NANZAN-TR-2007-02

Asymptotic Exactness of Parameter-Dependent Lyapunov Functions: An Error Bound and Exactness Verification

Yasuaki Oishi

July 2007

Technical Report of the Nanzan Academic Society
Mathematical Sciences and Information Engineering

Asymptotic Exactness of Parameter-Dependent Lyapunov Functions: An Error Bound and Exactness Verification*

Yasuaki Oishi†

This paper provides an approximate approach to a robust semidefinite programming problem with a functional variable and shows its asymptotic exactness. This problem covers a variety of control problems including a robust stability/performance analysis with a parameter-dependent Lyapunov function. In the proposed approach, an approximate semidefinite programming problem is constructed based on the division of the set of parameter values. This approach is asymptotically exact in the sense that, as the resolution of the division becomes higher, the optimal value of the constructed approximate problem converges to that of the original problem. Our convergence analysis is quantitative. In particular, this paper gives an *a priori* upper bound on the discrepancy between the optimal values of the two problems. Moreover, it discusses how to verify that an optimal solution of the approximate problem is actually optimal also for the original problem. Finally, the results are generalized for robust stability/performance analysis against a time-varying parameter and for stability analysis of a nonlinear system.

Keywords: parameter-dependent Lyapunov functions, robust semidefinite programming, approximation error, exactness verification, matrix dilation, linear matrix inequalities.

1. Introduction

Search of a parameter-dependent Lyapunov function is an important control problem for less conservative analysis and design of a parameter-dependent system. This problem is, however, not easy due to its infinite-dimensional nature. A widely accepted approach is to assume polynomial parameter dependence on a Lyapunov function and to apply recent techniques based on positive polynomials [14, 5, 19, 21, 22]. As a result, we have a standard semidefinite programming (SDP) problem, which is solvable with the efficient interior-point method. This approach is conservative by two reasons. First, an infinite-dimensional problem to find an unknown function is reduced to a finite-dimensional one. Second, a semi-infinite constraint necessary to hold for infinitely many parameter values is replaced by a finitely many linear matrix inequality (LMI) constraints.

Bliman [2] showed asymptotic exactness of a polynomially parameter-dependent Lyapunov function. That is, there is no conservatism if we assume polynomial parameter dependence

*July 28, 2007

†Department of Information Systems and Mathematical Sciences, Nanzan University, Seireicho 27, Seto 489-0863, Japan; Phone: +81-561-89-2000; Fax: +81-561-89-2082; Email: oishi@nanzan-u.ac.jp

of sufficiently high degree. This result combined with asymptotic exactness of the positive-polynomial techniques [10, 15, 20, 3, 24, 25] implies asymptotic exactness of the above approach. This analysis is, however, only qualitative. Namely, it states that conservatism asymptotically vanishes as the degrees of the Lyapunov function and the associated positive polynomial increase. It does not tell how large degrees are required to suppress conservatism to a certain level. This theoretical limitation prevents us from efficient use of the above approach to suppress conservatism with small computational complexity.

In this paper, we give an approximate approach to a parameter-dependent Lyapunov function and show its asymptotic exactness in a quantitative manner. We generally formulate the problem as a robust SDP problem with a functional variable. This problem covers a wide range of problems such as robust stability/performance analysis and design of a gain-scheduled state-feedback controller. The proposed approach is based on division of the parameter set. We make the infinite-dimensional problem to a finite-dimensional one by considering a piecewise polynomial consistent with the division. We make the semi-infinite constraint to a finite number of LMI constraints using the matrix-dilation technique in [17, 18]. Thus, a standard SDP problem is constructed to approximate the original robust SDP problem. As the resolution of the division goes up, the optimal value of the approximate problem converges to that of the original problem. In particular, we give an *a priori* upper bound on the approximation error, that is, the discrepancy between the two optimal values. This bound shows a quantitative relationship between the approximation error and the resolution of the division. This bound is also useful to make an efficient division, which gives good approximation with low computational complexity. Moreover, we discuss how to verify exactness of the approximation. With the proposed method, we can detect that the obtained approximate result is actually exact and higher resolution of the division is not necessary any more. At the end of the paper, the results are generalized for robust stability/performance analysis against a time-varying parameter and for stability analysis of a nonlinear system. There, the considered robust SDP problem includes not only an unknown function but also its derivatives.

These results are extension of the results of the present author [17, 18], which is on a robust SDP problem without a functional variable. It is notable that the extension is natural in spite of apparently large difference due to the existence of a functional variable. Many authors have considered the use of piecewise Lyapunov functions [4, 16, 9, 13, 23, 11, 6]. Although some of them discussed asymptotic exactness of their methods [9, 13, 23], their discussion was limited to qualitative one. Coutinho–Danès [6] gave an approach close to the present one without discussion on asymptotic exactness. Scherer [24] discussed verification of exactness in the case of no functional variable.

Construction of the paper is as follows. After a problem is formulated in Section 2, our approximate approach is presented in Section 3. Section 4 gives an upper bound on the approximation error and shows asymptotic exactness of our approach. Section 5 discusses verification of exactness. Section 6 provides a numerical example. After the results are generalized in Section 7, the paper is concluded in Section 8.

The symbol $\mathbb{R}^p$ stands for the set of p-dimensional real vectors while $\mathbb{Z}_+^p$ 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

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

For a set S, the symbol $|S|$ denotes its cardinality.

2. Problem

The problem to be considered in this paper is the following:

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

where the optimization variables are a vector $x \in \mathbb{R}^n$ and a function $\phi \in \Phi$. The set Φ is a linear space of piecewise continuous functions mapping Θ to $\mathbb{R}^{n_\phi}$. The set Θ is a given closed polytope in $\mathbb{R}^p$ and works as the domain of an uncertain p-dimensional parameter θ. The objective function is a linear function of x only, which is described as $c^{\mathrm{T}}x$ with given $c \in \mathbb{R}^n$. In the constraints, the value of $E(x)$ is an $\ell \times \ell$ symmetric matrix while the value of $\mathcal{F}(x, \phi(\theta); \theta)$ is an $m \times m$ symmetric matrix. The function $E(x)$ is affine in $x \in \mathbb{R}^n$. The function $\mathcal{F}(x, a; \theta)$ is affine in $x \in \mathbb{R}^n$ and $a \in \mathbb{R}^{n_\phi}$ while polynomial in θ. The optimal value of the problem P is denoted by $\inf P$ with the attention that the minimum may not be attained.

The problem P is called a *robust SDP problem* because the second constraint has to be satisfied for all possible values of the uncertain parameter θ. It is also notable that P has a functional variable $\phi \in \Phi$. Many control problems are formulated into the form of P. Here is a simple example.

Example 1. Consider a parameter-dependent linear system

$$\begin{aligned} \dot{\xi}(t) &= A(\theta)\xi(t) + B(\theta)u(t), \\ y(t) &= C(\theta)\xi(t), \end{aligned}$$

where the matrices $A(\theta)$, $B(\theta)$, and $C(\theta)$ are polynomials of a time-invariant but unknown parameter θ. Let the domain of θ be a closed polytope Θ. Suppose that $A(\theta)$ is Hurwitz for all $\theta \in \Theta$. Then, the maximum L^2-induced norm over Θ is computed through the problem:

$$\begin{aligned} &\text{minimize} \quad x \\ &\text{subject to} \quad \begin{bmatrix} -A(\theta)R(\theta) - R(\theta)A(\theta)^{\mathrm{T}} & -R(\theta)C(\theta)^{\mathrm{T}} & -B(\theta) \\ -C(\theta)R(\theta) & xI & O \\ -B(\theta)^{\mathrm{T}} & O & xI \end{bmatrix} \succeq O \quad (\forall \theta \in \Theta), \end{aligned}$$

where the optimization variables are a real scalar x and a piecewise continuous symmetric-matrix-valued function $R(\theta)$. No generality is lost by the piecewise continuity of $R(\theta)$ because $A(\theta)$ is continuous and Θ is compact [2]. This problem is in the form of P. $\square$

Various generalization is possible with this example. The matrices $A(\theta)$, $B(\theta)$, and $C(\theta)$ can be rational functions of θ. In this case, multiplication of an appropriate polynomial makes the constraint polynomial in θ. Design of a gain-scheduled state-feedback controller has a similar structure to Example 1 and can be formulated into P. When the system depends on a time-varying parameter, we need to modify the problem P. This will be discussed in Section 7.

The problem P is difficult to solve by two reasons: it has a functional variable $\phi \in \Phi$ and the semi-infinite constraint $\mathcal{F}(x, \phi(\theta); \theta) \succeq O$ $(\forall \theta \in \Theta)$. In the succeeding sections, we propose an approximate approach and discuss its properties.

3. Matrix-dilation approach

An approximate approach to the problem P is presented in this section. In the approach, we require ϕ to be a low-order polynomial and make the problem finite-dimensional. In order to improve the quality of approximation, we divide the parameter set Θ into several subregions and allow ϕ to be a piecewise polynomial. To deal with the semi-infinite constraint, we employ the matrix-dilation approach in [17, 18], which is again based on the division of the parameter set. Hence, we use the division for two purposes: to approximate the functional variable and to approximate the semi-infinite constraint. As the resolution of the division goes up, approximation is improved in the two senses. We will quantitatively investigate this property in the next section.

For the function variable $\phi(\theta)$, we use a fixed-order polynomial $\sum_{\alpha \in S} u_\alpha \theta^\alpha$ for some finite set $S \subset \mathbb{Z}_+^p$. We use the coefficients $u = (u_\alpha) \in \mathbb{R}^{n_u}$ to characterize the polynomial and write $\phi_u(\theta) = \sum_{\alpha \in S} u_\alpha \theta^\alpha$. Here, the dimension n_u is equal to $|S| n_\phi$. Substitution of $\phi_u(\theta)$ into $\mathcal{F}$ makes $\mathcal{F}$ dependent on the finite-dimensional variables x and u. This is made explicit by the notation

$$F(x, u; \theta) := \mathcal{F}(x, \phi_u(\theta); \theta).$$

Note that F is affine in x and u while polynomial in θ.

We next introduce a *division* $\mathit{\Delta}$ of the parameter set $\mathit{\Theta}$. This is a finite collection of p-dimensional closed convex polytopes $\{\mathit{\Theta}^{[1]}, \mathit{\Theta}^{[2]}, \ldots, \mathit{\Theta}^{[J]}\}$ such that the intersection $\mathit{\Theta}^{[j]} \cap \mathit{\Theta}^{[k]}$ has no interior point for any $j \neq k$ and the union $\cup_{j=1}^{J} \mathit{\Theta}^{[j]}$ is equal to the whole parameter set $\mathit{\Theta}$. We allow u to take a different value $u^{[j]}$ depending on the subregion $\mathit{\Theta}^{[j]}$ $(j = 1, 2, \ldots, J)$. Hence, our ϕ is a piecewise polynomial. We consider the following approximate problem:

$$\begin{array}{lll} P_0(\mathit{\Delta}): & \text{minimize} & c^{\mathrm{T}} x \\ & \text{subject to} & E(x) \succeq O, \quad F(x, u^{[j]}; \theta) \succeq O \quad (\forall \theta \in \mathit{\Theta}^{[j]},\ \forall j = 1, 2, \ldots, J), \end{array}$$

where the optimization variables are $x \in \mathbb{R}^n$ and $u^{[1]}, u^{[2]}, \ldots, u^{[J]} \in \mathbb{R}^{n_u}$. Since only a limited class of functions are considered for ϕ, we have $\inf P \leq \inf P_0(\mathit{\Delta})$.

The problem $P_0(\mathit{\Delta})$ is still difficult to solve because it has a semi-infinite constraint. We use the matrix-dilation approach to circumvent this difficulty. In particular, we replace the semi-infinite constraint with its sufficient condition that can be expressed by a finite number of LMIs. The resulting problem $P(\mathit{\Delta})$ is a standard SDP problem and thus solvable with the interior-point method. We have $\inf P \leq \inf P_0(\mathit{\Delta}) \leq \inf P(\mathit{\Delta})$ and expect that $\inf P(\mathit{\Delta})$ converges to $\inf P$ as the resolution of the division $\mathit{\Delta}$ becomes higher.

For construction of the problem $P(\mathit{\Delta})$, we use the expansion $F(x, u; \theta) = \sum_\alpha F_\alpha(x, u) \theta^\alpha$. In many practical situations, only a part of the coefficient matrices $F_\alpha(x, u)$ is nonzero. We would like to exploit this sparse structure to construct $P(\mathit{\Delta})$ of small size. To this end, we consider a directed graph (V, A) having the following properties.

(i) The vertex set V is a finite subset of $\mathbb{Z}_+^p$. It consists of the origin, all α's with nonzero coefficient matrices $F_\alpha(x, u)$, and possibly some additional points.

(ii) The arc set A is a finite set of arcs (α, β) with $\alpha, \beta \in V$ such that $\beta_i = \alpha_i + 1$ for one and only one $i = 1, 2, \ldots, p$ and $\beta_i = \alpha_i$ for all the remaining i's.

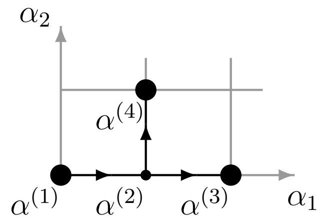

Figure 1. A rectilinear Steiner arborescence in the case of $p = 2$ and $\{\alpha \,|\, F_\alpha(x,u) \neq O\} = \{[0\ \ 0]^{\mathrm{T}}, [2\ \ 0]^{\mathrm{T}}, [1\ \ 1]^{\mathrm{T}}\}$.

(iii) For any vertex $\alpha \in V$, the directed graph (V, A) has one and only one directed path from the origin to α.

Figure 1 shows such a graph (V, A) in the case of $p = 2$ and $\{\alpha \,|\, F_\alpha(x,u) \neq O\} = \{[0\ \ 0]^{\mathrm{T}}, [2\ \ 0]^{\mathrm{T}}, [1\ \ 1]^{\mathrm{T}}\}$. Note that the vertex $[1\ \ 0]^{\mathrm{T}}$ is added for satisfaction of the conditions (ii) and (iii). A graph (V, A) satisfying (i)–(iii) is called a *rectilinear Steiner arborescence* for the set $\{\alpha \,|\, F_\alpha(x,u) \neq O\}$. When the set $\{\alpha \,|\, F_\alpha(x,u) \neq O\}$ has small cardinality, we can construct a rectilinear Steiner arborescence with small $|V|$.

We number the elements of V as $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(|V|)}$. The numbering is arbitrary as far as $\alpha^{(1)} = 0$. We here define the matrices:

$$G(x,u) := \begin{bmatrix} 2F_{\alpha^{(1)}}(x,u) & F_*(x,u) \\ F_*(x,u)^{\mathrm{T}} & O \end{bmatrix}, \tag{1}$$

$$F_*(x,u) := [F_{\alpha^{(2)}}(x,u)\ \ F_{\alpha^{(3)}}(x,u)\ \ \cdots\ \ F_{\alpha^{(|V|)}}(x,u)],$$

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

It is easy to see $2F(x,u;\theta) = M(\theta)^{\mathrm{T}} G(x,u) M(\theta)$. We next define the matrix $H(\theta)$ by

$$H(\theta) := \widetilde{H}(\theta) \otimes I_m \tag{2}$$

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

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

for $q = 1, 2, \ldots, |V|$ and $r = 1, 2, \ldots, |V| - 1$. It is straightforward to see $M(\theta)^{\mathrm{T}} H(\theta) = O$ and the matrix $[M(\theta)\ \ H(\theta)]$ is square and nonsingular for any θ. Hence, $H(\theta)$ is called an *orthogonal complement* of $M(\theta)$. It is notable that this particular orthogonal complement is affine in θ. In the case of Figure 1, the matrices $M(\theta)$ and $H(\theta)$ look as

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

With these preparations, we introduce the desired problem:

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

where the optimization variables are $x \in \mathbb{R}^n$, $u^{[1]}, u^{[2]}, \ldots, u^{[J]} \in \mathbb{R}^{n_u}$, and $|V|m \times (|V| - 1)m$ matrices $W^{[1]}, W^{[2]}, \ldots, W^{[J]}$. The symbol $\operatorname{ver} \mathit{\Theta}^{[j]}$ denotes the set of the vertices of the convex polytope $\mathit{\Theta}^{[j]}$. Since the second constraint is affine in θ, it actually holds for any $\theta \in \mathit{\Theta}^{[j]}$. Then, premultiplication of $M(\theta)^{\mathrm{T}}$ and postmultiplication of $M(\theta)$ give $F(x, u^{[j]}; \theta) \succeq O$. This leads to the next proposition [17, 18].

Proposition 2. *The feasible region of $P(\mathit{\Delta})$ projected onto the space of $(x, u^{[1]}, u^{[2]}, \ldots, u^{[J]})$ is contained in the feasible region of $P_0(\mathit{\Delta})$. In particular,* $\inf P \le \inf P_0(\mathit{\Delta}) \le \inf P(\mathit{\Delta})$.

We here have an approximate approach to P, which we call a *matrix-dilation approach*. That is, we first choose a division $\mathit{\Delta}$ for the parameter set $\mathit{\Theta}$ and solve the approximate problem $P(\mathit{\Delta})$. If the obtained optimal value $\inf P(\mathit{\Delta})$ is satisfactory, we stop. Otherwise, we subdivide $\mathit{\Delta}$ and solve again the new approximate problem $P(\mathit{\Delta})$.

This is a natural approach to the problem P. However, it is not clear how $\inf P(\mathit{\Delta})$ converges to $\inf P$. We quantitatively answer to this question in the next section.

4. An upper bound on the approximation error

4.1. Upper bound

In this section, we obtain an *a priori* upper bound on the *approximation error* $|\inf P(\mathit{\Delta}) - \inf P|$ in terms of the resolution of the division. This is a generalization of the upper bounds in [17, 18], which are for a robust SDP problem without a functional variable.

We make the following assumptions to have the result. The first assumption is a reasonable one, which means strict feasibility of the problem P. The second assumption means that search of bounded x and ϕ is sufficient for the problem P as well as its modification P_ϵ, to be introduced in Section 4.2. This assumption is mild enough because it is satisfied when P and P_ϵ have bounded optimal solutions. The last assumption is again mild one, which means that the polynomial $\phi_u(\theta)$ has a constant term.

Assumption 3.

(a) There exist $x \in \mathbb{R}^n$ and $\phi \in \Phi$ such that $E(x) \succ O$ and $\mathcal{F}(x, \phi(\theta); \theta) \succ O$ $(\forall\theta \in \Theta)$.

(b) There exist three positive numbers $\overline{\epsilon}$, $\overline{x}$, and $\overline{\phi}$ such that, for any $0 \le \epsilon \le \overline{\epsilon}$ and any $v \in \mathbb{R}$, the set $\{(x, \phi) \in \mathbb{R}^n \times \Phi \mid c^{\mathrm{T}}x \le v,\ E(x) \succeq O,\ \mathcal{F}(x, \phi(\theta); \theta) \succeq \epsilon I\ (\forall\theta \in \Theta)\}$ is either empty or having an element (x, ϕ) with $\|x\| < \overline{x}$ and $\|\phi(\theta)\| < \overline{\phi}$ $(\forall\theta \in \Theta)$.

(c) The set S in the polynomial $\phi_u(\theta) = \sum_{\alpha \in S} u_\alpha \theta^\alpha$ includes the origin. □

We next need a measure of the resolution of the division. For a division $\Delta = \{\Theta^{[1]}, \Theta^{[2]}, \ldots, \Theta^{[J]}\}$ of the parameter set Θ, the *radius* of a subregion $\Theta^{[j]}$ is $\operatorname{rad}\Theta^{[j]} := \min_{\theta\in\Theta^{[j]}} \max_{\theta'\in\Theta^{[j]}} \max_{i=1,2,\ldots,p} |\theta'_i - \theta_i|$. A $\theta \in \Theta^{[j]}$ that attains the minimum is called a *center* of $\Theta^{[j]}$. The *maximum radius* of the division Δ is defined as $\overline{\operatorname{rad}}\,\Delta := \max_{j=1,2,\ldots,J} \operatorname{rad}\Theta^{[j]}$, which measures the resolution of Δ.

With these preparations, we have the desired upper bound. The proof will be given in Section 4.2.

Theorem 4. *Under Assumption 3, there exist positive numbers C and C', which are independent of the division Δ, such that*

$$|\inf P(\Delta) - \inf P| \le C\ \overline{\operatorname{rad}}\,\Delta$$

for any Δ satisfying $\overline{\operatorname{rad}}\,\Delta \le C'$.

The concrete forms of C and C' will be given in (6) and (7).

This theorem immediately implies the asymptotic exactness of our approach. To see this, consider a sequence of divisions whose maximum radii converge to zero. Then, the corresponding approximate problems $P(\Delta)$ have the approximation errors converging to zero because C and C' are independent of Δ. Note that Theorem 4 also gives the rate of convergence. Namely, the convergence is at least in the linear order of the maximum radius.

Theorem 4 shows a tradeoff between the approximation error and the computational complexity of the present approach. Indeed, the theorem tells that the maximum radius should be small for small approximation error. A small maximum radius leads to a large number of subregions and, then, to a large number of variables and constraints in the approximate problem $P(\Delta)$. Since the number of subregions is in the order of $(\overline{\operatorname{rad}}\,\Delta)^{-p}$, a small maximum radius invites computational difficulty especially when p is large.

In order to address this issue, we can improve the theorem by replacing the maximum radius with a more sophisticated measure of the resolution. Consider an approximate problem $P(\mathit{\Delta})$ for some division $\mathit{\Delta}$ and suppose that its optimal value is attained. The constraint $G(x, u^{[j]}) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]}H(\theta)^{\mathrm{T}} \succeq O$ is called *active* for an optimal solution if the left-hand-side matrix has a zero eigenvalue for this optimal solution. A subregion $\Theta^{[j]}$ is called active if it has an active constraint. We define the *maximum active radius* of the division $\mathit{\Delta}$, denoted by $\overline{\text{a-rad}}\,\mathit{\Delta}$, as the maximum radius over all active subregions. When $P(\mathit{\Delta})$ has multiple optimal solutions, we define $\overline{\text{a-rad}}\,\mathit{\Delta}$ by taking the minimum of all possible values. In this case, actual computation of $\overline{\text{a-rad}}\,\mathit{\Delta}$ may be difficult while computation of its upper bound is easy. When the optimal value is not attained in $P(\mathit{\Delta})$, we define $\overline{\text{a-rad}}\,\mathit{\Delta} := \overline{\text{rad}}\,\mathit{\Delta}$. We here have the following result. The proof is given in Section 4.2.

Corollary 5. *Under Assumption 3, the inequality*

$$|\inf P(\mathit{\Delta}) - \inf P| \leq C\ \overline{\text{a-rad}}\,\mathit{\Delta}$$

holds for any $\mathit{\Delta}$ *satisfying* $\overline{\text{rad}}\,\mathit{\Delta} \leq C'$, *where* C *and* C' *are the same positive constants as in Theorem 4.*

This corollary states that we do not need to decrease the maximum radius but the maximum active radius for the reduction of the approximation error. Although direct decrease of the maximum active radius is difficult, a heuristic algorithm can be constructed.

Algorithm 6.

0. Choose a coarse division $\mathit{\Delta}$.
1. Solve $P(\mathit{\Delta})$ for the current division $\mathit{\Delta}$.
2. Stop if the obtained solution is satisfactory.
3. If an optimal solution is found, find an active subregion that attains $\overline{\text{a-rad}}\,\mathit{\Delta}$ and divide it into two subregions so that they have small radii. Otherwise, find a subregion that attains $\overline{\text{rad}}\,\mathit{\Delta}$ and divide it similarly.
4. Go back to Step 1 with the updated division $\mathit{\Delta}$. □

This type of algorithm was originally proposed for a robust SDP problem without a functional variable [17].

4.2. Proof

We prove Theorem 4 as well as Corollary 5 in this subsection. We present the full proof for the sake of completeness though it includes some discussion found also in [17, 18]. The key idea is to relate the following auxiliary problem with the approximate problem $P(\varDelta)$:

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

where ϵ is a nonnegative number. We will find below an ϵ that attains $\inf P \le \inf P(\varDelta) \le \inf P_\epsilon$. Note that $\inf P_\epsilon$ 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 $\inf P_\epsilon$, we have $|\inf P_\epsilon - \inf P| = \inf P_\epsilon - \inf P \le g\epsilon$. Since this implies $|\inf P(\varDelta) - \inf P| \le g\epsilon$, we obtain the desired upper bound on the approximation error.

We first prepare a $|V|m \times |V|m$ matrix $L(\theta)$, which will be used for simplification of the dilated LMI constraint $G(x, u^{[j]}) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]}H(\theta)^{\mathrm{T}} \succeq O$. The definition of $L(\theta)$ is based on the rectilinear Steiner arborescence (V, A). For $\alpha^{(q)}, \alpha^{(r)} \in V$, we say that $\alpha^{(q)}$ is *reachable* from $\alpha^{(r)}$ in (V, A) if these two vertices are identical or (V, A) has a directed path connecting $\alpha^{(r)}$ to $\alpha^{(q)}$. For the convenience of the proof, we assume that the numbering of the vertices is consistent with the partial order defined by (V, A). In other words, we have $q \ge r$ whenever $\alpha^{(q)}$ is reachable from $\alpha^{(r)}$. This assumption does not contradict with $\alpha^{(1)} = 0$ because the origin is the root of the arborescence (V, A). Also, this assumption does not harm the generality of the result because the result does not depend on the numbering.

Now, the matrix $L(\theta)$ is defined as

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

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 $q \ge r$. Moreover, its diagonal elements are all equal to unity. A consequence is nonsingularity of $\widetilde{L}(\theta)$ and also of $L(\theta)$.

Straightforward calculation gives the concrete forms of $L(\theta)^{\mathrm{T}}G(x, u)L(\theta)$ and $L(\theta)^{\mathrm{T}}H(\theta')$ for $\theta, \theta' \in \mathbb{R}^p$.

Lemma 7. *For the matrices $G(x, u)$ and $L(\theta)$ in (1) and (3), respectively, we have*

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

with

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

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

Lemma 8. *For the matrices $H(\theta)$ and $L(\theta)$ in (2) and (3), 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.

With these preparations, we relate P_ϵ and $P(\varDelta)$. By Assumption 3 (a), there exists $\epsilon_0 > 0$ such that, for any $0 \le \epsilon \le \epsilon_0$, the auxiliary problem P_ϵ is strictly feasible. This ϵ_0 can be chosen smaller than or equal to the $\bar\epsilon$ in Assumption 3 (b). We consider the set of all u such that the corresponding $\phi_u(\theta)$ is constantly equal to some vector a with $\|a\| < \overline{\phi}$. This set is nonempty by Assumption 3 (c). We let $\overline{F}$ be an upper bound of $\overline{\sigma}[F_{**}(x,u;\theta)]$ over all $\|x\| < \overline{x}$, all $\theta \in \varTheta$, and all u in the considered set.

We begin by the special case that $\varTheta \subseteq [-1,1]^p$. In this case, $|\theta_i| \le 1$ for any $\theta \in \varTheta$ and any $i = 1,2,\dots,p$.

Lemma 9. *Suppose that $\varTheta \subseteq [-1,1]^p$ and*

$$\overline{\mathrm{rad}}\,\varDelta \le \min\Big\{\frac{2\epsilon_0}{(\overline{F} + \sqrt{|V|m})^2},\ \frac{1}{|V|}\Big\}.$$

Then, we have $\inf P \le \inf P(\varDelta) \le \inf P_\epsilon$ *for*

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

Proof. Since $\inf P \le \inf P(\varDelta)$ by Proposition 2, we will show $\inf P(\varDelta) \le \inf P_\epsilon$. Let $\varDelta$ be $\{\varTheta^{[1]}, \varTheta^{[2]}, \dots, \varTheta^{[J]}\}$. The ϵ given in the lemma satisfies $0 \le \epsilon \le \epsilon_0$. Let (x,ϕ) be a feasible solution of P_ϵ satisfying $\|x\| \le \overline{x}$ and $\|\phi(\theta)\| \le \overline{\phi}$ $(\forall\theta \in \varTheta)$. Here, $\mathcal{F}(x,\phi(\theta);\theta) \succeq \epsilon I$ for any $\theta \in \varTheta$. We show below that there exists $(u^{[j]}, W^{[j]})$ for each $j = 1,2,\dots,J$ such that

$$G(x,u^{[j]}) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]}H(\theta)^{\mathrm{T}} \succeq O \quad (\forall\theta \in \mathrm{ver}\,\varTheta^{[j]}), \tag{4}$$

in other words, this x is feasible for $P(\mathit{\Delta})$. Then the proof is complete because (x, ϕ) can be chosen so that $c^{\mathrm{T}}x$ is arbitrarily close to $\inf P_{\epsilon}$. In fact, the inequality (4) holds if we choose $(u^{[j]}, W^{[j]})$ for each j so that $\phi_{u^{[j]}}(\theta) \equiv \phi(\theta^{\mathrm{c}})$ and $W^{[j]} = (1/\operatorname{rad}\mathit{\Theta}^{[j]})H(\theta^{\mathrm{c}})$, where θ^{c} is a center of $\mathit{\Theta}^{[j]}$. Note in this case that $\overline{\sigma}[F_{**}(x, u^{[j]}; \theta)] \leq \overline{F}$ for any j and $\theta \in \mathit{\Theta}$.

In order to show the desired inequality (4), we premultiply $L(\theta^{\mathrm{c}})^{\mathrm{T}}$ and postmultiply $L(\theta^{\mathrm{c}})$ to it. Lemmas 7 and 8 give the concrete forms of $L(\theta^{\mathrm{c}})^{\mathrm{T}}G(x, u^{[j]})L(\theta^{\mathrm{c}})$ and $(\operatorname{rad}\mathit{\Theta}^{[j]})(W^{[j]})^{\mathrm{T}}L(\theta^{\mathrm{c}}) = H(\theta^{\mathrm{c}})^{\mathrm{T}}L(\theta^{\mathrm{c}})$. By Lemma 8 again, the product $(1/\operatorname{rad}\mathit{\Theta}^{[j]})L(\theta^{\mathrm{c}})^{\mathrm{T}}H(\theta)$ has the form

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

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

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

whose magnitude is at most one since $|\theta_i^{\mathrm{c}}| \leq 1$ and $|\theta_i^{\mathrm{c}} - \theta_i| \leq \operatorname{rad}\mathit{\Theta}^{[j]}$ for $i = 1, 2, \ldots, p$. Let us write the product $(1/\operatorname{rad}\mathit{\Theta}^{[j]})L(\theta^{\mathrm{c}})^{\mathrm{T}}H(\theta)$ as $[H_1^{\mathrm{T}} \quad H_2^{\mathrm{T}}]^{\mathrm{T}}$ 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

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

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

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

Now, the left-hand side of (4) multiplied by $L(\theta^{\mathrm{c}})^{\mathrm{T}}$ and $L(\theta^{\mathrm{c}})$ has the upper-left $m \times m$ block equal to

$$2F(x, u^{[j]}; \theta^{\mathrm{c}}) = 2\mathcal{F}(x, \phi_{u^{[j]}}(\theta^{\mathrm{c}}); \theta^{\mathrm{c}}) = 2\mathcal{F}(x, \phi(\theta^{\mathrm{c}}); \theta^{\mathrm{c}}) \succeq 2\epsilon I.$$

Its Schur complement is

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

Noting that $2/\operatorname{rad}\Theta^{[j]} - |V| + 2 > 1/\operatorname{rad}\Theta^{[j]}$ and $\overline{\sigma}[F_{**}(x, u^{[j]}; \theta^{\mathrm{c}})] \le \overline{F}$, 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

$$\overline{\theta} := \max\{1,\ \max_{\theta\in\Theta} \max_{i=1,2,\ldots,p} |\theta_i|\} \tag{5}$$

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

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

we can regard $F_\alpha(x, u)\overline{\theta}^{|\alpha|}$ as a coefficient and $\theta/\overline{\theta}$ as a parameter. Since $\theta/\overline{\theta}$ moves in $[-1, 1]^p$, the discussion in the special case can be applied. To present the result in this case, we define $F'_{**}(x, u; \theta)$ by

$$F'_{**}(x, u; \theta) := \Big[\sum_{\alpha\in V^{(2)}} F_\alpha(x, u)\overline{\theta}^{|\alpha|} \Big(\frac{\theta}{\overline{\theta}}\Big)^{\alpha-\alpha^{(2)}} \quad \sum_{\alpha\in V^{(3)}} F_\alpha(x, u)\overline{\theta}^{|\alpha|} \Big(\frac{\theta}{\overline{\theta}}\Big)^{\alpha-\alpha^{(3)}} \quad \cdots \quad \sum_{\alpha\in V^{(|V|)}} F_\alpha(x, u)\overline{\theta}^{|\alpha|} \Big(\frac{\theta}{\overline{\theta}}\Big)^{\alpha-\alpha^{(|V|)}} \Big].$$

We also define $\overline{F'}$ as an upper bound of $\overline{\sigma}[F'_{**}(x, u; \theta)]$ over all $\|x\| < \overline{x}$, all $\theta \in \Theta$, and all u corresponding to the polynomial $\phi_u(\theta)$ constantly equal to a vector a with $\|a\| < \overline{\phi}$.

Lemma 10. *Suppose that*

$$\overline{\operatorname{rad}}\,\Delta \le \min\Big\{\frac{2\overline{\theta}\epsilon_0}{(\overline{F'} + \sqrt{|V|m})^2},\ \frac{\overline{\theta}}{|V|}\Big\}.$$

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

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

Proof. The replacement of $F_\alpha(x, u)$ by $F_\alpha(x, u)\overline{\theta}^{|\alpha|}$ and θ by $\theta/\overline{\theta}$ does not change the problems P and P_ϵ essentially. 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 9 is valid with $\overline{F}$ replaced by $\overline{F'}$ and $\overline{\operatorname{rad}}\,\Delta$ by $\overline{\operatorname{rad}}\,\Delta/\overline{\theta}$. This completes the proof.

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

$$G'(x, u) = \operatorname{diag}\{I_m,\ T\}G(x, u)\operatorname{diag}\{I_m,\ T\},$$

$$H'(\theta) = \mathrm{diag}\{I_m,\ T\}H(\theta)T^{-1},$$

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

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

is equivalent to the existence of W' satisfying

$$G'(x,u) + H'(\theta)(W')^{\mathrm{T}} + W'H'(\theta)^{\mathrm{T}} \succeq O$$

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

We now take the final step toward the proof of Theorem 4. Recall that we assume Assumption 3. The number ϵ_0 is such that the auxiliary problem P_ϵ is strictly feasible for any $0 \le \epsilon \le \epsilon_0$. It also satisfies $\epsilon_0 \le \overline{\epsilon}$ for the $\overline{\epsilon}$ in Assumption 3 (b). The number $\overline{\theta}$ is as in (5) while $\overline{F'}$ is as in the same paragraph. Finally, let g be an upper bound on the left derivative of $\inf P_\epsilon$ at $\epsilon = \epsilon_0$. Then, with

$$C = \frac{g(\overline{F'} + \sqrt{|V|m})^2}{2\overline{\theta}}, \tag{6}$$

$$C' = \min\Big\{\frac{2\overline{\theta}\epsilon_0}{(\overline{F'} + \sqrt{|V|m})^2},\ \frac{\overline{\theta}}{|V|}\Big\}, \tag{7}$$

we can prove the theorem.

Proof of Theorem 4. Lemma 10 implies that, when $\overline{\mathrm{rad}}\,\varDelta \le C'$, we have $\inf P \le \inf P(\varDelta) \le \inf P_\epsilon$ for $\epsilon = [(\overline{F'} + \sqrt{|V|m})^2/2\overline{\theta}]\,\overline{\mathrm{rad}}\,\varDelta$. Owing to the convexity of $\inf P_\epsilon$, the upper bound g is greater than or equal to the left derivative of $\inf P_\epsilon$ at this ϵ. Hence, the convexity of $\inf P_\epsilon$ again implies $\inf P \ge \inf P_\epsilon - g\epsilon$, from which $\inf P_\epsilon - \inf P \le g\epsilon$. Substitution of the concrete form of ϵ gives the theorem. $\square$

The proof of Corollary 5 follows from subdivision of $\varDelta$.

Proof of Corollary 5. The corollary is reduced to Theorem 4 when the optimal value of $P(\varDelta)$ is not attained. Hence, we assume that the optimal value is attained.

Consider an optimal solution and subdivide each inactive subregion of $\varDelta$, if necessary, so that each of the created subregion has the radius smaller than or equal to $\overline{\text{a-rad}}\,\varDelta$. Then, the resulting new division $\widetilde{\varDelta} = \{\widetilde{\varTheta}^{[k]}\}$ has $\overline{\mathrm{rad}}\,\widetilde{\varDelta} = \overline{\text{a-rad}}\,\varDelta$. Note also $\inf P(\widetilde{\varDelta}) \le \inf P(\varDelta)$. We next consider the SDP dual of $P(\varDelta)$. As is shown below, the dual problem of $P(\varDelta)$ has an optimal solution that attains $\inf P(\mathit{\Delta})$. Based on this solution, we construct a dual feasible solution of $P(\widetilde{\mathit{\Delta}})$. In particular, if a subregion $\widetilde{\mathit{\Theta}}^{[k]}$ is a newly created subregion, assign zero matrices to the dual variables corresponding to this subregion; If a subregion $\widetilde{\mathit{\Theta}}^{[k]}$ coincides with one subregion in $\mathit{\Delta}$, assign the same values as the dual optimal solution of $P(\mathit{\Delta})$. With these assigned values, the dual objective function takes the same value as before, *i.e.*, $\inf P(\mathit{\Delta})$. By weak duality, we have $\inf P(\mathit{\Delta}) \leq \inf P(\widetilde{\mathit{\Delta}}) \leq \inf P(\mathit{\Delta})$. Hence,

$$|\inf P(\mathit{\Delta}) - \inf P| = |\inf P(\widetilde{\mathit{\Delta}}) - \inf P| \leq C\, \overline{\mathrm{rad}}\, \widetilde{\mathit{\Delta}} = C\, \overline{\text{a-rad}}\, \mathit{\Delta},$$

which shows the claim.

We show that the dual problem of $P(\mathit{\Delta})$ has an optimal solution. Let $\epsilon = [(\overline{F'}+\sqrt{|V|m})^2/2\overline{\theta}]\times \overline{\mathrm{rad}}\, \mathit{\Delta}$. Since $0 < \epsilon \leq \epsilon_0 \leq \overline{\epsilon}$, the problem P_ϵ is strictly feasible and has a solution (x, ϕ) satisfying $\|x\| < \overline{x}$ and $\|\phi(\theta)\| < \overline{\phi}$ for any $\theta \in \mathit{\Theta}$. It is possible to assume strict feasibility of this (x, ϕ) because it can be perturbed, if necessary, to the direction of a strictly feasible solution. We construct from this (x, ϕ) a solution of $P(\mathit{\Delta})$ as in the proof of Lemma 9. We then obtain a strictly feasible solution of $P(\mathit{\Delta})$. On the other hand, the problem P is bounded from below by Assumption 3 (b) and so is $P(\mathit{\Delta})$. The duality theorem on SDP (Theorem 2.4.1 of [1]) implies that the dual problem of $P(\mathit{\Delta})$ has an optimal solution. □

5. Verification of exactness

Suppose that we solved an approximate problem $P(\mathit{\Delta})$ for some division $\mathit{\Delta}$ and obtained an optimal solution $(\widehat{x}, \{\widehat{u}^{[j]}\}, \{\widehat{W}^{[j]}\})$. In general, the attained value $c^{\mathrm{T}}\widehat{x} = \inf P(\mathit{\Delta})$ is larger than the true optimal value $\inf P$. In some cases, however, these two values happen to be identical. One of such cases is that there exist some j and $\widehat{\theta} \in \mathit{\Theta}^{[j]}$ such that $(\widehat{x}, \widehat{u}^{[j]})$ is optimal for the problem:

$$\begin{array}{lll} P(\widehat{\theta}): & \text{minimize} & c^{\mathrm{T}}x \\ & \text{subject to} & E(x) \succeq O, \quad F(x, u^{[j]}; \widehat{\theta}) \succeq O. \end{array}$$

Indeed, since only a particular parameter value $\widehat{\theta} \in \mathit{\Theta}$ is considered in this problem, the attained minimum value $c^{\mathrm{T}}\widehat{x}$ is smaller than or equal to $\inf P$. But we also have $\inf P \leq \inf P(\mathit{\Delta}) = c^{\mathrm{T}}\widehat{x}$, which implies that these three values are equal to each other. In this case, $\widehat{\theta}$ is understood as the worst-case parameter value, which prevents $\inf P$ from becoming smaller than the present value.

The purpose of this section is to give a condition for the above situation to occur. If this condition is satisfied, the obtained solution is optimal not only for the approximate problem $P(\mathit{\Delta})$ but also for the original problem P. This result is useful because, in this case, we do not have to improve the approximate problem any more. The condition is obtained by generalization of the result on a robust SDP problem without a functional variable [17]. See also [24] for a related result.

A feasible solution of an SDP problem is optimal if there exists a dual feasible solution satisfying the complementary slackness condition with the given (primal) solution (Section 4.1 of [1]). To make an explicit statement, we introduce the inner product $\langle A, B\rangle := \operatorname{tr} AB$ for symmetric matrices A and B and let e_i^n be the ith elementary n-dimensional vector ($i = 1, 2, \ldots, n$). Then, a feasible solution $(\widehat{x}, \widehat{u}^{[j]})$ of $P(\widehat{\theta})$ is optimal if there exists a matrix pair $(\widehat{Y}, \widehat{Z})$, where $\widehat{Y}$ is an $\ell \times \ell$ symmetric matrix and $\widehat{Z}$ is an $m \times m$ symmetric matrix, that satisfies the dual feasibility condition

$$\widehat{Y} \succeq O, \qquad \widehat{Z} \succeq O, \tag{8}$$

$$\left\langle \widehat{Y},\ E(e_i^n) - E(0)\right\rangle + \left\langle \widehat{Z},\ F(e_i^n, 0; \widehat{\theta}) - F(0, 0; \widehat{\theta})\right\rangle = c_i \quad (i = 1, 2, \ldots, n), \tag{9}$$

$$\left\langle \widehat{Z},\ F(0, e_i^{n_u}; \widehat{\theta}) - F(0, 0; \widehat{\theta})\right\rangle = 0 \quad (i = 1, 2, \ldots, n_u) \tag{10}$$

as well as the complementary slackness condition

$$\langle \widehat{Y},\ E(\widehat{x})\rangle = 0, \tag{11}$$

$$\langle \widehat{Z},\ F(\widehat{x}, \widehat{u}^{[j]}; \widehat{\theta})\rangle = 0. \tag{12}$$

We show below that the existence of such $(\widehat{Y}, \widehat{Z})$ is equivalent to the existence of a dual feasible solution of $P(\Delta)$ having some special structure. Hence, if we solve the approximate problem $P(\Delta)$ and the obtained dual solution has this structure, the obtained solution is exact in the sense that it optimizes the original problem P. Note that the primal-dual interior-point method gives not only a primal solution but also a dual solution. Here, the dual variables of $P(\Delta)$ consist of an $\ell \times \ell$ symmetric matrix, which corresponds to $E(x) \succeq O$, and $|V|m \times |V|m$ symmetric matrices, each of which corresponds to $G(x, u^{[j]}) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]} H(\theta)^{\mathrm{T}} \succeq O$ ($\theta \in \operatorname{ver} \Theta^{[j]}, j = 1, 2, \ldots, J$).

Theorem 11. *Let $(\widehat{x}, \{\widehat{u}^{[j]}\}, \{\widehat{W^{[j]}}\})$ be a primal feasible solution of the approximate problem $P(\Delta)$ for a division Δ. Let $\widehat{\theta}$ be a point in some subregion $\Theta^{[j]}$. Then, the problem $P(\widehat{\theta})$ has a dual feasible solution $(\widehat{Y}, \widehat{Z})$ satisfying the complementary slackness condition* (11) *and* (12) *with the primal feasible solution $(\widehat{x}, \widehat{u}^{[j]})$ if and only if the approximate problem $P(\Delta)$ has a dual feasible solution satisfying the complementary slackness condition with $(\widehat{x}, \{\widehat{u}^{[j]}\}, \{\widehat{W^{[j]}}\})$ and having the following structure.*

(a) The dual variable corresponding to the constraint $E(x) \succeq O$ is $\widehat{Y}$.

(b) Each remaining nonzero dual variable corresponds to the constraint $G(x, u^{[j]}) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]} H(\theta)^{\mathrm{T}} \succeq O$ *for some vertex* $\theta \in \operatorname{ver} \Theta^{[j]}$.

(c) Let the vertices in (b) be $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(Q)}$*. Then, the set* $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(Q)}\}$ *includes* $\widehat{\theta}$ *in the relative interior of its convex hull.*

(d) The nonzero dual variable corresponding to $\theta^{(q)}$ *is expressed as* $(a^{(q)}/2) M(\widehat{\theta}) \widehat{Z} M(\widehat{\theta})^{\mathrm{T}}$ *for each* $q = 1, 2, \ldots, Q$*, where* $a^{(q)}$ *is the coefficient of a positive convex combination* $\widehat{\theta} = a^{(1)}\theta^{(1)} + a^{(2)}\theta^{(2)} + \cdots + a^{(Q)}\theta^{(Q)}$.

The structure $(a^{(q)}/2) M(\widehat{\theta}) \widehat{Z} M(\widehat{\theta})^{\mathrm{T}}$ can be checked by solving an SDP problem. Indeed, a symmetric matrix A is equal to $(a^{(q)}/2) M(\widehat{\theta}) \widehat{Z} M(\widehat{\theta})^{\mathrm{T}}$ if and only if $AH(\widehat{\theta}) = O$ and the upper-left $m \times m$ submatrix of A is equal to $(a^{(q)}/2)\widehat{Z}$. Since $AH(\widehat{\theta}) = O$ is affine in $\widehat{\theta}$, we can search $\widehat{\theta}$ by solving an SDP problem. See Section 6 for an example.

Proof. Assume the existence of a dual feasible solution of $P(\Delta)$ with the described structure. Then, we have

$$\widehat{Y} \succeq O, \qquad \widehat{Z} \succeq O, \tag{13}$$

$$\Big\langle \widehat{Y},\ E(e_i^n) - E(0) \Big\rangle + \sum_{q=1}^{Q} \Big\langle \frac{a^{(q)}}{2} M(\widehat{\theta}) \widehat{Z} M(\widehat{\theta})^{\mathrm{T}},\ G(e_i^n, 0) - G(0,0) \Big\rangle = c_i \quad (i = 1, 2, \ldots, n), \tag{14}$$

$$\sum_{q=1}^{Q} \Big\langle \frac{a^{(q)}}{2} M(\widehat{\theta}) \widehat{Z} M(\widehat{\theta})^{\mathrm{T}},\ G(0, e_i^{n_u}) - G(0,0) \Big\rangle = 0 \quad (i = 1, 2, \ldots, n_u), \tag{15}$$

$$\sum_{q=1}^{Q} \Big\langle \frac{a^{(q)}}{2} M(\widehat{\theta}) \widehat{Z} M(\widehat{\theta})^{\mathrm{T}},\ H(\theta^{(q)}) W^{\mathrm{T}} + W H(\theta^{(q)})^{\mathrm{T}} \Big\rangle = 0$$

$$(\text{for any } |V|m \times (|V|-1)m \text{ matrix } W), \tag{16}$$

$$\langle \widehat{Y},\ E(\widehat{x}) \rangle = 0, \tag{17}$$

$$\Big\langle \frac{a^{(q)}}{2} M(\widehat{\theta}) \widehat{Z} M(\widehat{\theta})^{\mathrm{T}},\ G(\widehat{x}, \widehat{u}^{[j]}) + H(\theta^{(q)})(\widehat{W}^{[j]})^{\mathrm{T}} + \widehat{W}^{[j]} H(\theta^{(q)})^{\mathrm{T}} \Big\rangle = 0 \quad (q = 1, 2, \ldots, Q). \tag{18}$$

Here, (13)–(16) constitute the dual feasibility condition while (17) and (18) the complementary slackness condition.

We consider (14) first. Application of $a^{(1)} + a^{(2)} + \cdots + a^{(Q)} = 1$ and $M(\widehat{\theta})^{\mathrm{T}} G(x, u) M(\widehat{\theta}) = 2F(x, u; \widehat{\theta})$ gives (9). Similarly, (15) implies (10). Summing up (18) for $q = 1, 2, \ldots, Q$ and using $a^{(1)}\theta^{(1)} + a^{(2)}\theta^{(2)} + \cdots + a^{(Q)}\theta^{(Q)} = \widehat{\theta}$, we have (12). Hence, $(\widehat{Y}, \widehat{Z})$ satisfies (8)–(12), in other words, it is a dual feasible solution of $P(\widehat{\theta})$ satisfying the complementary slackness condition.

Conversely, assume that $(\widehat{Y}, \widehat{Z})$ satisfies (8)–(12). By the reversed reasoning, (14) and (15) follow from (9) and (10), respectively. The equality (16) is a consequence of $a^{(1)}\theta^{(1)} + a^{(2)}\theta^{(2)} + \cdots + a^{(Q)}\theta^{(Q)} = \widehat{\theta}$ and $M(\widehat{\theta})^{\mathrm{T}}H(\widehat{\theta}) = O$. We finally derive (18) from (12). The equality (12) implies

$$\sum_{q=1}^{Q} \left\langle \frac{a^{(q)}}{2} M(\widehat{\theta})\widehat{Z}M(\widehat{\theta})^{\mathrm{T}},\ G(\widehat{x}, \widehat{u}^{[j]}) + H(\theta^{(q)})(\widehat{W}^{[j]})^{\mathrm{T}} + \widehat{W}^{[j]}H(\theta^{(q)})^{\mathrm{T}} \right\rangle = 0.$$

Since each term is nonnegative, the equality (18) follows. □

6. Example

A numerical example is presented for illustration.

Consider a two-dimensional system

$$\dot{\xi}(t) = \begin{bmatrix} 0 & 1 \\ -1 & -1 - 0.1\theta_1 \end{bmatrix} \xi(t) + \begin{bmatrix} 0 \\ 1 + 0.2\theta_1 - 0.1\theta_1^2 - 0.1\theta_2 \end{bmatrix} u(t),$$
$$y(t) = [1 \quad 0]\xi(t)$$

depending on a two-dimensional parameter $\theta = [\theta_1 \quad \theta_2]^{\mathrm{T}} \in \Theta = [0,1]^2$. We compute its maximum L^2-induced norm over Θ. This problem is formulated into a robust SDP problem as in Example 1. Substituting a polynomial $R(\theta) = R_{00} + R_{10}\theta_1 + R_{20}\theta_1^2 + R_{01}\theta_2$ into the constraint there, we have an LMI with the terms of θ_1, θ_1^2, θ_1^3, θ_2, and $\theta_1\theta_2$ as well as the constant term.

With the coarsest division $\Delta = \{\Theta\}$, we constructed the approximate problem $P(\Delta)$ and solved it. The obtained optimal value was 1.20687. We used SeDuMi [26] for the SDP solver with the help of the parser YALMIP [12] on a computer equipped with Pentium M of 1.10 GHz and memory of 760 MBytes. The computational time was 7.25 seconds.

The region Θ has four vertices: $\theta^{(1)} = [0 \ \ 0]^{\mathrm{T}}$, $\theta^{(2)} = [1 \ \ 0]^{\mathrm{T}}$, $\theta^{(3)} = [0 \ \ 1]^{\mathrm{T}}$, and $\theta^{(4)} = [1 \ \ 1]^{\mathrm{T}}$. Among them, $\theta^{(1)}$ and $\theta^{(2)}$ are interesting because the obtained dual variables corresponding to these vertices are found nonzero. With the dual variable corresponding to $\theta^{(1)}$, we computed the worst-case parameter $\widehat{\theta}$ as in Section 5. The result was $\widehat{\theta} = [0.694205 \quad 0]^{\mathrm{T}}$. With this parameter, the system has the L^2-induced norm equal to 1.20687, which shows the accuracy of our approximate optimal value. Moreover, the obtained dual variables corresponding to $\theta^{(1)}$ and $\theta^{(2)}$ coincide with $(a^{(1)}/2)M(\widehat{\theta})\widehat{Z}M(\widehat{\theta})^{\mathrm{T}}$ and $(a^{(2)}/2)M(\widehat{\theta})\widehat{Z}M(\widehat{\theta})^{\mathrm{T}}$, respectively, up to the order of 10^{-8} for $a^{(1)} = 1 - 0.694205$, $a^{(2)} = 0.694205$ and appropriately chosen $\widehat{Z}$. This verifies the exactness of our approximation. We do not need to increase the order of $R(\theta)$ or divide Θ any more.

It is interesting that an exact result is obtained with a low-degree $R(\theta)$ and a coarse division. This is consistent with the reported performance of other approaches based on positive polynomials.

7. A robust SDP problem with function derivatives

In this section, we consider a robust SDP problem not only with a function but also with its derivatives:

$$\begin{aligned} P_{\mathrm{d}}: \quad & \text{minimize} \quad c^{\mathrm{T}} x \\ & \text{subject to} \quad E(x) \succeq O, \\ & \qquad \mathcal{F}\Big(x, \phi(\theta), \frac{\partial \phi(\theta)}{\partial \theta_1}, \frac{\partial \phi(\theta)}{\partial \theta_2}, \dots, \frac{\partial \phi(\theta)}{\partial \theta_p}; \theta\Big) \succeq O \quad (\forall \theta \in \Theta). \end{aligned}$$

Here, the optimization variables are a vector $x \in \mathbb{R}^n$ and a function $\phi \in \Phi_{\mathrm{d}}$. The function set Φ_{d} is a linear space of continuous and piecewise continuously differentiable functions defined on Θ. The function $\mathcal{F}$ is affine in x, $\phi(\theta)$, $(\partial/\partial\theta_1)\phi(\theta)$, $(\partial/\partial\theta_2)\phi(\theta)$, $\dots$, $(\partial/\partial\theta_p)\phi(\theta)$ while polynomial in θ. Other setting such as the size of $E(x)$ is the same as in P.

Various control problems are formulated into P_{d}.

Suppose in Example 1 that the parameter θ varies with time in the set Θ and its time derivative $\dot{\theta}$ belongs to the set $\dot{\Theta}$. Assume $0 \in \dot{\Theta}$. The L^2-induced norm of the system can be computed with an optimization problem similar to the one in Example 1 [27, 8, 28, 29]. The difference is that the $(1,1)$-block of the constraint is replaced by

$$\sum_{i=1}^{p} \frac{\partial R(\theta)}{\partial \theta_i} \dot{\theta}_i - A(\theta) R(\theta) - R(\theta) A(\theta)^{\mathrm{T}}$$

with $R(\theta)$ being a continuous and piecewise continuously differentiable function. Moreover, the parameter of the problem is $(\theta, \dot{\theta}) \in \Theta \times \dot{\Theta}$ here. This problem is in the form of P_{d}.

Another example is stability analysis of a nonlinear system.

Example 12. Consider a nonlinear system $\dot{\xi}(t) = f(\xi(t))$ with a polynomial $f(\xi)$ satisfying $f(0) = 0$. Then, the origin is asymptotically stable with this system if the optimum value of the following problem is negative:

$$\begin{aligned} & \text{minimize} \quad x \\ & \text{subject to} \quad \phi(0) = 0, \qquad 1 - \phi(\xi_0) \geq 0, \\ & \qquad \phi(\xi) + x\|\xi\|^2 \geq 0, \quad x\|\xi\|^2 - \Big(\frac{\partial \phi(\xi)}{\partial \xi}\Big)^{\mathrm{T}} f(\xi) \geq 0 \qquad (\forall \xi \in \Xi). \end{aligned}$$

Here, Ξ is some closed polytope containing the origin in its interior; $\xi_0 \neq 0$ is a point in Ξ. The optimization variables are a scaler x and a Lyapunov function candidate ϕ, which is continuous and piecewise continuously differentiable. This problem is again in the form of P_{d}. $\square$

We consider an approximate approach to the problem P_{d}.

First, a general function $\phi(\theta)$ is replaced by a polynomial $\phi_u(\theta) = \sum_{\alpha \in S} u_\alpha \theta^\alpha$ characterized by $u = (u_\alpha) \in \mathbb{R}^{n_u}$. The notation

$$F(x, u; \theta) := \mathcal{F}\Big(x, \phi_u(\theta), \frac{\partial \phi_u(\theta)}{\partial \theta_1}, \frac{\partial \phi_u(\theta)}{\partial \theta_2}, \ldots, \frac{\partial \phi_u(\theta)}{\partial \theta_p}; \theta\Big)$$

is used for explicit representation of the finite dimensionality. Here, F is affine in x and u while polynomial in θ.

Next, a division $\Delta = \{\Theta^{[j]}\}_{j=1}^J$ of the parameter set Θ is prepared and a polynomial $\phi_{u^{[j]}}(\theta)$ is considered for each subregion $\Theta^{[j]}$. Recall that the function $\phi(\theta)$ has to be continuous over the whole set Θ. This requirement is expressed by a linear equality constraint on $u^{[1]}, u^{[2]}, \ldots, u^{[J]}$, which is denoted by

$$d(u^{[1]}, u^{[2]}, \ldots, u^{[J]}) = 0.$$

Finally, the matrices $G(x, u)$, $M(\theta)$, $H(\theta)$ are defined in the same way as in P. Then, we have an approximate problem:

$$\begin{aligned} P_{\mathrm{d}}(\Delta): \quad & \text{minimize} && c^{\mathrm{T}} x \\ & \text{subject to} && E(x) \succeq O, \quad G(x, u^{[j]}) + H(\theta)(W^{[j]})^{\mathrm{T}} + W^{[j]} H(\theta)^{\mathrm{T}} \succeq O \\ &&& (\forall \theta \in \operatorname{ver} \Theta^{[j]}, \forall j = 1, 2, \ldots, J), \\ &&& d(u^{[1]}, u^{[2]}, \ldots, u^{[J]}) = 0, \end{aligned}$$

which can be solved with the standard interior-point method.

It is natural to expect an upper bound on the approximation error $|\inf P_{\mathrm{d}}(\Delta) - \inf P_{\mathrm{d}}|$. For the present, such a bound is obtained in a rather special setting: the parameter set Θ is a multi-dimensional interval $\prod_{i=1}^p [\underline{\theta}_i, \overline{\theta}_i]$, its division Δ is a Cartesian division, and the polynomial $\phi_u(\theta)$ is cubic in each θ_i. This is due to the requirement of continuity and the technical difficulty to bound $u^{[j]}$ in the limit of $\overline{\operatorname{rad}} \, \Delta \to 0$.

A Cartesian division of $\Theta = \prod_{i=1}^p [\underline{\theta}_i, \overline{\theta}_i]$ is defined as follows. For each $i = 1, 2, \ldots, p$, divide the interval $[\underline{\theta}_i, \overline{\theta}_i]$ as $\underline{\theta}_i = \theta_i^0 < \theta_i^1 < \cdots < \theta_i^{Q_i - 1} < \theta_i^{Q_i} = \overline{\theta}_i$. Then, the collection of all the multi-dimensional intervals $\prod_{i=1}^p [\theta_i^{q_i}, \theta_i^{q_i+1}]$ $(q_i = 0, 1, \ldots, Q_i - 1,\ i = 1, 2, \ldots, p)$ is a Cartesian division of Θ.

When the polynomial $\phi_u(\theta)$ is cubic in each θ_i $(i = 1, 2, \ldots, p)$, the dimension n_u is equal to $4^p n_\phi$. Let us use this $\phi_u(\theta)$ to interpolate a sufficiently smooth function $\phi(\theta)$ on Θ. Pick up one subregion $\Theta^{[j]}$ from a Cartesian division Δ. Then, we can choose $\phi_{u^{[j]}}(\theta)$ so as to interpolate the function value $\phi(\theta)$ as well as the first derivatives $(\partial/\partial\theta_1)\phi(\theta)$, $(\partial/\partial\theta_2)\phi(\theta)$, …, $(\partial/\partial\theta_p)\phi(\theta)$ at all the 2^p vertices of $\Theta^{[j]}$. Since the required dimension is $2^p(p+1)n_\phi$, the choice of $u^{[j]}$ is made from the space of the dimension $4^p n_\phi - 2^p(p+1)n_\phi$. Whatever the particular choice is, if each $u^{[j]}$ is chosen as such the continuity condition $d(u^{[1]}, u^{[2]}, \ldots, u^{[J]}) = 0$ is automatically satisfied. Moreover, it is possible to bound $u^{[j]}$ irrespectively of the maximum radius $\overline{\mathrm{rad}}\,\Delta$ by exploiting the freedom in the choice of $u^{[j]}$. More discussion can be found in Chapters 14–16 of [7].

We make the following two assumptions to obtain the bound. The first assumption is a natural generalization of Assumption 3 (a). Although the second assumption is a generalization of Assumption 3 (b), it may look restrictive due to the bounds on the high-order derivatives of $\phi(t)$. Note, however, that it is satisfied in a practically reasonable situation that the problem P_{d} and its modification $P_{\mathrm{d},\epsilon}$, to be introduced, are optimized by some x and smooth enough $\phi(\theta)$.

Assumption 13.

(a) There exist $x \in \mathbb{R}^n$ and $\phi \in \Phi_{\mathrm{d}}$ such that $E(x) \succ O$ and $\mathcal{F}(x, \phi(\theta), (\partial/\partial\theta_1)\phi(\theta), (\partial/\partial\theta_2)\phi(\theta), \ldots, (\partial/\partial\theta_p)\phi(\theta); \theta) \succ O$ $(\forall\theta \in \Theta)$.

(b) There exist positive numbers $\bar{\epsilon}$, $\bar{x}$, and $\bar{\phi}_\alpha$, where $\alpha = [\alpha_1 \;\; \alpha_2 \;\; \cdots \;\; \alpha_p]^{\mathrm{T}}$ and $\alpha_i = 0, 1, 2, 3$ for each i, having the following property: For any $0 \le \epsilon \le \bar{\epsilon}$ and any $v \in \mathbb{R}$, the set $\{(x, \phi) \in \mathbb{R}^n \times \Phi_{\mathrm{d}} \mid c^{\mathrm{T}}x \le v,\; E(x) \succeq O,\; \mathcal{F}(x, \phi(\theta), (\partial/\partial\theta_1)\phi(\theta), (\partial/\partial\theta_2)\phi(\theta), \ldots, (\partial/\partial\theta_p)\phi(\theta); \theta) \succeq \epsilon I \;(\forall\theta \in \Theta)\}$ is either empty or having an element (x, ϕ) with $\|x\| \le \bar{x}$ and $\|(\partial/\partial\theta_1)^{\alpha_1}(\partial/\partial\theta_2)^{\alpha_2} \cdots (\partial/\partial\theta_p)^{\alpha_p}\phi(\theta)\| \le \bar{\phi}_\alpha$ for any $\theta \in \Theta$ and $\alpha_i = 0, 1, 2, 3$ $(i = 1, 2, \ldots, p)$. □

With this setting, we have the following bound on the approximation error of $P_{\mathrm{d}}(\Delta)$.

Theorem 14. *Let Θ be a multi-dimensional interval and Δ be its Cartesian division. Let also $\phi_u(\theta)$ be cubic in each θ_i. Then, under Assumption 13, there exist positive numbers C_{d} and C'_{d}, which are independent of the division Δ, such that*

$$|\inf P_{\mathrm{d}}(\Delta) - \inf P_{\mathrm{d}}| \le C_{\mathrm{d}}\,\overline{\mathrm{rad}}\,\Delta$$

for any Δ satisfying $\overline{\mathrm{rad}}\,\Delta \le C'_{\mathrm{d}}$.

The implication of this theorem is parallel to that of Theorem 4. This theorem immediately implies the asymptotic exactness of $P_{\mathrm{d}}(\Delta)$, that is, the approximation error $|\inf P_{\mathrm{d}}(\Delta) - \inf P_{\mathrm{d}}|$ converges to zero as the maximum radius $\overline{\mathrm{rad}}\,\Delta$ of the division goes to zero. The convergence is at least proportional to $\overline{\mathrm{rad}}\,\Delta$. Exactness verification is possible in a similar fashion to Section 5.

Adaptive division as in Section 4 is not possible in the present setting because the division is restricted to be Cartesian. Improvement on this point is a future research subject. In the literature [4, 16, 9, 13, 23, 11, 6], a low-order polynomial is used for $\phi_u(\theta)$ in place of the present cubic function. An upper bound in this setting is another interesting subject.

The proof of Theorem 14 is similar to that of Theorem 4. We show only a sketch of the proof for avoidance of repetition.

Sketch of the proof of Theorem 14. We first consider the following auxiliary problem:

$$\begin{aligned} P_{\mathrm{d},\epsilon}: \quad & \text{minimize} && c^{\mathrm{T}}x \\ & \text{subject to} && E(x) \succeq O, \\ &&& \mathcal{F}\Big(x, \phi(\theta), \frac{\partial \phi(\theta)}{\partial \theta_1}, \frac{\partial \phi(\theta)}{\partial \theta_2}, \ldots, \frac{\partial \phi(\theta)}{\partial \theta_p}; \theta\Big) \succeq \epsilon I \quad (\forall \theta \in \Theta) \end{aligned}$$

for $\epsilon \geq 0$.

The critical step of the proof is to show $\inf P_{\mathrm{d}}(\Delta) \leq \inf P_{\mathrm{d},\epsilon}$ for Δ with small $\overline{\mathrm{rad}}\,\Delta$ and for consistently small $\epsilon > 0$. To this aim, we construct a solution $(x, \{u^{[j]}\}, \{W^{[j]}\})$ of $P_{\mathrm{d}}(\Delta)$ from a solution (x, ϕ) of $P_{\mathrm{d},\epsilon}$.

Suppose $\Theta \subseteq [-1, 1]^p$ first. If $\epsilon > 0$ is small enough, the auxiliary problem $P_{\mathrm{d},\epsilon}$ is feasible due to Assumption 13 (a). Moreover, a feasible solution (x, ϕ) can be chosen so that $\|x\| \leq \overline{x}$ and $\|(\partial/\partial\theta_1)^{\alpha_1}(\partial/\partial\theta_2)^{\alpha_2} \cdots (\partial/\partial\theta_p)^{\alpha_p}\phi(\theta)\| \leq \overline{\phi}_\alpha$, for any $\theta \in \Theta$ and $\alpha_i = 0, 1, 2, 3$ owing to Assumption 13 (b). Now, we choose a cubic function $\phi_{u^{[j]}}(\theta)$ for each $j = 1, 2, \ldots, J$ so as to interpolate $\phi(\theta)$ and its first derivatives at the vertices of $\Theta^{[j]}$. We also choose $W^{[j]}$ as $(1/2\,\mathrm{rad}\,\Theta^{[j]})H(\theta^0)$ for each j, where θ^0 is the vertex of $\Theta^{[j]}$ with the smallest coordinate in each axis. Then, if the maximum radius of Δ satisfies some linear relationship with ϵ, this $(x, \{u^{[j]}\}, \{W^{[j]}\})$ forms a feasible solution of the approximate problem $P_{\mathrm{d}}(\Delta)$. This implies $\inf P_{\mathrm{d}}(\Delta) \leq \inf P_{\mathrm{d},\epsilon}$.

The extension to the case of $\Theta \not\subseteq [-1, 1]^p$ is as in Section 4.2. Finally, the convexity of $\inf P_{\mathrm{d},\epsilon}$ gives the desired bound. $\square$

8. Conclusion

In this paper, we considered a robust SDP problem with a functional variable, which covers many control problems with a parameter-dependent Lyapunov function. In particular, we present an approximate approach to solve this problem and show its asymptotic exactness. A quantitative error bound is given in terms of the resolution of the division. Verification of exactness is also discussed. A numerical example shows the possibility of exact optimization through a finite-dimensional approximate problem.

Acknowledgment

Part of this work is supported by a Grant-in-Aid of the Ministry of Education, Culture, Sports, Science and Technology of Japan and the 2007 Pache Research Fund I-A-2 of Nanzan University, Japan.

References

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

[2] P.-A. Bliman, "An existence result for polynomial solutions of parameter-dependent LMIs," *Systems & Control Letters*, vol. 51, nos. 3–4, pp. 165–169, 2004.

[3] 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.

[4] R. Brayton and C. Tong, "Stability of dynamical systems: a constructive approach," *IEEE Transactions on Circuits and Systems*, vol. 26, no. 4, pp. 224–234, 1979.

[5] G. Chesi, A. Garulli, A. Tesi, and A. Vicino, "Polynomially parameter-dependent Lyapunov functions for robust stability of polytopic systems: an LMI approach," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 365–370, 2005.

[6] D. F. Coutinho and P. Danès, "Piecewise Lyapunov functions to the stability analysis of rational systems subject to multiple state constraints," in *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, USA, December 2006, pp. 5801–5806.

[7] G. Farin, *Curves and Surfaces for CAGD: A Practical Guide*, 5th ed. San Francisco, USA: Morgan Kaufmann, 2002.

[8] P. Gahinet, P. Apkarian, and M. Chilali, "Affine parameter-dependent Lyapunov functions and real parametric uncertainty," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp. 436–442, 1996.

[9] M. Johansson and A. Rantzer, "On the computation of piecewise quadratic Lyapunov functions," in *Proceedings of the 36th IEEE Conference on Decision and Control*, San Diego, USA, December 1997, pp. 3515–3520.

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

[11] S. Lim and J. P. How, "Analysis of linear parameter-varying systems using a non-smooth dissipative systems framework," *International Journal of Robust and Nonlinear Control*, vol. 12, no. 12, pp. 1067–1092, 2002.

[12] 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

[13] I. Masubuchi, A. Kume, and E. Shimemura, "Spline-type solution to parameter-dependent LMIs," in *Proceedings of the 37th IEEE Conference on Decision and Control*, Tampa, USA, December 1998, pp. 1753–1758.

[14] 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, December 1999, pp. 2341–2346.

[15] 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.

[16] Y. Ohta, H. Imanishi, L. Gong, and H. Haneda, "Computer generated Lyapunov functions for a class of nonlinear systems," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 40, no. 5, pp. 343–354, 1993.

[17] Y. Oishi, "A matrix-dilation approach to robust semidefinite programming and its error bound," in *Proceedings of the 2006 American Control Conference*, Minneapolis, USA, June 2006, pp. 123–129.

[18] 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, July 2007, pp. 6169–6176.

[19] R. C. L. F. Oliveira and P. L. D. Peres, "LMI conditions for robust stability analysis based on polynomially parameter-dependent Lyapunov functions," *Systems & Control Letters*, vol. 55, no. 1, pp. 52–61, 2006.

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

[21] 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.

[22] M. Sato and D. Peaucelle, "Robust stability/performance analysis for linear time-invariant polynomially parameter-dependent systems using polynomially parameter-dependent Lyapunov functions," in *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, USA, December 2006, pp. 5807–5813.

[23] C. W. Scherer, "Robust performance analysis for parameter dependent systems using tensor product splines," in *Proceedings of the 37th IEEE Conference on Decision and Control*, Tampa, USA, December 1998, pp. 2216–2221.

[24] 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.

[25] 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.

[26] 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.

[27] R. Watanabe, K. Uchida, M. Fujita, and E. Shimemura, "L^2 gain and H^∞ control of linear systems with scheduling parameter," in *Proceedings of the 33rd IEEE Conference on Decision and Control*, Lake Buena Vista, USA, December 1994, pp. 1412–1414.

[28] F. Wu, X. H. Yang, A. Packard, and G. Becker, "Induced L^2-norm control for LPV systems with bounded parameter variation rates," *International Journal of Robust and Nonlinear Control*, vol. 6, nos. 9–10, pp. 983–998, 1996.

[29] J. Yu and A. Sideris, "H^∞ control with parametric Lyapunov functions," *Systems & Control Letters*, vol. 30, nos. 2–3, pp. 57–69, 1997.