

## EXPLOITING SPARSITY IN THE MATRIX-DILATION APPROACH TO ROBUST SEMIDEFINITE PROGRAMMING

Yasuaki Oishi  
Nanzan University

Yusuke Isaka  
The University of Tokyo

(Received February 23, 2008; Revised December 16, 2008)

*Abstract* 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 a given problem has only a small number of nonzero terms when it is expressed as a polynomial in 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 accuracy of the approximation can be analyzed quantitatively. In particular, it is shown that the accuracy can be improved to any level by dividing the parameter region into small subregions.

**Keywords:** Optimization, robust semidefinite programming, linear matrix inequalities, matrix dilation, sparsity, rectilinear Steiner arborescences

### 1. Introduction

A *robust semidefinite programming problem* (robust SDP problem, in short) to be considered in this paper is the following:

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

Here,  $c$  is a given vector in  $\mathbb{R}^n$ , which is the set of  $n$ -dimensional real vectors;  $x = (x_1 \ x_2 \ \cdots \ x_n)^T \in \mathbb{R}^n$  is a design variable;  $\theta = (\theta_1 \ \theta_2 \ \cdots \ \theta_p)^T \in \Theta$  is an uncertain parameter, whose domain  $\Theta$  is a closed polytope (*i.e.*, 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  $\theta$ , whose value is an  $m \times m$  real symmetric matrix. The inequalities in the constraint mean that the symmetric matrices  $E(x)$  and  $F(x, \theta)$  are positive semidefinite. Inequalities of this type are called *linear matrix inequalities* (LMIs). The robust SDP problem has many applications in nonlinear optimization and robust control [1, 3, 4, 7, 31].

The robust SDP problem is NP-hard [20] due to the semi-infinite constraint  $F(x, \theta) \succeq O$  ( $\forall \theta \in \Theta$ ). Hence, approximate approaches have been proposed [2, 5, 8, 22, 27, 30, 32]. There, a standard and thus solvable SDP problem is constructed as an approximation of a given robust SDP problem. The matrix-dilation approach is one of such approaches and was proposed by the group of the first author [9, 23] based on the robust control techniques of [18, 19, 21, 36]. An advantage of this approach is availability of an upper bound on the approximation error, that is, the discrepancy between the optimal values of the original

robust SDP problem and its approximate problem. This bound gives better comprehension of the approach and an efficient method for reduction of the approximation error. On the other hand, this approach and others share a drawback that the approximate problem tends to have a large size when applied to practical problems. This may be a natural consequence of the difficulty of the original problem. However, there is still a possibility that a small-sized 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 exploit a sparse structure of  $F(x, \theta)$  in the constraint of  $P$ . Here, a sparse structure means that, when  $F(x, \theta)$  is expanded as a polynomial in  $\theta$ , only a part of the terms is nonzero. Such a structure has a practical application as seen in Example 5.3. This structure is captured with a special graph called a rectilinear Steiner arborescence and is used for construction of a small-sized approximate problem. The effect of the reduction is especially evident when the number of nonzero terms is small and the degrees of the polynomial  $F(x, \theta)$  are high. It is also notable that an upper bound on the approximation error is still available even with the reduction.

Exploitation of sparsity has been considered in the actively investigated sum-of-squares approach [11, 13, 14, 16, 26, 37]. (See [12, 15, 25, 32] for the approach itself.) The techniques used there are, however, quite different from the present one. Comparison with these techniques is a future research subject. Quantitative analysis of the approximation error is rarely made in other approaches to robust SDP problems. Although related analysis is made in [32], it is unknown whether the analysis is useful for reduction of the approximation error or adaptable to the case where sparsity is exploited. Parts of the results of the present paper have been reported in [10, 24].

This paper is organized as follows. In Section 2, the existing matrix-dilation approach is outlined. Section 3 gives a reduced-size approximate problem and Section 4 provides an upper bound on the approximation error. After numerical examples are presented in Section 5, the paper is concluded in Section 6.

The symbol  $\mathbb{Z}_+^p$  stands for the set of  $p$ -dimensional vectors of nonnegative integers. The symbol  $^T$  denotes the transpose of a matrix or a vector. For  $\theta = (\theta_1 \ \theta_2 \ \cdots \ \theta_p)^T \in \mathbb{R}^p$  and  $\alpha = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_p)^T \in \mathbb{Z}_+^p$ , the symbol  $\theta^\alpha$  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  $\bar{\sigma}(A)$ . For a real symmetric matrix  $A$ , the inequality  $A \succeq O$  expresses its positive semidefiniteness while  $A \succ O$  its positive definiteness. 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

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

There hold  $(A \otimes B)^T = A^T \otimes B^T$  and  $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$  for matrices of appropriate size. For a set  $S$ , the symbol  $|S|$  denotes its cardinality. For a minimization problem  $P$ , its optimal value is expressed by  $\inf P$  with an attention that the optimal value may not be attained.

## 2. Matrix-Dilation Approach

In this section, we present the matrix-dilation approach of [9, 23], which is for solving the robust SDP problem  $P$ . The idea is to make a standard SDP problem by replacing the semi-infinite constraint of  $P$  by a stronger constraint.

We need some preparations. In the constraint of  $P$ , we let  $d_i$  be the maximum degree of  $F(x, \theta)$  as a polynomial in  $\theta_i$  for  $i = 1, 2, \dots, p$ . We assume that  $\max_{i=1,2,\dots,p} d_i \geq 1$  without loss of generality. Otherwise,  $F(x, \theta) \succeq O$  is independent of  $\theta$  and the problem  $P$  is easy to solve. With  $V_f := \{\alpha \in \mathbb{Z}_+^p \mid 0 \leq \alpha_i \leq d_i \text{ for } i = 1, 2, \dots, p\}$ , we can write  $F(x, \theta) = \sum_{\alpha \in V_f} F_\alpha(x) \theta^\alpha$ . Here,  $|V_f| = \prod_{i=1}^p (d_i + 1)$ . The subscript ‘‘f’’ stands for the ‘‘full size’’ and is put for distinction from the reduced-size counterpart given in Section 3. We number the elements of  $V_f$  as  $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(|V_f|)}$  with  $\alpha^{(1)}$  being the origin.

We now define

$$\begin{aligned} F_{*f}(x) &:= (F_{\alpha^{(2)}}(x) \quad F_{\alpha^{(3)}}(x) \quad \cdots \quad F_{\alpha^{(|V_f|)}}(x)), \\ G_f(x) &:= \begin{pmatrix} 2F_{\alpha^{(1)}}(x) & F_{*f}(x) \\ F_{*f}(x)^T & O \end{pmatrix}, \\ M_f(\theta) &:= (\theta^{\alpha^{(1)}} I_m \quad \theta^{\alpha^{(2)}} I_m \quad \cdots \quad \theta^{\alpha^{(|V_f|)}} I_m)^T. \end{aligned}$$

An example is found in Example 2.1 below. The matrix  $G_f(x)$  is  $|V_f|m \times |V_f|m$  and the matrix  $M_f(\theta)$  is  $|V_f|m \times m$ . They satisfy  $2F(x, \theta) = M_f(\theta)^T G_f(x) M_f(\theta)$ . Furthermore, we consider a  $|V_f|m \times (|V_f| - 1)m$  matrix  $H_f(\theta)$  such that the square matrix  $(M_f(\theta) \quad H_f(\theta))$  is nonsingular and the relation  $M_f(\theta)^T H_f(\theta) = O$  holds for all  $\theta \in \mathbb{R}^p$ . Such  $H_f(\theta)$  is called an *orthogonal complement* of  $M_f(\theta)$ .

**Example 2.1** We present explicit forms of  $F_{*f}(x)$ ,  $M_f(\theta)$ , and  $H_f(\theta)$  on an example problem considered in [23]: maximization of  $f(\theta) = 9\theta_1\theta_2 - 5\theta_1\theta_2^2 - 5\theta_1^2\theta_2$  in  $\theta \in \Theta = [0, 1]^2$ . Its maximum value 1.08 is attained at  $\theta = (0.6 \ 0.6)^T$ . It is formulated as the robust SDP problem:

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

Hence,  $F(x, \theta) = x - f(\theta) = x - 9\theta_1\theta_2 + 5\theta_1\theta_2^2 + 5\theta_1^2\theta_2$  with  $n = 1$ ,  $\ell = 0$ ,  $m = 1$ ,  $p = 2$ ,  $d_1 = 2$ , and  $d_2 = 2$ . The required matrices are defined as

$$\begin{aligned} F_{\alpha^{(1)}}(x) &= x, & F_{*f}(x) &= (0 \ 0 \ 0 \ -9 \ 5 \ 0 \ 5 \ 0), \\ M_f(\theta) &= \begin{pmatrix} 1 \\ \theta_1 \\ \theta_1^2 \\ \theta_2 \\ \theta_1\theta_2 \\ \theta_1^2\theta_2 \\ \theta_2^2 \\ \theta_1\theta_2^2 \\ \theta_1^2\theta_2^2 \end{pmatrix}, & H_f(\theta) &= \begin{pmatrix} -\theta_1 & & -\theta_2 & & & & & & \\ 1 & -\theta_1 & & -\theta_2 & & & & & \\ & 1 & & -\theta_2 & & & & & \\ & & 1 & & -\theta_2 & & & & \\ & & & 1 & & -\theta_2 & & & \\ & & & & 1 & & -\theta_2 & & \\ & & & & & 1 & & & \\ & & & & & & 1 & & \\ & & & & & & & 1 & \\ & & & & & & & & 1 \end{pmatrix} \end{aligned}$$

for appropriate numbering of  $V_f = \{\alpha \in \mathbb{Z}_+^2 \mid \alpha_1, \alpha_2 = 0, 1, 2\}$ .

The key fact is that the orthogonal complement  $H_f(\theta)$  can be chosen affine in  $\theta$  [23]. This establishes the following fact, which provides a basis for the present approach.

**Lemma 2.1** *Let  $x$  be any point in  $\mathbb{R}^n$  and  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(Q)}$  be any  $Q$  points in  $\mathbb{R}^p$ . Then,  $F(x, \theta) \succeq O$  holds for all  $\theta$  in the convex hull of  $\{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(Q)}\}$  if there exists  $W_f$  satisfying*

$$G_f(x) + H_f(\theta^{(q)})W_f^T + W_fH_f(\theta^{(q)})^T \succeq O \quad (\forall q = 1, 2, \dots, Q). \tag{2.1}$$

*Proof.* Choose any  $\theta$  in the convex hull of  $\{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(Q)}\}$  and express it as a convex combination. The convex combination of (2.1) with the same coefficients gives  $G_f(x) + H_f(\theta)W_f^T + W_fH_f(\theta)^T \succeq O$  for the chosen  $\theta$  because of the affinity of  $H_f(\theta)$ . Premultiplication of  $M_f(\theta)^T$  and postmultiplication of  $M_f(\theta)$  to this inequality provide  $M_f(\theta)^T G_f(x) M_f(\theta) = 2F(x, \theta) \succeq O$ .  $\square$

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

$$\begin{aligned} P_f(\Delta) : \quad & \text{minimize} \quad c^T x \\ & \text{subject to} \quad E(x) \succeq O, \quad G_f(x) + H_f(\theta)(W_f^{[j]})^T + W_f^{[j]}H_f(\theta)^T \succeq O \\ & \quad (\forall \theta \in \text{ver } \Theta^{[j]}; \quad \forall j = 1, 2, \dots, J), \end{aligned}$$

where the design variables are  $x \in \mathbb{R}^n$  and the  $|V_f| m \times (|V_f| - 1)m$  matrices  $W_f^{[1]}, W_f^{[2]}, \dots, W_f^{[J]}$ . This problem  $P_f(\Delta)$  is a standard SDP problem, which can be solved with the interior-point method. From Lemma 2.1, a relationship between this problem and the original problem can be established, which is stated in the next proposition. Here, noting that the feasible region of  $P_f(\Delta)$  is in the space of  $(x, \{W_f^{[j]}\}_{j=1}^J)$ , we mean its projection to the  $x$ -space by the *projected feasible region* of  $P_f(\Delta)$ .

**Proposition 2.1** *Let  $\Delta$  be a division of  $\Theta$ . Then, the projected feasible region of the approximate problem  $P_f(\Delta)$  is contained in the feasible region of the original problem  $P$ . In particular,  $\inf P \leq \inf P_f(\Delta)$ .*

In the matrix-dilation approach, we solve the approximate problem  $P_f(\Delta)$  in place of the original problem  $P$ . Although there is a nonzero *approximation error*  $\inf P_f(\Delta) - \inf P$  in general, we can make it smaller by subdividing the polytopes in  $\Delta$  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  $\Theta^{[j]}$  is  $\text{rad } \Theta^{[j]} := \min_{\theta \in \Theta^{[j]}} \max_{\theta' \in \Theta^{[j]}} \max_{i=1,2,\dots,p} |\theta_i - \theta'_i|$ , where a  $\theta$  that attains the minimum is called a *center* of  $\Theta^{[j]}$ . The *maximum radius* of a division  $\Delta = \{\Theta^{[j]}\}_{j=1}^J$  is defined as  $\overline{\text{rad}} \Delta := \max_{j=1,2,\dots,J} \text{rad } \Theta^{[j]}$ .

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. One major reason is that  $|V_f| = \prod_{i=1}^p (d_i + 1)$  increases rapidly as the degrees  $d_i$  increase.

In the next section, we assume a kind of sparsity in the given problem  $P$  and reduce the sizes of  $G_f(x)$ ,  $M_f(\theta)$ , and  $H_f(\theta)$ . This leads us to a reduced-size approximate problem. Even with this improvement, an upper bound on the approximation error remains available, whose derivation is a task of Section 4.

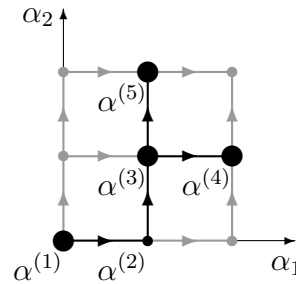


Figure 1: The directed graph  $(V_f, A_f)$  (gray) in the case of  $p = 2$ ,  $d_1 = 2$ , and  $d_2 = 2$  together with a rectilinear Steiner arborescence  $(V, A)$  (black) for the set  $S$  whose elements are shown by the large circles

### 3. A Reduced-Size Approximate Problem

In this section, we construct a reduced-size approximate problem. The idea is to exploit sparsity in the given robust SDP problem  $P$  and to consider matrices smaller in size than  $G_f(x)$ ,  $M_f(\theta)$ , and  $H_f(\theta)$ , which will be denoted by  $G(x)$ ,  $M(\theta)$ , and  $H(\theta)$  without the subscript “f.” Let us review the requirements on these matrices:  $2F(x, \theta)$  has to be equal to  $M(\theta)^T G(x) M(\theta)$ ;  $H(\theta)$  has to be an orthogonal complement of  $M(\theta)$  and be affine in  $\theta$ .

Recall that  $F(x, \theta) = \sum_{\alpha \in V_f} F_\alpha(x) \theta^\alpha$ . We assume a kind of sparsity that the coefficient matrices  $F_\alpha(x)$  are nonzero for a small number of  $\alpha$ 's. We let  $S$  be the *support* of  $F(x, \theta)$  defined as  $S := \{\alpha \in V_f \subset \mathbb{Z}_+^p \mid F_\alpha(x) \not\equiv O\}$ . Since  $F(x, \theta)$  is not independent of  $\theta$ , the support  $S$  contains at least one element other than the origin.

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

For the support  $S$ , we consider a subgraph  $(V, A)$  of  $(V_f, A_f)$  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  $\alpha$  is said to be *reachable* from a vertex  $\beta$  if either  $\alpha = \beta$  or there is a path connecting  $\beta$  to  $\alpha$  through the arcs in the directed way. Such a graph  $(V, A)$  is referred to as a *rectilinear Steiner arborescence* for  $S$  (see [29] and the references therein). For a given  $S$ , a rectilinear Steiner arborescence always exists though not unique. Figure 1 gives an example in the case of Example 2.1, where  $p = 2$ ,  $d_1 = 2$ , and  $d_2 = 2$ . There, the directed graph  $(V_f, A_f)$  is presented in gray. Since  $F(x, \theta) = x - 9\theta_1\theta_2 + 5\theta_1\theta_2^2 + 5\theta_1^2\theta_2$ , the support  $S$  is  $\{(0 \ 0)^T, (1 \ 1)^T, (1 \ 2)^T, (2 \ 1)^T\}$ , whose elements are expressed by the large circles. A rectilinear Steiner arborescence for this  $S$  is shown in black.

With these preparations, we now construct the desired matrices  $G(x)$ ,  $M(\theta)$ , and  $H(\theta)$ . Let the vertices in  $V$  be  $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(|V|)}$ . 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  $q \leq r$ . This is consistent with  $\alpha^{(1)}$  being the origin. Figure 1 is an example of this numbering. With this notation, we define

$$F_*(x) := \begin{pmatrix} F_{\alpha^{(2)}}(x) & F_{\alpha^{(3)}}(x) & \cdots & F_{\alpha^{(|V|)}}(x) \end{pmatrix},$$

$$G(x) := \begin{pmatrix} 2F_{\alpha^{(1)}}(x) & F_*(x) \\ F_*(x)^T & O \end{pmatrix}, \quad (3.1)$$

$$M(\theta) := (\theta^{\alpha^{(1)}} I_m \quad \theta^{\alpha^{(2)}} I_m \quad \cdots \quad \theta^{\alpha^{(|V|)}} I_m)^T.$$

See Example 3.1 below for an example. It is easy to see that  $2F(x, \theta) = M(\theta)^T G(x) M(\theta)$ . We next define

$$H(\theta) := \tilde{H}(\theta) \otimes I_m \quad (3.2)$$

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

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

for  $q = 1, 2, \dots, |V|$  and  $r = 1, 2, \dots, |V| - 1$ . Obviously,  $H(\theta)$  is affine in  $\theta$ . The size of  $G(x)$  is  $|V|m \times |V|m$ , that of  $M(\theta)$  is  $|V|m \times m$ , and that of  $H(\theta)$  is  $|V|m \times (|V| - 1)m$ . Since  $|V| \leq |V_f|$ , the matrices  $G(x)$ ,  $M(\theta)$ , and  $H(\theta)$  have the sizes smaller than or equal to those of  $G_f(x)$ ,  $M_f(\theta)$ , and  $H_f(\theta)$ .

**Example 3.1** We consider the problem of Example 2.1. With the rectilinear Steiner arborescence and the numbering of Figure 1, we obtain the matrices

$$F_{\alpha^{(1)}}(x) = x, \quad F_*(x) = (0 \quad -9 \quad 5 \quad 5),$$

$$M(\theta) = \begin{pmatrix} 1 \\ \theta_1 \\ \theta_1 \theta_2 \\ \theta_1^2 \theta_2 \\ \theta_1 \theta_2^2 \end{pmatrix}, \quad H(\theta) = \begin{pmatrix} -\theta_1 & & & \\ 1 & -\theta_2 & & \\ & 1 & -\theta_1 & -\theta_2 \\ & & 1 & \\ & & & 1 \end{pmatrix}.$$

These matrices are smaller in size than their counterparts in Example 2.1. This is a consequence that the rectilinear Steiner arborescence has only  $|V| = 5$  vertices while the original graph has  $|V_f| = 9$ .

The matrix  $H(\theta)$  in (3.2) has the following required property.

**Lemma 3.1** *The matrix  $H(\theta)$  is an orthogonal complement of  $M(\theta)$ .*

*Proof.* By the definition, the matrix  $\tilde{H}(\theta)$  is upper triangular in the sense that  $\tilde{H}(\theta)_{qr} \neq 0$  only if  $q \leq r + 1$ . Moreover,  $\tilde{H}(\theta)_{qr} = 1$  if  $q = r + 1$ . Hence, the matrix  $\tilde{H}(\theta)$  is of column full rank and, thus, so is  $H(\theta)$ . Since  $M(\theta)$  is clearly of column full rank, the nonsingularity of  $(M(\theta) \quad H(\theta))$  follows when the orthogonality  $M(\theta)^T H(\theta) = O$  holds. Let us show the orthogonality. Write  $M(\theta) = \tilde{M}(\theta) \otimes I_m$  with  $\tilde{M}(\theta) = (\theta^{\alpha^{(1)}} \quad \theta^{\alpha^{(2)}} \quad \cdots \quad \theta^{\alpha^{(|V|)}})^T$ . Noting that  $M(\theta)^T H(\theta) = [\tilde{M}(\theta)^T \tilde{H}(\theta)] \otimes I_m$ , we consider the product between  $\tilde{M}(\theta)^T$  and the  $r$ th column vector of  $\tilde{H}(\theta)$ . 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  $\hat{q}$  and suppose that the arc  $(\alpha^{(\hat{q})}, \alpha^{(r+1)})$  is parallel to the  $i$ th axis. Then,  $\tilde{H}(\theta)_{qr}$  is equal to  $-\theta_i$  for  $q = \hat{q}$ , equal to unity for  $q = r + 1$ , and equal to zero otherwise. Hence, the considered product is equal to  $\theta^{\alpha^{(\hat{q})}}(-\theta_i) + \theta^{\alpha^{(r+1)}}$ , which is equal to zero. Repeating this reasoning for each  $r$ , we arrive at the desired orthogonality.  $\square$

Because the matrices  $G(x)$ ,  $M(\theta)$ , and  $H(\theta)$  have the required properties, they can be used in the matrix-dilation approach. Namely, with  $\Delta = \{\Theta^{[j]}\}_{j=1}^J$  being a division of  $\Theta$ , we consider a reduced-size approximate problem:

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

Then, Proposition 2.1 holds on this problem with necessary modification. As is mentioned in Section 2, we can decrease the approximation error by subdividing  $\Delta$ .

We close this section by discussing when the discrepancy between  $|V|$  and  $|V_f|$  becomes large.

**Proposition 3.1** *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  $\alpha = (\alpha_1 \ \alpha_2 \ \dots \ \alpha_p)^T \in S$ , consider a directed path connecting  $(0 \ 0 \ \dots \ 0)^T, (1 \ 0 \ \dots \ 0)^T, \dots, (\alpha_1 \ 0 \ \dots \ 0)^T, (\alpha_1 \ 1 \ \dots \ 0)^T, \dots, (\alpha_1 \ \alpha_2 \ \dots \ 0)^T, \dots, (\alpha_1 \ \alpha_2 \ \dots \ \alpha_p)^T$  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  $|V| \leq |S| \sum_{i=1}^p d_i + 1$ .  $\square$

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

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 heuristic algorithm for this purpose is given in [29].

**Remark 3.1** Suppose that the support  $S$  is equal to the whole vertex set  $V_f$ . If we construct a rectilinear Steiner arborescence for this  $S$  as the proof of Proposition 3.1, we obtain the full-size matrices  $G_f(x)$ ,  $M_f(\theta)$ , and  $H_f(\theta)$ , which were considered in [23]. In this sense, the present approach is a generalization of the conventional approach in [23].

**Remark 3.2** The optimal value of the reduced-size approximate problem may be different from that of the full-size approximate problem. Although no quantitative result is known on this effect, 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  $P(\Delta)$  to know its quantitative properties.

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

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

where  $\epsilon$  is a nonnegative number. Note that  $P_0$  is identical with the original problem  $P$ . Moreover,  $\inf P_\epsilon$  is convex as a function in  $\epsilon$ . To see this, suppose that  $x_1$  is feasible in  $P_{\epsilon_1}$  and so is  $x_2$  in  $P_{\epsilon_2}$ . Then,  $(x_1 + x_2)/2$  is feasible in  $P_{(\epsilon_1 + \epsilon_2)/2}$ , which implies  $(\inf P_{\epsilon_1} + \inf P_{\epsilon_2})/2 \geq \inf P_{(\epsilon_1 + \epsilon_2)/2}$ . We will show below that  $\epsilon$  can be chosen so that  $\inf P \leq \inf P(\Delta) \leq \inf P_\epsilon$ . We then have  $0 \leq \inf P_\epsilon - \inf P \leq g\epsilon$  with  $g$  being an upper bound on a subgradient of

$\inf P_\epsilon$ . This implies  $0 \leq \inf P(\Delta) - \inf P \leq g\epsilon$  to give the desired upper bound on the approximation error. This basic idea is the same as the full-size case [23]. A difference is that the technique used in [23] for the establishment of  $\inf P \leq \inf P_f(\Delta) \leq \inf P_\epsilon$  is not applicable to the reduced-size approximate problem  $P(\Delta)$  because the whole set of  $F_\alpha(x)$ ,  $\alpha \in V_f$ , is not used this time. We hence introduce a different technique with a coordinate transformation matrix  $L(\theta)$ .

We need the following assumption to have the result. A corresponding assumption is common in the context of the interior-point method for linear programming [28, Section 5.7].

**Assumption 4.1**

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

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

**Theorem 4.1** *Suppose that Assumption 4.1 holds. Then, the reduced-size approximate problem  $P(\Delta)$  satisfies*

$$0 \leq \inf P(\Delta) - \inf P \leq C \overline{\text{rad}} \Delta. \quad (4.1)$$

for any division  $\Delta$  with  $\overline{\text{rad}} \Delta \leq C'$ , where  $C$  and  $C'$  are positive numbers independent of  $\Delta$ .

The specific forms of  $C$  and  $C'$  will be given in (4.6) and (4.7), respectively.

The bound (4.1) reveals quantitative properties of our matrix-dilation approach. First, it implies the asymptotic exactness. Namely, the approximation error of the reduced-size approximate problem converges to zero as the maximum radius of the division goes to zero. The bound also gives the convergence rate, which is at most a linear order of the maximum radius. Evaluation of  $C$  and  $C'$  is possible as in [23], though often conservative.

The bound (4.1) is useful for complexity analysis of the approach. For reduction of the approximation error, decrease of the maximum radius is required, which results in increase of the number of subregions in the order of  $O((\overline{\text{rad}} \Delta)^{-p})$ . Especially when the parameter dimension  $p$  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. In fact, the maximum radius in (4.1) can be replaced by the *maximum active radius*, which is the maximum radius only over the subregions having active constraints. This means that we do not need to make a uniform division but need to divide only an important area. This idea leads us to adaptive construction of an efficient division. The discussion is parallel to that in [23]. The details are omitted.

In the rest of this section, we prove Theorem 4.1. We first prepare the  $|V|m \times |V|m$  matrix  $L(\theta)$  for simplification of the dilated LMI constraint  $G(x) + H(\theta)(W^{[j]})^T + W^{[j]}H(\theta)^T \succeq O$ . This matrix is defined as

$$L(\theta) := \tilde{L}(\theta) \otimes I_m \quad (4.2)$$

with

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

The matrix  $\tilde{L}(\theta)$  is lower triangular, *i.e.*,  $\tilde{L}(\theta)_{qr} \neq 0$  only if  $q \geq r$  due to the numbering consistent with  $(V, A)$ . Moreover, its diagonal elements are all equal to unity. A consequence



is nonsingularity of  $\tilde{L}(\theta)$  and also of  $L(\theta)$ . For example,

$$L(\theta) = \begin{pmatrix} 1 & & & & \\ \theta_1 & 1 & & & \\ \theta_1\theta_2 & \theta_2 & 1 & & \\ \theta_1^2\theta_2 & \theta_1\theta_2 & \theta_1 & 1 & \\ \theta_1\theta_2^2 & \theta_2^2 & \theta_2 & & 1 \end{pmatrix}$$

in the case of Example 3.1.

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

**Lemma 4.1** For the matrices  $G(x)$  and  $L(\theta)$  in (3.1) and (4.2), respectively, we can write

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

with

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

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.  $\square$

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

**Lemma 4.2** For the matrices  $H(\theta)$  and  $L(\theta)$  in (3.2) and (4.2), respectively, we can write

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

where the elements expressed by  $*$  are 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, all the elements expressed by  $*$  are equal to zero.

*Proof.* Noting that  $L(\theta)^T H(\theta') = [\tilde{L}(\theta)^T \tilde{H}(\theta')] \otimes I_m$ , we will evaluate the  $(s, r)$ -element of  $\tilde{L}(\theta)^T \tilde{H}(\theta')$ , which is the inner product between the  $s$ th column of  $\tilde{L}(\theta)$  and the  $r$ th column of  $\tilde{H}(\theta')$ , where  $s = 1, 2, \dots, |V|$  and  $r = 1, 2, \dots, |V| - 1$ . As in the proof of Lemma 3.1, the element  $\tilde{H}(\theta')_{qr}$  is equal to  $-\theta'_i$  if  $q = \hat{q}$ , equal to unity if  $q = r + 1$ , and equal to zero otherwise. Here, the unique arc ending at  $\alpha^{(r+1)}$  is supposed to be  $(\alpha^{(\hat{q})}, \alpha^{(r+1)})$  and be parallel to the  $i$ th axis. The value of the considered  $(s, r)$ -element depends on the reachability of the vertices  $\alpha^{(\hat{q})}$  and  $\alpha^{(r+1)}$  from  $\alpha^{(s)}$  in  $(V, A)$ . Namely, there are three possible cases.

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

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

Note that  $\alpha^{(\hat{q})} - \alpha^{(s)} \in \mathbb{Z}_+^p$ .

(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  $\alpha^{(\hat{q})}$  nor  $\alpha^{(r+1)}$  is reachable from  $\alpha^{(s)}$ . In this case, the quantity is equal to zero.

Summarizing these results, we obtain the lemma. □

With these preparations, we relate  $P(\Delta)$  and  $P_\epsilon$ . By Assumption 4.1 (a), there exists  $\epsilon_0 > 0$  such that, for any  $0 \leq \epsilon \leq \epsilon_0$ , the auxiliary problem  $P_\epsilon$  is strictly feasible. Let  $v_0$  be a number with  $\inf P_{\epsilon_0} \leq v_0$  and define the level set

$$X := \{x \in \mathbb{R}^n \mid c^T x \leq v_0, E(x) \succeq O, \text{ and } F(x, \theta) \succeq O (\forall \theta \in \Theta)\}.$$

This set is nonempty and bounded by Assumption 4.1 (b). For each  $0 \leq \epsilon \leq \epsilon_0$ , the infimum of  $P_\epsilon$  is attained in  $X$  and only in  $X$ .

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

$$\max_{x \in X} \max_{\theta \in \Theta} \bar{\sigma}[F_{**}(x, \theta)] \leq \bar{F}$$

with the maximum singular value  $\bar{\sigma}$ .

**Lemma 4.3** *Suppose that  $\Theta \subseteq [-1, 1]^p$  and*

$$\overline{\text{rad}} \Delta \leq \min \left\{ \frac{2\epsilon_0}{(\bar{F} + \sqrt{|V|m})^2}, \frac{1}{|V|} \right\}.$$

*Then, we have  $\inf P \leq \inf P(\Delta) \leq \inf P_\epsilon$  for*

$$\epsilon = \frac{(\bar{F} + \sqrt{|V|m})^2 \overline{\text{rad}} \Delta}{2}.$$

*Proof.* Since  $\inf P \leq \inf P(\Delta)$  as noticed in Section 3, we will show  $\inf P(\Delta) \leq \inf P_\epsilon$ . Let  $\Delta$  be  $\{\Theta^{[j]}\}_{j=1}^J$ . The  $\epsilon$  given in the lemma satisfies  $0 \leq \epsilon \leq \epsilon_0$ . Hence, the infimum of  $P_\epsilon$  is attained at a point in  $X$ . Let such a minimizing point be  $x^* \in X$ . 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  $j = 1, 2, \dots, J$  such that

$$G(x^*) + H(\theta)(W^{[j]})^T + W^{[j]}H(\theta)^T \succeq O \quad (\forall \theta \in \text{ver } \Theta^{[j]}). \tag{4.3}$$

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

In order to show the inequality (4.3) for a fixed  $j$ , we premultiply  $L(\theta^c)^T$  and postmultiply  $L(\theta^c)$  to its left-hand-side matrix. Lemmas 4.1 and 4.2 give the concrete forms of  $L(\theta^c)^T G(x^*) L(\theta^c)$  and  $(W^{[j]})^T L(\theta^c) = (1/\text{rad } \Theta^{[j]})H(\theta^c)^T L(\theta^c)$ . By Lemma 4.2 again, the product  $L(\theta^c)^T H(\theta)$  divided by  $\text{rad } \Theta^{[j]}$  has the form

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

where the elements expressed by  $*$  are either equal to zero or of the form

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

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

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

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

$$H_2 + H_2^\text{T} \succeq \left( \frac{2}{\text{rad } \Theta^{[j]}} - |V| + 2 \right) I.$$

Now, the left-hand side of (4.3) multiplied by  $L(\theta^c)^\text{T}$  and  $L(\theta^c)$  has the upper-left  $m \times m$  block equal to  $2F(x^*, \theta^c)$ , which satisfies  $2F(x^*, \theta^c) \succeq 2\epsilon I$  due to the feasibility of  $x^*$  in  $P_\epsilon$ . The Schur complement of this upper-left block is

$$\begin{aligned} & H_2 + H_2^\text{T} - [F_{**}(x^*, \theta^c) + H_1]^\text{T} [2F(x^*, \theta^c)]^{-1} [F_{**}(x^*, \theta^c) + H_1] \\ & \succeq \left( \frac{2}{\text{rad } \Theta^{[j]}} - |V| + 2 \right) I - \{ \bar{\sigma}[F_{**}(x^*, \theta^c)] + \sqrt{(|V| - 1)m} \}^2 \frac{1}{2\epsilon} I. \end{aligned}$$

Noting that  $2/\text{rad } \Theta^{[j]} - |V| + 2 \geq 1/\text{rad } \Theta^{[j]}$  and  $\bar{\sigma}[F_{**}(x^*, \theta^c)] \leq \bar{F}$ , we see the positive semidefiniteness of the right-hand-side matrix. This completes the proof.  $\square$

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

$$\bar{\theta} := \max\{1, \max_{\theta \in \Theta} \max_{i=1,2,\dots,p} |\theta_i|\} \quad (4.4)$$

and  $\langle \alpha \rangle := \alpha_1 + \alpha_2 + \dots + \alpha_p$ . Since

$$F(x, \theta) = \sum_{\alpha \in V} F_\alpha(x) \theta^\alpha = \sum_{\alpha \in V} F_\alpha(x) \bar{\theta}^{\langle \alpha \rangle} \left( \frac{\theta}{\bar{\theta}} \right)^\alpha,$$

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

$$\begin{aligned} F'_{**}(x, \theta) := & \left( \sum_{\alpha \in V^{(2)}} F_\alpha(x) \bar{\theta}^{\langle \alpha \rangle} \left( \frac{\theta}{\bar{\theta}} \right)^{\alpha - \alpha^{(2)}} \sum_{\alpha \in V^{(3)}} F_\alpha(x) \bar{\theta}^{\langle \alpha \rangle} \left( \frac{\theta}{\bar{\theta}} \right)^{\alpha - \alpha^{(3)}} \dots \right. \\ & \left. \sum_{\alpha \in V^{(|V|)}} F_\alpha(x) \bar{\theta}^{\langle \alpha \rangle} \left( \frac{\theta}{\bar{\theta}} \right)^{\alpha - \alpha^{(|V|)}} \right) \end{aligned}$$

and

$$\bar{F}' := \max_{x \in X} \max_{\theta \in \Theta} \bar{\sigma}[F'_{**}(x, \theta)]. \quad (4.5)$$

**Lemma 4.4** Suppose that

$$\overline{\text{rad}} \Delta \leq \min \left\{ \frac{2\bar{\theta}\epsilon_0}{(\bar{F}' + \sqrt{|V|m})^2}, \frac{\bar{\theta}}{|V|} \right\}.$$

Then, we have  $\inf P \leq \inf P(\Delta) \leq \inf P_\epsilon$  for

$$\epsilon = \frac{(\bar{F}' + \sqrt{|V|m})^2 \overline{\text{rad}} \Delta}{2\bar{\theta}}.$$

*Proof.* We observed that the replacement of  $F_\alpha(x)$  by  $F_\alpha(x)\bar{\theta}^{(\alpha)}$  and  $\theta$  by  $\theta/\bar{\theta}$  does not change the problems  $P$  and  $P_\epsilon$ . As is shown below, this replacement does not change either the approximate problem  $P(\Delta)$ . The parameter after the replacement, *i.e.*,  $\theta/\bar{\theta}$ , moves in  $[-1, 1]^p$ . Hence, the result of Lemma 4.3 is valid with  $\bar{F}$  replaced by  $\bar{F}'$  and  $\overline{\text{rad}} \Delta$  by  $\overline{\text{rad}} \Delta/\bar{\theta}$ . This completes the proof.

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

$$\begin{aligned} G'(x) &= \text{diag}\{I_m, T\}G(x)\text{diag}\{I_m, T\}, \\ H'(\theta) &= \text{diag}\{I_m, T\}H(\theta)T^{-1}, \end{aligned}$$

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

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

is equivalent to the existence of  $W'$  satisfying

$$G'(x) + H'(\theta)(W')^T + W'H'(\theta)^T \succeq O$$

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

We now take the final step for the proof of Theorem 4.1. We assume Assumption 4.1. The number  $\epsilon_0$  is chosen so that the auxiliary problem  $P_\epsilon$  is strictly feasible for any  $0 \leq \epsilon \leq \epsilon_0$ . The numbers  $\bar{\theta}$  and  $\bar{F}'$  are as in (4.4) and (4.5), respectively. Finally, let  $g$  be an upper bound on a subgradient of  $\inf P_\epsilon$  at  $\epsilon = \epsilon_0$ . Then, with

$$C = \frac{g(\bar{F}' + \sqrt{|V|m})^2}{2\bar{\theta}}, \quad (4.6)$$

$$C' = \min \left\{ \frac{2\bar{\theta}\epsilon_0}{(\bar{F}' + \sqrt{|V|m})^2}, \frac{\bar{\theta}}{|V|} \right\}, \quad (4.7)$$

we can prove the theorem.

*Proof of Theorem 4.1.* Lemma 4.4 implies that, when  $\overline{\text{rad}} \Delta \leq C'$ , we have  $\inf P \leq \inf P(\Delta) \leq \inf P_\epsilon$  for  $\epsilon = [(\bar{F}' + \sqrt{|V|m})^2/2\bar{\theta}]\overline{\text{rad}} \Delta$ . Owing to the convexity of  $\inf P_\epsilon$ , the number  $g$  serves as an upper bound on a subgradient of  $\inf P_\epsilon$  also at this  $\epsilon$ . Hence, the convexity of  $\inf P_\epsilon$  again implies  $\inf P \geq \inf P_\epsilon - g\epsilon$ , from which  $0 \leq \inf P(\Delta) - \inf P \leq \inf P_\epsilon - \inf P \leq g\epsilon$ . Substitution of the concrete form of  $\epsilon$  gives the theorem.  $\square$

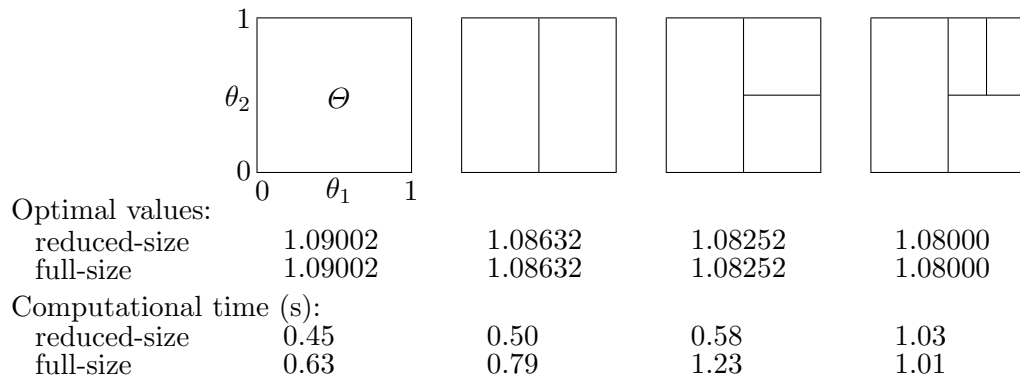


Figure 2: Comparison between the reduced-size and the full-size approximate problems in maximization of  $f(\theta) = 9\theta_1\theta_2 - 5\theta_1\theta_2^2 - 5\theta_1^2\theta_2$  ( $\theta \in [0, 1]^2$ ). Four different divisions were used.

## 5. Numerical Examples

We compare the proposed reduced-size approximate problem with the full-size approximate problem. Three numerical examples show the superiority of the former.

**Example 5.1** This is a continuation of Examples 2.1 and 3.1, where the maximization of  $f(\theta) = 9\theta_1\theta_2 - 5\theta_1\theta_2^2 - 5\theta_1^2\theta_2$  was considered in  $\theta \in \Theta = [0, 1]^2$ . We constructed and solved the reduced-size and the full-size approximate problems with the coarsest division  $\Delta = \{\Theta\}$ . The problem size measured by the number of rows and columns of  $G(x)$  was  $|V|m = 5$  in the reduced-size approximate problem while  $|V_f|m = 9$  in the full-size approximate problem. Solving the reduced-size approximate problem, we obtained the optimal value 1.08000, which was very close to the exact value 1.08. The computational time was 0.43 seconds. With the full-size approximate problem, on the other hand, we obtained a worse approximation 1.09002 after longer computational time 0.63 seconds. The reduced-size approximate problem was hence more useful. For the computation, the SDP solver SeDuMi [34] was used with the modeling language YALMIP [17]. The computer was equipped with Pentium 4 of 2.4 GHz and 2 GByte memory.

The results may change with a different rectilinear Steiner arborescence. We next used the arborescence constructed as the proof of Proposition 3.1. In this case, the size of the reduced-size approximate problem was  $|V|m = 6$ . We used the four different divisions of  $\Theta$  adaptively constructed in [23]. The results are summarized in Figure 2. The obtained optimal values showed almost no difference between the reduced-size and the full-size approximate problems. As the resolution became higher, they approached the exact value 1.08. On the other hand, the computational time was generally smaller in the reduced-size approximate problem. In this case, we obtained  $C = 102.038$  and  $C' = 0.166667$ . Since the used divisions have larger maximum radii (and also larger maximum active radii) than this  $C'$ , the upper bound in Theorem 4.1 is not directly applicable. We can see however that the actual approximation error is much smaller than the theoretical upper bound  $C \overline{\text{rad}} \Delta$ .

**Example 5.2** We next consider maximization of two-dimensional random polynomials with the degrees  $d_1 = d_2 = \mu$ , where  $\mu$  is an integer not less than three. Each polynomial has ten nonzero terms: one is the maximum-degree term  $\theta_1^\mu \theta_2^\mu$  and the remaining nine are randomly chosen from  $\{\theta_1^{\alpha_1} \theta_2^{\alpha_2} \mid \alpha_1, \alpha_2 = 0, 1, \dots, \mu, (\alpha_1, \alpha_2) \neq (\mu, \mu)\}$ . The coefficient of each term is randomly chosen from  $[-1, 1]$ . For each of  $\mu = 3, 4, \dots, 8$ , we generated 100 random polynomials and computed their maxima over  $\theta \in \Theta = [0, 1]^2$  with the reduced-size and the

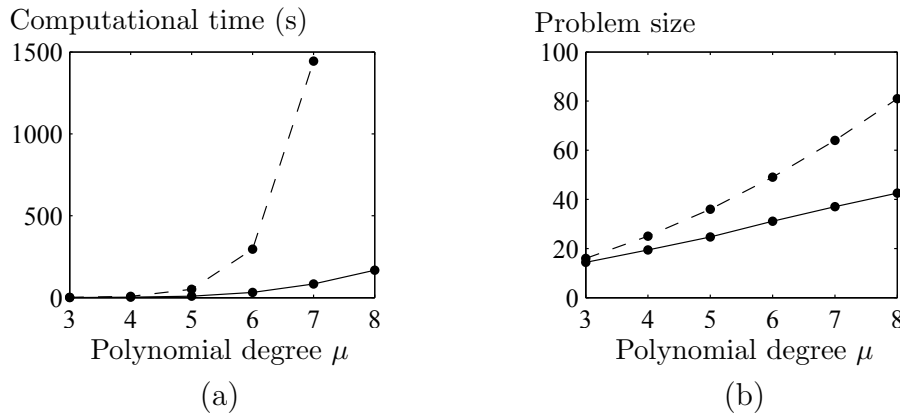


Figure 3: Comparison between the reduced-size (the solid lines) and the full-size (the broken lines) approximate problems in maximization of random polynomials

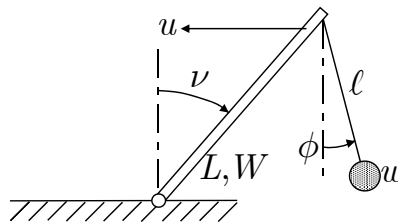


Figure 4: A crane model, for which a controller was designed

full-size approximate problems. The used division was the coarsest one  $\Delta = \{\Theta\}$ . As  $\mu$  increases, the degrees  $d_1$  and  $d_2$  grow up while the size of the support  $|S|$  does not. Hence, the superiority of the reduced-size approximate problem should become clearer according to Proposition 3.1.

The averaged computational time is presented in Figure 3 (a). The environment of the computation was the same as the previous example. We can see in the figure that the reduced-size approximate problem had much smaller computational time than the full-size approximate problem. Indeed, for  $\mu = 8$ , the full-size approximate problems could not be solved due to the crash of the solver. Figure 3 (b) shows the averaged sizes of the approximate problems measured by  $|V|/m$  or  $|V_f|/m$ . The reduced-size approximate problem had a smaller size, which explains its small computational time.

The accuracy of approximation was reasonably good if the coarseness of the division is taken into account. Indeed, the obtained value, which was an upper bound on the true maximum, differed from a lower bound by 0.01 or less for 80.3% of the polynomials and by  $10^{-6}$  or less for 53.0% of the polynomials. Here, a lower bound was computed by evaluating a polynomial at 2500 grid points in  $\Theta$  and taking the maximum. The averaged value of the upper bound was 0.945 while that of the lower bound was 0.932.

**Example 5.3** As the last example, we consider a more realistic problem: design of a state-feedback controller for an uncertain system. Figure 4 shows a schematic view of the system to be considered, which is a simplification of a crane model investigated in [35]. Here,  $g$  denotes the acceleration of gravity,  $9.8 \text{ m/s}^2$ ;  $L$  the length of the boom,  $0.71 \text{ m}$ ;  $W$  the weight of the boom,  $0.205 \text{ kg}$ ;  $w$  the weight of the load,  $0.3 \text{ kg}$ ;  $l$  the length of the rope, which can take any value between  $1 \text{ m}$  and  $1.5 \text{ m}$ . We want to stabilize the system at a given angle of the boom,  $\nu_0$ , which can take any value between  $40$  degrees and  $50$  degrees. The dynamics

of the system linearized at  $\nu = \nu_0$  and  $\phi = 0$ , which is the angle of the rope, is

$$\frac{d}{dt} \begin{pmatrix} \nu(t) - \nu_0 \\ \phi(t) \\ \dot{\nu}(t) \\ \dot{\phi}(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{g(W/2+w)}{La \cos \nu_0} & \frac{gw \cos \nu_0}{La} & 0 & 0 \\ -\frac{g(W/2+w)}{\ell a} & -\frac{g(W/3+w)}{\ell a} & 0 & 0 \end{pmatrix} \begin{pmatrix} \nu(t) - \nu_0 \\ \phi(t) \\ \dot{\nu}(t) \\ \dot{\phi}(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -\frac{\cos \nu_0}{La} \\ \frac{\cos^2 \nu_0}{\ell a} \end{pmatrix} (u(t) - u_0).$$

Here,  $a$  is  $W/3 + w - w \cos^2 \nu_0$  and  $u_0$  is the force to make  $\nu = \nu_0$  and  $\theta = 0$  equilibrium. We write this equation as  $(d/dt)\xi(t) = A(\nu_0, \ell)\xi(t) + B(\nu_0, \ell)\tilde{u}(t)$  making explicit the dependence on the uncertain parameters  $\nu_0$  and  $\ell$ .

We design a state-feedback controller  $\tilde{u}(t) = K\xi(t)$  to stabilize this system for all the possible values of  $\nu_0$  and  $\ell$ . This type of problems is called robust control. A robust control problem has been difficult to solve when a given system has nonlinear dependence on uncertain parameters like our problem. In that case, it has been common to adopt conservative design. We tackle our robust control problem by formulating it into an optimization problem:

$$\begin{aligned} & \text{minimize} && x \\ & \text{subject to} && I - Y \succeq O, \quad \begin{pmatrix} 1 & Z \\ Z^T & I \end{pmatrix} \succeq O, \quad Y + xI \succeq O, \\ & && -A(\nu_0, \ell)Y - B(\nu_0, \ell)Z - YA(\nu_0, \ell)^T - Z^TB(\nu_0, \ell)^T + xI \succeq O \\ & && \text{(for all possible values of } \nu_0 \text{ and } \ell). \end{aligned}$$

If the optimal  $x$  is negative, the optimal  $Y$  and  $Z$  give a robustly stabilizing controller by  $K = ZY^{-1}$  [3, 6].

To write this problem in the form of  $P$ , multiply  $a \cos \nu_0$  to both sides of the last constraint and set  $\theta_1 = \cos \nu_0$  and  $\theta_2 = 1/\ell$ . The resulting constraint is a polynomial in the parameters  $\theta_1$  and  $\theta_2$ , whose degrees are  $d_1 = 3$  and  $d_2 = 1$ . Although the number of possible terms is  $(d_1 + 1)(d_2 + 1) = 8$ , two terms  $\theta_2$  and  $\theta_1^2\theta_2$  have zero coefficients. In this sense, our problem has sparsity. The parameter set  $\Theta$  is written as  $[\underline{\theta}_1, \bar{\theta}_1] \times [\underline{\theta}_2, \bar{\theta}_2]$ , where  $\underline{\theta}_i$  and  $\bar{\theta}_i$  ( $i = 1, 2$ ) are the possible minimum and maximum values of the parameters, respectively.

We solved this problem with the reduced-size and the full-size approximate problems. We chose the division  $\Delta$  to consist of  $[\underline{\theta}_1, \bar{\theta}_1] \times [\underline{\theta}_2, (\underline{\theta}_2 + \bar{\theta}_2)/2]$  and  $[\underline{\theta}_1, \bar{\theta}_1] \times [(\underline{\theta}_2 + \bar{\theta}_2)/2, \bar{\theta}_2]$ . The reduced-size approximate problem had the optimal value  $-0.0127419$ , which was an upper bound on the true minimum. This upper bound was accurate because a lower bound computed from 2500 grid points was  $-0.0127419$ . The size of the approximate problem was  $|V|m = 24$  and the computational time was 11.11 seconds. Although the full-size approximate problem gave almost the same optimal value, the size of the approximate problem was larger with  $|V_f|m = 32$ , which resulted in larger computational time, 50.48 seconds.

## 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 is especially evident when the LMI constraint of the given problem has a small support and high degrees. Even with this improvement, a 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. They are indebted to Kazuo Murota for the careful reading of the manuscript and the helpful comments. Their thanks go to the anonymous reviewers for the 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 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 H. Wolkowicz, R. Saigal, and L. Vandenberghe (eds.): *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications* (Kluwer, Boston, USA, 2000), 139–162.
- [2] A. Ben-Tal and A. Nemirovski: Robust convex optimization. *Mathematics of Operations Research*, **23** (1998), 769–805.
- [3] A. Ben-Tal and A. Nemirovski: *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications* (SIAM, Philadelphia, USA, 2001).
- [4] A. Ben-Tal and A. Nemirovski: Robust optimization: methodology and applications. *Mathematical Programming*, **92** (2002), 453–480.
- [5] P.-A. Bliman: On robust semidefinite programming. In *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems (MTNS2004)*, Leuven, Belgium, 2004.
- [6] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan: *Linear Matrix Inequalities in System and Control Theory* (SIAM, Philadelphia, USA, 1994).
- [7] L. El Ghaoui and S.-I. Niculescu: Robust decision problems in engineering: a linear matrix inequality approach. In L. El Ghaoui and S.-I. Niculescu (eds.): *Advances in Linear Matrix Inequality Methods in Control* (SIAM, Philadelphia, USA, 2000), 3–37.
- [8] L. El Ghaoui, F. Oustry, and H. Lebret: Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, **9** (1998), 33–52.
- [9] 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*, **41** (2005), 314–321.
- [10] Y. Isaka and Y. Oishi, Exploiting sparsity in a matrix-dilation approach to robust semidefinite programming (in Japanese). *Transactions of the Society of Instrument and Control Engineers*, **43** (2007), 93–101.
- [11] S. Kim, M. Kojima, and H. Waki: Generalized Lagrangean duals and sums of squares relaxations of sparse polynomial optimization problems. *SIAM Journal on Optimization*, **15** (2005), 697–719.
- [12] 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, 2003.
- [13] M. Kojima, S. Kim, and H. Waki: Sparsity in sums of squares of polynomials. *Mathematical Programming*, **103** (2005), 45–62.
- [14] M. Kojima and M. Muramatsu: A note on sparse SOS and SDP relaxations for polynomial optimization problems over symmetric cones. *Computational Optimization and Applications*, **42** (2009), 31–41.



- [15] J.B. Lasserre: Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, **11** (2001), 796–817.
- [16] J.B. Lasserre: Convergent SDP-relaxations in polynomial optimization with sparsity. *SIAM Journal on Optimization*, **17** (2006), 822–843.
- [17] 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, 2004, 284–289, available from <http://control.ee.ethz.ch/~joloef/wiki/pmwiki.php>
- [18] 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, 2003, 6115–6120.
- [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*, **12** (1999), 390–394.
- [20] A. Nemirovskii: Several NP-hard problems arising in robust stability analysis. *Mathematics of Control, Signals, and Systems*, **6** (1993), 99–105.
- [21] A.T. Neto: Parameter dependent Lyapunov functions for a class of uncertain linear systems: an LMI approach. In *Proceedings of the 38th Conference on Decision and Control*, Phoenix, USA, 1999, 2341–2346.
- [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, 2001, 3593–3594.
- [23] Y. Oishi: A region-dividing approach to robust semidefinite programming and its error bound. In *Proceedings of the 2006 American Control Conference*, Minneapolis, USA, 2006, 123–129. Also submitted to *SICE Journal of Control, Measurement, and System Integration*.
- [24] Y. Oishi and Y. Isaka: Exploiting sparsity in the matrix-dilation approach to robust semidefinite programming. In *Proceedings of the 2007 American Control Conference*, New York, USA, 2007, 6169–6176.
- [25] P.A. Parrilo: Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, **96** (2003), 293–320.
- [26] P.A. Parrilo: Exploiting algebraic structure in sum of squares programs. In D. Henrion and A. Garulli (eds.): *Positive Polynomials in Control* (Springer, Berlin, Germany, 2005), 181–194.
- [27] D. Peaucelle and M. Sato: LMI tests for positive definite polynomials: slack variable approach. *IEEE Transactions on Automatic Control*, **54** (2009), 886–891.
- [28] C. Roos, T. Terlaky, and J.-P. Vial: *Interior Point Methods for Linear Optimization*, 2nd ed. (Springer, New York, USA, 2006).
- [29] S.K. Rao, P. Sadayappan, F.K. Hwang, and P.W. Shor: The rectilinear Steiner arborescence problem. *Algorithmica*, **7** (1992), 277–288.
- [30] C.W. Scherer: Relaxations for robust linear matrix inequality problems with verifications for exactness. *SIAM Journal on Matrix Analysis and Applications*, **27** (2005), 365–395.
- [31] C.W. Scherer: LMI relaxations in robust control. *European Journal of Control*, **12** (2006), 3–29.

- [32] C.W. Scherer and C.W.J. Hol: Matrix sum-of-squares relaxations for robust semi-definite programs. *Mathematical Programming*, **107** (2006), 189–211.
- [33] W. Shi and C. Su: The rectilinear Steiner arborescence problem is NP-complete. *SIAM Journal on Computing*, **35** (2006), 729–740.
- [34] J.F. Sturm: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, **11–12** (1999), 625–653, available from <http://sedumi.mcmaster.ca/>
- [35] K. Takagi and H. Nishimura, Gain-scheduling control of a tower crane considering varying load-rope length. *JSME International Journal*, Ser. C, **42** (1999), 914–921.
- [36] A. Trofino and C.E. de Souza: Biquadratic stability of uncertain linear systems. *IEEE Transactions on Automatic Control*, **46** (2001), 1303–1307.
- [37] 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*, **17** (2006), 218–242.

Yasuaki Oishi  
Department of Systems Design and Engineering  
Nanzan University  
Seireicho 27, Seto 489-0863, Japan  
E-mail: oishi@nanzan-u.ac.jp