge 1/\operatorname{rad}\Theta^{[j]}$ and $\overline{\sigma}[F_{**}(x,\theta)] \le U$, we see the positive semidefiniteness of the right-hand side matrix. This completes the proof. □

The general case that not necessarily $\Theta \subseteq [-1,1]^p$ can be reduced to the special case. Let us write

$$\overline{\theta} := \max\{1, \max_{\theta\in\Theta} \max_{i=1,2,\ldots,p} |\theta_i|\} \tag{7}$$

and $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_p$. Since

$$F(x,\theta) = \sum_{\alpha\in V} F_\alpha(x)\theta^\alpha = \sum_{\alpha\in V} F_\alpha(x)\overline{\theta}^{|\alpha|}\Big(\frac{\theta}{\overline{\theta}}\Big)^\alpha,$$

we can regard $F_\alpha(x)\overline{\theta}^{|\alpha|}$ as a coefficient and $\theta/\overline{\theta}$ as a parameter. The problems P and P_ϵ essentially remain the same with this replacement. Since $\theta/\overline{\theta}$ moves in $[-1,1]^p$, the discussion in the special case can be applied. To state the result, we define

$$\begin{aligned} \overline{F}_{**}(x,\theta) := \Big[&\sum_{\alpha\in V^{(2)}} F_\alpha(x)\overline{\theta}^{|\alpha|}\Big(\frac{\theta}{\overline{\theta}}\Big)^{\alpha-\alpha^{(2)}} \quad \sum_{\alpha\in V^{(3)}} F_\alpha(x)\overline{\theta}^{|\alpha|}\Big(\frac{\theta}{\overline{\theta}}\Big)^{\alpha-\alpha^{(3)}} \quad \cdots \\ &\sum_{\alpha\in V^{(|V|)}} F_\alpha(x)\overline{\theta}^{|\alpha|}\Big(\frac{\theta}{\overline{\theta}}\Big)^{\alpha-\alpha^{(|V|)}} \Big] \end{aligned}$$

and

$$\overline{U} := \max_{x\in X}\max_{\theta\in\Theta} \overline{\sigma}[\overline{F}_{**}(x,\theta)]. \tag{8}$$

Lemma 13. *Suppose that*

$$\overline{\mathrm{rad}}\,\Delta \le \min\left\{\frac{2\overline{\theta}\epsilon_0}{(\overline{U}+\sqrt{|V|m})^2},\ \ \frac{\overline{\theta}}{|V|}\right\}.$$

Then, we have $\min P \le \min P(\Delta) \le \min P_\epsilon$ *for*

$$\epsilon = \frac{(\overline{U}+\sqrt{|V|m})^2}{2\overline{\theta}}\overline{\mathrm{rad}}\,\Delta.$$

Proof. We observed that the replacement of $F_\alpha(x)$ by $F_\alpha(x)\overline{\theta}^{|\alpha|}$ and θ by $\theta/\overline{\theta}$ does not change the problems P and P_ϵ. As is shown below, this replacement does not change either the approximate problem $P(\Delta)$. The parameter after the replacement, *i.e.*, $\theta/\overline{\theta}$, moves in $[-1,1]^p$. Hence, the result of Lemma 12 is valid with U replaced by $\overline{U}$ and $\overline{\mathrm{rad}}\,\Delta$ by $\overline{\mathrm{rad}}\,\Delta/\overline{\theta}$. This completes the proof.

We show that the approximate problem $P(\Delta)$ does not change by the replacement above. Let $\overline{G}(x)$ and $\overline{H}(\theta)$ be the matrices obtained from $G(x)$ and $H(\theta)$, respectively, by this replacement. It is routine to confirm that

$$\begin{aligned}\overline{G}(x) &= \mathrm{diag}\{I_m,\ T\}G(x)\mathrm{diag}\{I_m,\ T\},\\ \overline{H}(\theta) &= \mathrm{diag}\{I_m,\ T\}H(\theta)T^{-1},\end{aligned}$$

where $T := \mathrm{diag}\{\overline{\theta}^{|\alpha(2)|},\ \overline{\theta}^{|\alpha(3)|},\ \ldots,\ \overline{\theta}^{|\alpha(|V|)|}\}\otimes I_m$ and diag denotes a block-diagonal matrix. Therefore, the existence of W satisfying

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

is equivalent to the existence of $\overline{W}$ satisfying

$$\overline{G}(x) + \overline{H}(\theta)\overline{W}^{\mathrm{T}} + \overline{W}\overline{H}(\theta)^{\mathrm{T}} \succeq O$$

with the correspondence $\overline{W} = \mathrm{diag}\{I_m,\ T\}WT$. This means that the approximate problem $P(\Delta)$ does not change essentially by the replacement. □

We now take the final step for the proof of Theorem 9. Recall that we assume Assumption 8 and that ϵ_0 is a number such that the auxiliary problem P_ϵ is strictly feasible for any $0 \le \epsilon \le \epsilon_0$. The numbers $\overline{\theta}$ and $\overline{U}$ are as in (7) and (8), respectively. Finally, let g be an upper bound on the left derivative of $\min P_\epsilon$ at $\epsilon = \epsilon_0$. Then, with

$$C_1 = \frac{g(\overline{U}+\sqrt{|V|m})^2}{2\overline{\theta}}, \tag{9}$$

$$C_2 = \min\left\{\frac{2\overline{\theta}\epsilon_0}{(\overline{U}+\sqrt{|V|m})^2},\ \ \frac{\overline{\theta}}{|V|}\right\}, \tag{10}$$

we can prove the theorem.

Proof of Theorem 9. Lemma 13 implies that, when $\overline{\mathrm{rad}}\,\Delta \le C_2$, we have $\min P \le \min P(\Delta) \le \min P_\epsilon$ for $\epsilon = [(\overline{U} + \sqrt{|V|m})^2/2\overline{\theta}]\overline{\mathrm{rad}}\,\Delta$. Owing to the 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, the convexity of $\min P_\epsilon$ again implies $\min P \ge \min P_\epsilon - g\epsilon$, from which $\min P_\epsilon - \min P \le g\epsilon$. Substitution of the concrete form of ϵ gives the theorem. □

5. A numerical example

We compare our reduced-size approximate problem with the full-size approximate problem in its computational cost. A numerical experiment shows the superiority of the former.

Let us consider maximization of $f_\mu(\theta) = -100(\theta_2^\mu - \theta_1^{2\mu})^2 - (1-\theta_1^\mu)^2$ in $\theta \in \Theta = [0,2]^2$, where μ is a positive integer. As is described in Example 4, this type of maximization is formulated into a robust SDP problem and is solvable with our approach. For $\mu = 1$, in particular, this problem is identical with that in Example 4 and the corresponding reduced-size approximate problem was given there. For a general μ, the support S of $F(x,\theta) = x - f_\mu(\theta)$ has the coordinate μ times larger than that of $x - f_1(\theta)$. Hence, magnifying μ times the rectilinear Steiner arborescence for $\mu = 1$, we obtain an arborescence for a general μ and then the reduced-size approximate problem. Construction of the full-size approximate problem is as in [23]. For each approximate problem, we use the coarsest division $\Delta = \{\Theta\}$, that is, we use the whole parameter region Θ as one subregion.

The computational time to solve the approximate problems is shown in Figure 2 (a). Here, we used SeDuMi [34] for the SDP solver with the help of YALMIP [18]. The computer was equipped with Pentium 4 of 2.4 GHz and 2 GByte memory. Each approximate problem gave sufficiently good approximation, whose error was less than 10^{-4}. We can see in the figure that the reduced-size approximate problem had much smaller computational time than the full-size approximate problem. Indeed, the full-size approximate problem was not solvable for $\mu \ge 3$ owing to numerical difficulties. Figure 2 (b) shows the sizes of the approximate problems. The reduced-size approximate problem had a much smaller size, which explains its small computational time.

6. Conclusion

A reduced-size approximate problem is proposed in the matrix-dilation approach to robust SDP. This reduction results from exploitation of sparsity of a given robust SDP problem. Its effect

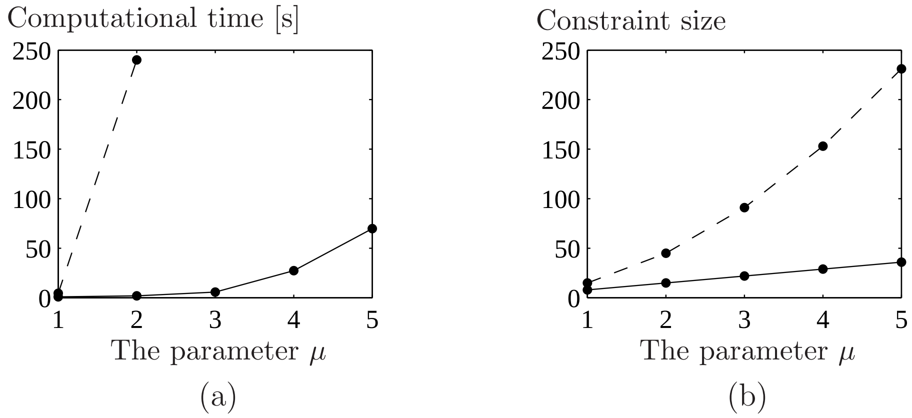

Figure 2. Comparison between the reduced-size (the solid lines) and the full-size (the broken lines) approximate problems

is especially evident when the LMI constraint of the given problem has a small support and high degrees. Even with this improvement, the good property of the matrix-dilation approach is still kept. Namely, a quantitative relationship is derived between the approximation error and the maximum radius of the division. This relationship implies the asymptotic exactness of the approach.

Acknowledgments

The authors are thankful to Tomomi Matsui for the information on a rectilinear Steiner arborescence. Their thanks go to Kazuo Murota for the careful reading of the manuscript and the helpful 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 for Scientific Research of the Japan Society for the Promotion of Science.

References

[1] 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.

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

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

[4] 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.

[5] A. Ben-Tal and A. Nemirovski, "Robust optimization: methodology and applications," *Mathematical Programming*, vol. 92, no. 3, pp. 453–480, 2002.

[6] 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.

[7] B.-D. Chen and S. Lall, "Degree bounds for polynomial verification of the matrix cube problem," to be presented at *the 45th IEEE Conference on Decision and Control*, San Diego, USA, December 2006.

[8] M. C. de Oliveira, J. Bernussou, and J. C. Geromel, "A new discrete-time robust stability condition," *Systems & Control Letters*, vol. 37, no. 4, pp. 261–265, 1999.

[9] 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.

[10] 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.

[11] 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.

[12] S. Kim, M. Kojima, and H. Waki, "Generalized Lagrangean duals and sums of squares relaxations of sparse polynomial optimization problems," *SIAM Journal on Optimization*, vol. 15, no. 3, pp. 697–719, 2005.

[13] M. Kojima, "Sums of squares relaxations of polynomial semidefinite programs," Technical Report B-397, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, November 2003.

[14] M. Kojima, S. Kim, and H. Waki, "Sparsity in sums of squares of polynomials," *Mathematical Programming*, vol. 103, no. 1, pp. 45–62, 2005.

[15] M. Kojima and M. Muramatsu, "A note on sparse SOS and SDP relaxations for polynomial optimization problems over symmetric cones," Technical Report B-421, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, January 2006.

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

[17] J. B. Lasserre, "Convergent SDP-relaxations in polynomial optimization with sparsity," *SIAM Journal on Optimization*, vol. 17, no. 3, pp. 822–843, 2006.

[18] J. Löfberg, "YALMIP: a toolbox for modeling and optimization in MATLAB," in *Proceedings of the 2004 IEEE International Symposium on Computer Aided Control Systems Design*, Taipei, Taiwan, September 2004, pp. 284–289, available from http://control.ee.ethz.ch/~joloef/yalmip.php

[19] 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.

[20] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, "Testing unconstrained optimization software," *ACM Transactions on Mathematical Software*, vol. 7, no. 1, pp. 17–41, 1981.

[21] A. Nemirovskii, "Several NP-hard problems arising in robust stability analysis," *Mathematics of Control, Signals, and Systems*, vol. 6, no. 2, pp. 99–105, 1993.

[22] A. Ohara and Y. Sasaki, "On solvability and numerical solutions of parameter-dependent differential matrix inequality," in *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, USA, December 2001, pp. 3593–3594.

[23] Y. Oishi, "A region-dividing approach to robust semidefinite programming and its error bound," Technical Report METR2006-10, Department of Mathematical Informatics, The University of Tokyo, Tokyo, Japan, available from http://www.keisu.t.u-tokyo.ac.jp/Research/techrep.0.html; also in *Proceedings of the 2006 American Control Conference*, Minneapolis, USA, June 2006, pp. 123–129.

[24] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Mathematical Programming*, vol. 96, no. 2, pp. 293–320, 2003.

[25] 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.

[26] D. Peaucelle, D. Arzelier, O. Bachelier, and J. Bernussou, "A new robust D-stability condition for real convex polytopic uncertainty," *Systems & Control Letters*, vol. 40, no. 1, pp. 21–30, 2000.

[27] D. Peaucelle, Y. Ebihara, D. Arzelier, and T. Hagiwara, "General polynomial parameter-dependent Lyapunov functions for polytopic uncertain systems," in *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS2006)*, Kyoto, Japan, July 2006, pp. 2238–2242.

[28] S. K. Rao, P. Sadayappan, F. K. Hwang, and P. W. Shor, "The rectilinear Steiner arborescence problem," *Algorithmica*, vol. 7, nos. 2–3, pp. 277–288, 1992.

[29] M. Sato and D. Peaucelle, "Robust stability/performance analysis for linear time-invariant polynomially parameter-dependent systems using polynomially parameter-dependent Lyapunov functions," to be presented at *the 45th IEEE Conference on Decision and Control*, San Diego, USA, December 2006.

[30] 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.

[31] C. W. Scherer, "LMI relaxations in robust control," *European Journal of Control*, vol. 12, no. 1, pp. 3–29, 2006.

[32] C. W. Scherer and C. W. J. Hol, "Matrix sum-of-squares relaxations for robust semi-definite programs," *Mathematical Programming*, vol. 107, nos. 1–2, pp. 189–211, 2006.

[33] W. Shi and C. Su, "The rectilinear Steiner arborescence problem is NP-complete," *SIAM Journal on Computing*, vol. 35, no. 3, pp. 729–740, 2006.

[34] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vols. 11–12, pp. 625–653, 1999.

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